

07.2026

διαΝΕΟσις

ΟΡΓΑΝΙΣΜΟΣ ΕΡΕΥΝΑΣ & ΑΝΑΛΥΣΗΣ

[Πώς] μπορεί η Τεχνητή Νοημοσύνη να είναι ηθική και σύννομη;

Λίλιαν Μήτρου

Γιάννης Ασσαέλ, Ηρακλής Βογιατζής, Δημήτρης Γραμμένος,
Σέργιος Θεοδωρίδης, Ιρένε Καμάρα, Αντώνης Κουρουτάκης,
Μιχάλης Κρητικός, Δημήτρης-Φραγκίσκος Λέκκας,
Βαγγέλης Παπακωνσταντίνου, Παναγιώτης Περάκης,
Κλειώ Σγουροπούλου, Χαράλαμπος Τσέκερης,
Σταυρούλα Τσινόρεμα, Κατερίνα Χρυσανθοπούλου

Ιούλιος 2026

Ομάδα Έργου

Συντονισμός

Λίλιαν Μήτρου

Καθηγήτρια Πανεπιστημίου Αιγαίου, Δικηγόρος, Αντιπρόεδρος της Εθνικής Επιτροπής Βιοηθικής και Τεχνοηθικής (ΕΕΒΤ)

Συντελεστές

Γιάννης Ασσαέλ

Ερευνητής Επιστήμονας Τεχνητής Νοημοσύνης, Εθνικός Εκπρόσωπος στο Global Partnership on Artificial intelligence του ΟΟΣΑ, Ειδικός Σύμβουλος Τεχνητής Νοημοσύνης του Υπουργείου Ψηφιακής Διακυβέρνησης και Τεχνητής Νοημοσύνης

Ηρακλής Βογιατζής

Μεταδιδακτορικός ερευνητής στο ΕΑΠ, Συνιδρυτής της Visible Machines

Δημήτρης Γραμμένος

Κύριος Ερευνητής του Ινστιτούτου Πληροφορικής του Ιδρύματος Τεχνολογίας και Έρευνας (ΙΤΕ)

Σέργιος Θεοδωρίδης

Ομότιμος Καθηγητής ΕΚΠΑ-ΗΡΩΝ Κέντρο Αριστείας στη Ρομποτική και ΤΝ, ΑΘΗΝΑ ΕΚ

Ιρένε Καμάρα

Αναπληρώτρια Καθηγήτρια Δικαίου Κυβερνοασφάλειας και Κυβερνοεγκλήματος, Νομική Σχολή Πανεπιστημίου Τίλμπουργκ, Ολλανδία

Αντώνης Κουρουτάκης

Αναπληρωτής Καθηγητής της Νομικής Σχολής του Πανεπιστημίου ΙΕ, Μαδρίτη, Ισπανία

Μιχάλης Κρητικός

PhD-Secretary of the European Group on Ethics in Science and New Technologies (EGE), European Commission

Δημήτρης-Φραγκίσκος Λέκκας

Καθηγητής Πανεπιστημίου Αιγαίου-Τμήμα Περιβάλλοντος

Βαγγέλης Παπακωνσταντίνου

Καθηγητής Νομικής, Vrije Universiteit Brussel

Παναγιώτης Περάκης

Δικηγόρος, Πρόεδρος του Παρατηρητηρίου Δικαιοσύνης

Κλειώ Σγουροπούλου

Καθηγήτρια Τμήματος Μηχανικών Πληροφορικής και Υπολογιστών, Μέλος του Συμβουλίου Διοίκησης Πανεπιστημίου Δυτικής Αττικής

Χαράλαμπος Τσέκερης

Διευθυντής Ερευνών, Εθνικό Κέντρο Κοινωνικών Ερευνών (ΕΚΚΕ), Πρόεδρος της Εθνικής Επιτροπής Βιοηθικής & Τεχνοηθικής (ΕΕΒΤ)

Σταυρούλα Τσινόρεμα

Καθηγήτρια Φιλοσοφίας και Βιοηθικής-Πανεπιστήμιο Κρήτης, Μέλος της Εθνικής Επιτροπής Βιοηθικής και Τεχνοηθικής (ΕΕΒΤ)

Κατερίνα Χρυσανθοπούλου

Δρ Παιδείας στα ΜΜΕ και Ψηφιακών Γραμματισμών

Η παρούσα μελέτη αντανακλά την πληροφόρηση που ήταν διαθέσιμη έως και τις 30 Ιουνίου 2026.

Περιεχόμενα

1	Τι είναι η Τεχνητή Νοημοσύνη, γιατί και πώς συζητάμε για αυτήν	7
1.1.	Τι είναι η Τεχνητή Νοημοσύνη;	8
1.2.	Γιατί συζητάμε για την Τεχνητή Νοημοσύνη;	13
1.3.	Ποιοι είναι οι κίνδυνοι από την Τεχνητή Νοημοσύνη;	16
1.4.	Αποτελεί η Τεχνητή Νοημοσύνη απειλή για την υπαρξιακή μας υπόσταση;	20
2	Η πρωτοβουλία για την ηθική πλαισίωση της Τεχνητής Νοημοσύνης	32
2.1.	Η αξιόπιστη Τεχνητή Νοημοσύνη	33
2.2.	Αρχές της αξιόπιστης Τεχνητής Νοημοσύνης	34
2.3.	Οι απαιτήσεις σε σχέση με την αξιόπιστη Τεχνητή Νοημοσύνη	36
2.4.	Η Σύσταση της UNESCO για την Ηθική της Τεχνητής Νοημοσύνης	39
2.5.	Κανονιστικό πλαίσιο και όχι μόνο αρχές και κατευθύνσεις ηθικής	40
2.6.	Η Ηθική στην Τεχνητή Νοημοσύνη	43
3	Το ενωσιακό και το εθνικό κανονιστικό πλαίσιο της Τεχνητής Νοημοσύνης	55
3.1.	Η Πράξη για την Τεχνητή Νοημοσύνη	57
3.2.	Η Πράξη για την Τεχνητή Νοημοσύνη και ο διεθνής κατακερματισμός των ρυθμιστικών προσεγγίσεων	60
3.3.	Πώς η Ευρώπη ρυθμίζει τις ψηφιακές τεχνολογίες (τρίτων) και πώς αυτό επηρεάζει την ψηφιακή της κυριαρχία	62
3.4.	Το όραμα της ψηφιακής κυριαρχίας της Ευρωπαϊκής Ένωσης	71
3.5.	Digital Omnibus: Απλοποίηση ή απορρύθμιση;	73
3.6.	Η ρύθμιση ως άξονας μιας κοινωνικά βιώσιμης καινοτομίας	77
4	Τα θεμελιώδη δικαιώματα, οι δημοκρατικές διαδικασίες και η ρύθμιση της Τεχνητής Νοημοσύνης	83
4.1.	Η Τεχνητή Νοημοσύνη και τα θεμελιώδη δικαιώματα	85
4.2.	Η Σύμβαση-Πλαίσιο του Συμβουλίου της Ευρώπης για την Τεχνητή Νοημοσύνη	89
4.3.	Η Τεχνητή Νοημοσύνη και οι δημοκρατικές διαδικασίες	91
4.4.	Τεχνητή Νοημοσύνη, ενημέρωση και παραπληροφόρηση	93

5	Οι προϋποθέσεις εφαρμογής και αποτελεσματικότητας των νομικών κανόνων	98
5.1.	Ο μηχανισμός εφαρμογής της Πράξης για την Τεχνητή Νοημοσύνη	99
5.2.	Η διακυβέρνηση της Τεχνητής Νοημοσύνης.....	101
5.3.	Πρότυπα και προδιαγραφές.....	103
5.4.	Η τεχνική τυποποίηση και πιστοποίηση ως πυλώνες για αξιόπιστα συστήματα Τεχνητής Νοημοσύνης	105
5.5.	Ο ψηφιακός γραμματισμός.....	112
5.6.	Ο γραμματισμός στην Τεχνητή Νοημοσύνη στην Ελλάδα	115
6	Τεχνητή Νοημοσύνη, κράτος δικαίου και διοίκηση	125
6.1.	Η συνεισφορά της Τεχνητής Νοημοσύνης στην ανάπτυξη μιας αποτελεσματικής διοίκησης.....	126
6.2.	Μπορεί η Τεχνητή Νοημοσύνη να εφαρμόσει τον νόμο;	127
6.3.	Η Τεχνητή Νοημοσύνη και οι απαιτήσεις του κράτους δικαίου	129
6.4.	Τα ζητήματα της διαφάνειας και της αιτιολογίας	132
6.5.	Ο «ανθρώπινος παράγοντας» και η «μεροληψία αυτοματισμού»	136
6.6.	Ζητήματα από τη χρήση Παραγωγικής Τεχνητής Νοημοσύνης από τη/στη διοίκηση	139
7	Για μια ηθική Τεχνητή Νοημοσύνη στο πεδίο της Δικαιοσύνης	141
7.1.	Σημείο εκκίνησης-Εισαγωγική αναγκαία επισήμανση	142
7.2.	Οι βασικές αρχές για μια ηθική Τεχνητή Νοημοσύνη στο πεδίο της Δικαιοσύνης.....	144
7.3.	Η χρήση Τεχνητής Νοημοσύνης στην Ποινική Δικαιοσύνη.....	151
7.4.	Το ζήτημα της υποχρεωτικής αποκάλυψης της χρήσης Τεχνητής Νοημοσύνης σε δικαστήριο	153
7.5.	Συμπέρασμα.....	155
8	Οι επιπτώσεις της Τεχνητής Νοημοσύνης στον κόσμο της εργασίας και τη διαμόρφωση των εργασιακών σχέσεων	160
8.1.	Το «τέλος της εργασίας» ή νέες μορφές ψηφιακής εργασίας	162
8.2.	Ευκαιρίες και κίνδυνοι της χρήσης Τεχνητής Νοημοσύνης στην αγορά εργασίας	173
8.3.	Από την αυτοματοποίηση στην επαύξηση: Η Τεχνητή Νοημοσύνη ως καταλύτης δομικού μετασχηματισμού της εργασίας και των δεξιοτήτων στην Ελλάδα.....	175
8.4.	Οι επιπτώσεις της χρήσης Τεχνητής Νοημοσύνης στα δικαιώματα των εργαζομένων	188
8.5.	Η αλγοριθμική διοίκηση (algorithmic management) και οι εργασιακές σχέσεις.....	190
8.6.	Το κανονιστικό πλαίσιο για τη χρήση Τεχνητής Νοημοσύνης στο πλαίσιο των εργασιακών σχέσεων	193

9	Τεχνητή Νοημοσύνη και έρευνα	196
9.1.	Η Τεχνητή Νοημοσύνη και η έρευνα	197
9.2.	Η ποιότητα και η ακεραιότητα της έρευνας	199
9.3.	Θέτοντας τις βάσεις για μια νέα εθνική πολιτική στην έρευνα και την καινοτομία της Τεχνητής Νοημοσύνης	202
10	Τεχνητή Νοημοσύνη, δημιουργία και πνευματική ιδιοκτησία	208
10.1.	Musa ex machina: Τεχνητή Νοημοσύνη και ανθρώπινη δημιουργικότητα.....	210
10.2.	Τεχνητή Νοημοσύνη και ζητήματα πνευματικής ιδιοκτησίας	222
11	Ο ρόλος και η ευθύνη της Πολιτείας	225
11.1.	Ποιες είναι οι υποχρεώσεις του κράτους στην εποχή της Τεχνητής Νοημοσύνης;.....	226
11.2.	Τεχνητή Νοημοσύνη και Συνταγματική Αναθεώρηση	228
11.3.	Ψηφιακός συνταγματισμός ή/και ένα νέο «κοινωνικό συμβόλαιο»;.....	232
11.4.	Η ρύθμιση ως έκφραση της ψηφιακής κυριαρχίας: Η Τεχνητή Νοημοσύνη, το «κοινωνικό συμβόλαιο» και η εμφάνιση του ψηφιακού συνταγματισμού	236
12	Η μεγάλη εικόνα	250
12.1.	Τεχνητή Νοημοσύνη: Διακυβέρνηση, ασφάλεια και κοινωνική συμπερίληψη	251
12.2.	Καινοτομία με κάθε κόστος; Και η κλιματική κρίση;	257
12.3.	Τεχνητή Νοημοσύνη και Κυκλική Οικονομία: Ψηφιακός μετασχηματισμός της περιβαλλοντικής διακυβέρνησης στην Ευρώπη και τη Μεσόγειο	259
12.4.	Η Τεχνητή Νοημοσύνη πήρε το όπλο της;	263
12.5.	Είναι θέμα σχεδιασμού;	265
12.6.	Κάποιες τελικές σκέψεις.....	267

[ΠΩΣ] ΜΠΟΡΕΙ Η ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΝΑ ΕΙΝΑΙ
ΗΘΙΚΗ ΚΑΙ ΣΥΝΝΟΜΗ;
Ιούλιος 2026

Τι είναι η Τεχνητή Νοημοσύνη, γιατί και πώς συζητάμε για αυτήν



Σέργιος Θεοδωρίδης, Λίλιαν Μήτρου

Συγγραφική συνεισφορά ανά υποκεφάλαιο:

1.1., 1.2., 1.3. – Λίλιαν Μήτρου

1.4. – Σέργιος Θεοδωρίδης

1.1. Τι είναι η Τεχνητή Νοημοσύνη;

Την τελευταία πενταετία η Τεχνητή Νοημοσύνη (εφεξής και ΤΝ) βρίσκεται στο επίκεντρο του δημόσιου ενδιαφέροντος. Φαίνεται πως δεν πρόκειται για μια πρόσκαιρη «άνθηση», μιας και έχει βιώσει «χειμώνες και καλοκαίρια», αλλά πως η εβδομηκοντούτις τεχνολογία ήλθε για να μείνει, αν και δεν λείπουν και κάποιοι που προειδοποιούν για τη «φούσκα» της ΤΝ. Χωρίς αμφιβολία, το «νέο» αυτό τεχνολογικό κύμα επιδρά και θα συνεχίσει να επιδρά καθοριστικά στο παρόν και συνακόλουθα στο μέλλον μας, επηρεάζοντας την κοινωνία και επαναδιαμορφώνοντας μοντέλα ζωής, εργασίας, μάθησης, κοινωνικής και πολιτικής αλληλεπίδρασης.

Τι είναι, όμως, η Τεχνητή Νοημοσύνη; Η ΤΝ δεν είναι απλώς μια «τεχνολογία γενικού σκοπού», που –όπως ο ατμός ή ο ηλεκτρισμός– μπορεί να χρησιμοποιηθεί σε απεριόριστο αριθμό δραστηριοτήτων, αλλά, πολύ περισσότερο, μια τεχνολογία που φαίνεται να κατέχει τα κλειδιά για την εκθετική ανάπτυξη της ανθρώπινης ευημερίας. Η διαθεσιμότητα των δεδομένων μεγάλης κλίμακας, η αύξηση της υπολογιστικής ισχύος και η πρόοδος ως προς τους αλγόριθμους κατέστησαν την ΤΝ την πιο σημαντική τεχνολογία του 21ου αιώνα. Επί του παρόντος, οι ρυθμοί ανάπτυξης είναι τόσο ραγδαίοι που, περισσότερο από ποτέ, δεν είμαστε σε θέση να «προβλέψουμε» την περαιτέρω εξέλιξη.

Η ΤΝ προσεγγίζεται ως η τεχνολογία που αντιλαμβάνεται το περιβάλλον και προσαρμόζει τις ενέργειες κατά τρόπο ώστε να μεγιστοποιεί τις πιθανότητες επιτυχίας,¹ το «εργαλείο» που προοπτικά θα υποκαταστήσει τις ανθρώπινες δραστηριότητες.² Θα μπορούσε να παραθέσει κανείς πολλούς ορισμούς της ΤΝ. Ακριβέστερα, θα έπρεπε να μιλάμε για «πολλές Τεχνητές Νοημοσύνες». Το κοινό σημείο των πολλών ορισμών και προσεγγίσεων εντοπίζεται στον συσχετισμό ή στην ταύτιση των συστημάτων ΤΝ με ανθρώπινες ή έλλογες ιδιότητες. Υιοθετώντας μια απλή και περιεκτική προσέγγιση κι έχοντας κατά

¹ Janssen, M., Hartog, M., Matheus, R., Yi Ding, A., & Kuk, G. (2021). Will Algorithms Blind People? The Effect of Explainable AI and Decision-Makers' Experience on AI-supported Decision-Making in Government, *Social Science Computer Review*, 40(2), 478-493. <https://doi.org/10.1177/0894439320980118>

² Makridakis, S. (2017). The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms, *Futures*, 90(Supplement C), 46-60. <https://doi.org/10.1016/j.futures.2017.03.006>

νου τη Δοκιμή του Turing,³ θα μπορούσαμε να μιλήσουμε για την «τεχνολογία που επιτρέπει στους υπολογιστές να κάνουν πράγματα, τα οποία απαιτούν νοημοσύνη όταν γίνονται από ανθρώπους» (Marvin Minsky).

Θα μπορούσαμε να απαριθμήσουμε πολλούς ορισμούς. Καθώς η προσέγγισή μας δεν είναι θεωρητική, αλλά φιλοδοξεί να συνδέσει τα όποια ζητήματα θέτει η χρήση της ΤΝ, κρίνουμε σκόπιμο να παραθέσουμε και την κανονιστική αποτύπωση. Σύμφωνα με τον σχετικό ορισμό που υιοθετεί η Πράξη για την Τεχνητή Νοημοσύνη (Artificial Intelligence Act – εφεξής και AI Act)⁴ στο άρθρο 3 παρ. 1, «το Σύστημα ΤΝ συνιστά κάθε μηχανικό σύστημα που έχει σχεδιαστεί για να λειτουργεί με διαφορετικά επίπεδα αυτονομίας και μπορεί να παρουσιάζει προσαρμοστικότητα μετά την εφαρμογή του και το οποίο, για ρητούς ή σιωπηρούς στόχους, συνάγει, από τα “στοιχεία εισόδου” που λαμβάνει, πώς να παράγει “στοιχεία εξόδου”, όπως προβλέψεις, περιεχόμενο, συστάσεις ή αποφάσεις που μπορούν να επηρεάσουν υλικά ή εικονικά περιβάλλοντα». Ο ορισμός αυτός δεν συνιστά πρωτογενή επινόηση του ενωσιακού νομοθέτη, καθώς συνειδητά επιδιώχθηκε η εναρμόνιση με τον αντίστοιχο ορισμό του ΟΟΣΑ,⁵ ενώ προφανής είναι η επίδραση των ορισμών και τεχνικών προδιαγραφών που περιλαμβάνονται σε διεθνή πρότυπα ISO/IEC, όπως το ISO/IEC 22989:2022.⁶

Η έκταση της μελέτης δεν επιτρέπει μια λεπτομερή ανάλυση, αλλά θα πρέπει να αναδείξουμε κυρίως τα στοιχεία ενός συστήματος ΤΝ:

α) της αυτονομίας, δηλαδή της δυνατότητας να εκτελεί λειτουργίες ή και να παράγει συστάσεις και αποφάσεις χωρίς να υφίσταται συνεχής ή άμεση ανθρώπινη παρέμβαση,⁷

³ Θεμελιώδες πείραμα για την Τεχνητή Νοημοσύνη, που προτάθηκε το 1950 από τον κορυφαίο Βρετανό μαθηματικό και πρωτοπόρο στην ανάπτυξη της θεωρίας των υπολογιστών, Alan Turing.

⁴ Κανονισμός (ΕΕ) 2024/1689. Πρόκειται για τον Κανονισμό που τιτλοφορείται Πράξη για την Τεχνητή Νοημοσύνη.

⁵ Ο αρχικός ορισμός που είχε διατυπωθεί από τον ΟΟΣΑ το 2019 (OECD, Recommendation of the Council on Artificial Intelligence) όριζε ως σύστημα ΤΝ «κάθε σύστημα που επεξεργάζεται δεδομένα για να δώσει προβλέψεις, συστάσεις ή αποφάσεις, με κάποιο επίπεδο αυτονομίας, για καθορισμένο σκοπό». Ο ορισμός αυτός επικαιροποιήθηκε το 2024 για να προσαρμοστεί πλέον σε αυτόν της AI Act. Ο επικαιροποιημένος ορισμός έχει ως εξής: «AI system: An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployments».

⁶ International Organization for Standardization & International Electrotechnical Commission (ISO/IEC) (2022). Information technology – Artificial intelligence – Artificial intelligence concepts and terminology (ISO/IEC 22989:2022). <https://www.iso.org/standard/74296.html>

⁷ Η ύπαρξη ανθρώπινης επίβλεψης δεν αναιρεί την αυτονομία, αρκεί το σύστημα να μπορεί να ανταποκρίνεται και να λειτουργεί σε ορισμένο βαθμό ανεξάρτητα. Βλ. European Commission. (2025). Commission guidelines on the definition of an artificial intelligence system established by Regulation (EU) 2024/1689 (AI Act) (p. 4). Publications Office of the European Union.

β) της προσαρμοστικότητας, δηλαδή της ικανότητας του συστήματος να τροποποιεί ή να εξελίσσει τη συμπεριφορά του,⁸ και

γ) της συναγωγής συμπερασμάτων, δηλαδή της ικανότητας του συστήματος να εξάγει συμπεράσματα από τα δεδομένα εισόδου, καθώς η ΤΝ διαφέρει από τις κλασικές τεχνολογίες αυτοματισμού κατά το ότι δεν παράγει αποφάσεις επί τη βάση μιας προ-προγραμματισμένης λογικής «εάν-τότε» (if-then logic).⁹ Η ικανότητα αυτή επιτρέπει στα συστήματα της ΤΝ να προσαρμόζουν τη συμπεριφορά τους και να διαχειρίζονται αβεβαιότητα, πολυπλοκότητα ή νέες καταστάσεις με τρόπους που τα παραδοσιακά λογισμικά δεν μπορούν.¹⁰

Σε μεταγενέστερη φάση της διαδικασίας συζήτησης της AI Act, λόγω της μάλλον αναπάντεχης και ταχύτατης, σχεδόν εκθετικής, διάδοσης των Μεγάλων Γλωσσικών Μοντέλων (Large Language Models – LLMs), όπως αυτών της OpenAI, προστέθηκαν ορισμοί και κανόνες που αφορούσαν τα Μοντέλα Τεχνητής Νοημοσύνης Γενικού Σκοπού (General-Purpose AI Models). Ως τέτοιο μοντέλο ορίζεται, μεταξύ άλλων, ένα μοντέλο που έχει εκπαιδευτεί με μεγάλο όγκο δεδομένων χρησιμοποιώντας αυτοεποπτεία σε κλίμακα, το οποίο παρουσιάζει σημαντική γενικότητα και είναι ικανό να εκτελεί αποτελεσματικά ευρύ φάσμα διακριτών καθηκόντων, ανεξάρτητα από τον τρόπο με τον οποίο το μοντέλο διατίθεται στην αγορά και μπορεί να ενσωματωθεί σε διάφορα κατάντη συστήματα ή εφαρμογές» [άρθρο 3 (63) AI Act].¹¹

Η κατηγορία των Μοντέλων Τεχνητής Νοημοσύνης Γενικού Σκοπού περιλαμβάνει, μεταξύ άλλων, τα Μεγάλα Γλωσσικά Μοντέλα, τα οποία έχουν σχεδιαστεί και εκπαιδευτεί σε ευρείας κλίμακας δεδομένα, με σκοπό την ανάπτυξη ικανότητας κατανόησης και παραγωγής λόγου. Μπορούν να ανταποκριθούν σε πληθώρα λειτουργιών: από την απάντηση ερωτημάτων, τη σύνοψη κειμένων και τη μετάφραση, έως τη συγγραφή κώδικα

⁸ Βλ. την αιτιολογική σκέψη 12 της AI Act, στην οποία διευκρινίζεται ότι η «προσαρμοστικότητα» αναφέρεται σε δυνατότητες αυτοεκπαίδευσης, που επιτρέπουν στη συμπεριφορά του συστήματος να αλλάζει κατά τη χρήση. Η νέα συμπεριφορά του προσαρμοσμένου συστήματος μπορεί να παράγει διαφορετικά αποτελέσματα από το προηγούμενο σύστημα για τις ίδιες εισροές.

⁹ Όπου οι οδηγίες αναφορικά με το τι εισέρχεται παράγουν ακριβώς το ίδιο αποτέλεσμα. Αντίθετα, η ΤΝ χαρακτηρίζεται από αυτονομία στη λειτουργία και την εκμάθηση. Ahn, M. J., & Chen, Y. C. (2020). Artificial intelligence in government: Potentials, challenges, and the future. In S. J. Eom & J. Lee (Eds.), *dg.o '20: Proceedings of the 21st Annual International Conference on Digital Government Research* (pp. 243-252). Association for Computing Machinery.

¹⁰ Η ικανότητα συναγωγής συμπερασμάτων είναι μια σημαντική ιδιότητα αυτών των συστημάτων, ιδίως των συστημάτων μηχανικής μάθησης (machine learning), αλλά ενέχει και ιδιαίτερους κινδύνους για τα δικαιώματα των προσώπων, καθώς παρέχουν τη δυνατότητα εντοπισμού υποκείμενων μοτίβων στα δεδομένα, με αποτέλεσμα να αποκαλύπτονται ή να συνάγονται νέες προσωπικές πληροφορίες για πρόσωπα, που μπορεί να αναφέρονται σε κατηγοριοποιήσεις, προβλέψεις ή αποφάσεις με επιπτώσεις σε διάφορες πτυχές της ζωής τους. Βλ. European Parliament (2020). *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence*. European Parliamentary Research Service, καθώς και την ανάλυση του Solove, D. (2024). *Artificial intelligence and privacy*. *Florida Law Review*, 77(1), 1-73. <https://scholarship.lawufl.edu/flr/vol77/iss1/1>

¹¹ Ως «κατάντη συστήματα» μεταφράζεται ο όρος «downstream systems» της αγγλικής εκδοχής του κειμένου της Πράξης για την Τεχνητή Νοημοσύνη.

και τη δημιουργία πρωτότυπου περιεχομένου.¹² Όπως αναφέρεται στην αιτιολογική σκέψη 99 της AI Act, τα Μεγάλα Γλωσσικά Μοντέλα Παραγωγικής ΤΝ αποτελούν κλασικό παράδειγμα Μοντέλου ΤΝ Γενικού Σκοπού, δεδομένου ότι επιτρέπουν την ευέλικτη παραγωγή περιεχομένου, π.χ. με τη μορφή κειμένου, ήχου, εικόνων ή βίντεο, που μπορεί εύκολα να εξυπηρετήσει ευρύ φάσμα διακριτών εργασιών. Ο ΟΟΣΑ περιγράφει την Παραγωγική Τεχνητή Νοημοσύνη (Generative AI) ως μια μορφή ΤΝ που δημιουργεί νέο περιεχόμενο ως απόκριση σε προτροπές (prompts), συμπεριλαμβανομένων κειμένου, εικόνων, βίντεο και μουσικής.¹³

Ιδίως η Παραγωγική ΤΝ φαίνεται να είναι σε θέση να μεταμορφώνει την οικονομία, το κράτος και την κοινωνία με πολλούς τρόπους, θετικούς και αρνητικούς. Είναι, επίσης, εμφανές ότι τα τελευταία τρία-τέσσερα χρόνια, κυρίως από το σημείο καμπής της εμφάνισης του, όπως αποδείχθηκε, εξαιρετικά δημοφιλούς στο ευρύ κοινό ChatGPT, τον Νοέμβριο του 2022, πολλοί ταυτίζουν την ΤΝ με τα Μεγάλα Γλωσσικά Μοντέλα (όπως το ChatGPT, το GEMINI, το Claude, το Deep Seek). Μια ταύτιση που, όπως σημειώνουν οι Γιάννης Ιωαννίδης και Τίμος Σελλής,¹⁴ θολώνει τη συζήτηση για πολλές εφαρμογές της ΤΝ, όπως η διάγνωση ή η πρόβλεψη.

Η συνήθης «εικόνα» με την οποία το ευρύ κοινό συνδέει την ΤΝ με ένα chatbot που απαντά ερωτήσεις απέχει, όλο και περισσότερο, από την πραγματικότητα της Παραγωγικής ΤΝ, όπως αποδίδεται συνήθως η Agentic AI. Όπως παρατηρείται στην έκθεση του ΟΟΣΑ, οι όροι «AI agents» και «Agentic AI» βρίσκονται πλέον στο επίκεντρο των συζητήσεων για τα προηγμένα συστήματα ΤΝ. Η Agentic AI έχει αναδειχθεί εντονότερα μόλις τα τελευταία χρόνια, κυρίως λόγω της εμφάνισης ικανότερων Μεγάλων Γλωσσικών Μοντέλων, τα οποία μπορούν να σχεδιάζουν, να ενεργούν και να αλληλεπιδρούν με τρόπους που προσομοιάζουν σε παράγοντες (agents).¹⁵

Τα πλέον προηγμένα συστήματα ΤΝ καθίστανται ολοένα και περισσότερο παραγοντικά: Όταν τους τίθεται ένας στόχος, τον αποσυνθέτουν σε επιμέρους υποκαθήκοντα, επιλέγουν εργαλεία, εκτελούν πολυσταδιακά σχέδια και προσαρμόζονται όταν προκύπτουν εμπόδια.¹⁶ Σύμφωνα με τον ορισμό στον οποίο καταλήγει ο ΟΟΣΑ, η Παραγωγική ΤΝ αναφέρεται σε συστήματα που αποτελούνται από πολλαπλώς συντονισμένους παράγοντες ΤΝ,

¹² Τέτοια είναι το GPT-4 και το πρόσφατο GPT-5165 της OpenAI, το PaLM της Google, το LLaMA της META κ.ά.

¹³ Lorenz, P., Perset, K., & Berryhill, J. (2023). Initial policy considerations for generative artificial intelligence (p. 8 ff.). OECD Artificial Intelligence Papers, 1. OECD Publishing. <https://doi.org/10.1787/7ae2d1e6-en>

¹⁴ Βλ. τη συνέντευξη των Γ. Ιωαννίδη και Τ. Σελλή στο: Ιωαννίδης, Γ., & Σελλής, Τ. (2026). Όλα πλέον τα βαφτίζουμε τεχνητή νοημοσύνη, κάτι που θεωρώ επικίνδυνο. Στο *The Key Issue: Τεχνητή Νοημοσύνη – Στρατηγική, Οικονομία, Κοινωνία, Τεχνολογία* (σσ. 184-191). Το Βήμα.

¹⁵ Organisation for Economic Co-operation and Development (OECD) (2026). The agentic AI landscape and its conceptual foundations. OECD Artificial Intelligence Papers, 56. <https://doi.org/10.1787/396cf758-en>

¹⁶ Miller, R. (2024). What exactly is an AI agent?. TechCrunch. <https://techcrunch.com/2024/12/15/what-exactly-is-an-ai-agent>

οι οποίοι μπορούν να αναλύουν σύνθετα καθήκοντα σε επιμέρους στάδια, να συνεργάζονται και να επιδιώκουν αυτόνομα πολύπλοκους στόχους για εκτεταμένα χρονικά διαστήματα. Τα συστήματα αυτά σχεδιάζονται ώστε να λειτουργούν σε περισσότερο ανοιχτά και λιγότερο προβλέψιμα φυσικά ή εικονικά περιβάλλοντα και να επιχειρούν με ελάχιστη ανθρώπινη επίβλεψη.¹⁷

¹⁷ Organisation for Economic Co-operation and Development (OECD), ό.π., p. 20.

1.2. Γιατί συζητάμε για την Τεχνητή Νοημοσύνη;

Οι άνθρωποι πάντα «επένδυαν» στην τεχνολογία του καιρού τους για να αντιμετωπίσουν τις δυσκολίες του περιβάλλοντος και να καλυτερεύσουν τη ζωή τους. Χωρίς αμφιβολία ο –κατά τον Ιωσήφ Σηφάκη– «παροξυσμός αισιοδοξίας»¹⁸ για την Τεχνητή Νοημοσύνη συνοδεύεται από –συχνά αντίστοιχα έντονους– προβληματισμούς για τους όρους, τα όρια και τις επιπτώσεις της ανάθεσης ανθρώπινων δραστηριοτήτων σε τέτοια συστήματα. Συνδεόμενη με την προσδοκία να εκτοξεύσει τις δυνατότητες της ανθρώπινης νοημοσύνης και να διευρύνει τα όρια κατανόησης (και τιθάσευσης;) του κόσμου, η ιδέα μιας ΤΝ μη υποκείμενης στον ανθρώπινο έλεγχο προκαλούσε (και εξακολουθεί να προκαλεί) ταυτόχρονα δέος, αν όχι φόβο.

Διατυπώνοντας το αυτονόητο, ας σημειώσουμε ότι –καταρχάς– δεν είναι η τεχνολογία καθ' εαυτή που αναπτύσσει επιπτώσεις, αλλά η χρήση της – και ο τρόπος με τον οποίο επιδρά σε διαδικασίες και δομές της κοινωνίας, του κράτους, της οικονομίας. Συνεπώς, οι επιπτώσεις μιας τεχνολογίας διαφέρουν ανάλογα με τους σκοπούς που επιδιώκονται και το πλαίσιο εντός του οποίου αναπτύσσονται, χρησιμοποιούνται και επενεργούν. Ισχύει αυτό και για την ΤΝ; Παρά το γεγονός ότι η ΤΝ συνδέεται με αυτό που αποκαλείται «έλευση των μηχανών» (rise of the machines), θα πρέπει να αναγνωρίσουμε ότι οι άνθρωποι εμπλέκονται σε κάθε στάδιο της ανάπτυξης και της εφαρμογής τεχνολογιών ΤΝ.

Ωστόσο, τα συστήματα αυτά, σύνθετα και όχι αμιγώς τεχνικά, χαρακτηρίζονται από πολυπλοκότητα, δυναμισμό και τη δυναμική και συνεχή αλληλεπίδραση με το περιβάλλον μέσα στο οποίο λειτουργούν. Είναι σύνθετα στη λειτουργική λογική τους, παράγοντας αποτελέσματα, τα οποία είναι δύσκολο να προβλεφθούν, ιδίως όταν χρησιμοποιούνται αλγόριθμοι μηχανικής μάθησης. Η ΤΝ, ως βασική τεχνολογία του 21ου αιώνα, έχει τη δυνατότητα να παρεμβαίνει βαθιά στις κοινωνικές και οικονομικές σχέσεις των σύγχρονων κοινωνιών –και επομένως και στις σχέσεις εξουσίας και κυριαρχίας ειδικότερα– και να τις αλλάζει ουσιαστικά. Οι τεχνολογίες ΤΝ είναι ταυτόχρονα και διαμορφωτές συμπεριφορών, αντιλήψεων και κανόνων.

¹⁸ Σηφάκης, Ι. (2020). Κατανοώντας και αλλάζοντας τον κόσμο (σ. 121). Αρμός.

Κατά τούτο δεν πρέπει να νοείται, τουλάχιστον αποκλειστικά, ως τεχνολογική καινοτομία, ένας *terminus technicus*, αλλά πρέπει να προσεγγίζεται ως ένα καθολικής έκτασης φαινόμενο με εκτεταμένες και μη ευχερώς προδιαγνώσιμες τεχνολογικές, οικονομικές, κοινωνικές, πολιτικές και πολιτιστικές επιπτώσεις. Με τις επενδύσεις σε τεχνολογίες και υποδομές να αυξάνονται ραγδαία, τον καταϊγισμό των σχετικών ειδήσεων και τον εντεινόμενο ενθουσιασμό των κυβερνήσεων, δεν προκαλεί έκπληξη ότι η ΤΝ εκλαμβάνεται ευρέως ως αναπόφευκτη, ένα «πεπρωμένο» εξελισσόμενο σε «κανονικότητα».

Ταυτόχρονα, δεν πρέπει να μας διαφεύγει ότι τόσο σε επίπεδο δημόσιας και επιστημονικής συζήτησης όσο και στην αντίληψη του μέσου χρήστη η ΤΝ συνδέεται με προσδοκίες, αλλά και κινδύνους. Κινδύνους που αφορούν τη θέση του μέσα στην κοινωνία, τον καταμερισμό εργασίας, την ποιότητα της ζωής και των ανθρώπινων σχέσεων, τη διαμόρφωση της προσωπικότητας, το επίπεδο και την ποιότητα της πληροφόρησης, τα προσωπικά δεδομένα.¹⁹ Η δημοφιλία των συστημάτων ΤΝ ισορροπεί ανάμεσα στον τεχνολογικό ενθουσιασμό και στην τεχνοφοβία, γράφει ο Ηρακλής Βογιατζής.²⁰ Όπως αναφέρει, η ΤΝ εμφανίζει απειλητικά χαρακτηριστικά τόσο για τα άτομα που τη χρησιμοποιούν όσο και για αυτά που την αποφεύγουν, ενώ επισημαίνει ότι ο φόβος της αλλοτρίωσης ή ενός μέλλοντος όπου οι μηχανές θα διαλύσουν τους ανθρώπους είναι διάχυτος στη δημόσια σφαίρα, αντανakλώντας ένα δυτικό φαντασιακό γύρω από την κυριαρχία της μηχανής επί του ανθρώπου.

Η συζήτηση για την ΤΝ είναι ευρύτατη και διεξάγεται σε διάφορα επίπεδα: φιλοσοφικό, ηθικό, θεσμικό, κοινωνικό, πολιτισμικό. Είναι μια συζήτηση που χαρακτηρίζεται από ιδιαίτερα ενδιαφέρουσες, πολύπτυχες προσεγγίσεις, συχνές παρανοήσεις και πολύ «θόρυβος». Κινείται συχνά σε επίπεδα γενικής θεώρησης ή αφορά πολύ συγκεκριμένα ζητήματα. Η συζήτηση χαρακτηρίζεται ενίοτε από μανιαϊσμό, αιωρούμενη μεταξύ της απεριόριστης αισιοδοξίας και προσδοκίας για τις δυνατότητες της ΤΝ –ακόμη και ως μέσου υπέρβασης των βιολογικών περιορισμών του ανθρώπου μέσω της σύζευξής της με τις νευροεπιστήμες– και της απαισιοδόξης θεώρησής της ως ενός σύγχρονου κουτιού της Πανδώρας. Ιδίως ύστερα από την εκθετική αύξηση της χρήσης εφαρμογών ΤΝ από τον

¹⁹ διαΝΕΟσις (2026). Τι πιστεύουν οι Έλληνες για την Τεχνητή Νοημοσύνη Α΄ Μέρος – Έκθεση αποτελεσμάτων. https://www.dianeosis.org/wp-content/uploads/2026/05/AI_Dimoskopiki_Apotelesmata_190526.pdf και Μπαλαμπανίδης, Γ., Φαναράς, Σ., & Αποστολοπούλου, Π. (2026). Τι πιστεύουν οι Έλληνες για την Τεχνητή Νοημοσύνη Α΄ Μέρος. διαΝΕΟσις. https://www.dianeosis.org/wp-content/uploads/2026/05/AI_METRON_ANALYSIS_essay_190526.pdf. Βλ. τις παρατηρήσεις των Γιάννη Μπαλαμπανίδη, Στράτου Φαναρά και Πέννυς Αποστολοπούλου, οι οποίοι επισημαίνουν, ωστόσο, ότι «η έστω και μικρή επαφή και εξοικείωση με το “μαύρο κουτί” της ΤΝ μπορεί να οδηγήσει σε στάσεις περισσότερο δεκτικές απ’ ό,τι η απουσία επαφής με αυτό το πρωτόγνωρο εργαλείο. Όταν απουσιάζει πλήρως η επαφή με την Τεχνητή Νοημοσύνη, τότε φαίνεται να εντείνεται και η αίσθηση ότι αποτελεί κάτι το απειλητικό».

²⁰ Βογιατζής, Η. (2026). Οι Έλληνες για την Τεχνητή Νοημοσύνη: Ελπίδα και ανησυχία για ένα αβέβαιο μέλλον. διαΝΕΟσις. https://www.dianeosis.org/wp-content/uploads/2026/03/AI_VOGLATZIS_essay.pdf

μέσο χρήστη, που εναποθέτει σε Μεγάλα Γλωσσικά Μοντέλα τη διαχείριση της καθημερινότητας και της εργασίας του,²¹ μπορούμε να διαπιστώσουμε πως «το τζίνι βγήκε από το μπουκάλι».

²¹ διαΝΕΟσις (2026). Τι πιστεύουν οι Έλληνες για την Τεχνητή Νοημοσύνη Α΄ Μέρος – Έκθεση αποτελεσμάτων.
https://www.dianeosis.org/wp-content/uploads/2026/05/AI_Dimoskopiki_Apotelesmata_190526.pdf

1.3. Ποιοι είναι οι κίνδυνοι από την Τεχνητή Νοημοσύνη;

Δεν είναι συνεπώς παράδοξο να αντιλαμβάνεται κανείς την Τεχνητή Νοημοσύνη είτε ως πανάκεια είτε ως απειλή. Δεν είναι, επίσης, παράδοξο που διατυπώνονται απόψεις οι οποίες φτάνουν να εξισώνουν την ΤΝ με έναν υπαρξιακό κίνδυνο (existential risk).²² Ο φιλόσοφος Nick Bostrom ορίζει ως «υπαρξιακό κίνδυνο» εκείνον που «απειλεί να προκαλέσει την εξαφάνιση της νοήμονος ζωής που προέρχεται από τη Γη ή να καταστρέψει με άλλον τρόπο μόνιμα και δραστικά τις δυνατότητές της για μελλοντική επιθυμητή ανάπτυξη».²³ Οι αυτόνομες μηχανές²⁴ προσιοιώνονται σενάρια στα οποία διακυβεύονται το ίδιο το νόημα, αλλά και η ίδια η επιβίωση της ανθρωπότητας.²⁵ Η ιδέα ότι η ΤΝ μπορεί να ενέχει έναν τέτοιο κίνδυνο συμβαδίζει με τη φιλοδοξία κατασκευής υπερπανίσχυρων υπολογιστών, ικανών για πλήρως αυτόνομη λειτουργία, που θα οδηγούσαν σε μια «εκθετική ανάπτυξη» προς όφελος της ανθρωπότητας.²⁶

Παρά τον κίνδυνο να κινηθεί κανείς στα όρια της τεκνοφοβίας ή και του λουδιτισμού, η διάγνωση των κινδύνων είναι κρίσιμη για τη συζήτηση γύρω από τα ηθικά, θεσμικά και δικαιοπολιτικά ζητήματα που μπορεί να θέσει η χρήση της ΤΝ. Μια ενδιαφέρουσα κατηγοριοποίηση αυτών των κινδύνων

²² Ως υπαρξιακές απειλές, τα λεγόμενα existential risks διαφέρουν από αυτό που αντιλαμβανόμαστε ως κινδύνους σε μεγάλη, παγκόσμια κλίμακα, όπως η υπερθέρμανση του πλανήτη ή μεγάλες πανδημικές κρίσεις, καθώς στις περιπτώσεις αυτές, παρά την έκταση και την οξύτητα των κρίσεων και των διαφορετικών αντιλήψεων για την αξιολόγηση της σοβαρότητας και της αντιμετώπισής τους, οι κρίσεις αυτές υπό το φως της ιστορικής εξέλιξης του ανθρώπου θεωρούνται παροδικές και αντιμετωπίσιμες.

²³ Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies* (p. 115). Oxford University Press [Digital Edition].

²⁴ Ο όρος λουδιτισμός αναφέρεται στο κίνημα που αναπτύχθηκε κυρίως στην Αγγλία στις αρχές του 19ου αιώνα, όταν ομάδες εργατών αντέδρασαν στην εισαγωγή μηχανών, τις οποίες θεωρούσαν ότι απειλούσαν την εργασία, τα εισοδήματα και τον τρόπο ζωής τους. Ιστορικά, ο όρος συνδέεται ιδίως με τις επιθέσεις εναντίον μηχανών της κλωστοϋφαντουργίας κατά την περίοδο 1811-1813. Σήμερα ο όρος χρησιμοποιείται ευρύτερα, δηλώνοντας γενικά την επιφυλακτική ή αρνητική στάση απέναντι στην τεχνολογική καινοτομία. Η ερμηνεία αυτή είναι εν μέρει ανακόλουθη σε σχέση με την ιστορική πραγματικότητα, καθώς οι λουδίτες δεν ήταν απλώς «εχθροί της τεχνολογίας», αλλά αντιδρούσαν σε συγκεκριμένες κοινωνικο-οικονομικές συνέπειες της εκβιομηχάνισης. Βλ. αναλυτικότερα Navickas, K. (2005). The search for 'General Ludd': the mythology of Luddism. *Social History*, 30(3), 281-295. <https://doi.org/10.1080/03071020500185406>

²⁵ Asp, K. (2019). Autonomy of Artificial Intelligence, Ecology, and Existential Risk: A Critique. In: T. Heffernan (Ed.). *Cyborg Futures. Social and Cultural Studies of Robots and AI* (p. 63 ff.). Palgrave Macmillan.

²⁶ Βλ. χαρακτηριστικά το κείμενο του Good, I. J. (1966). Speculations Concerning the First Ultra Intelligent Machine. *Advances in Computers*, 6, 31-88. [https://doi.org/10.1016/S0065-2458\(08\)60418-0](https://doi.org/10.1016/S0065-2458(08)60418-0)

προτείνουν οι Matthew Tokson και Yonathan Arbel,²⁷ οι οποίοι διακρίνουν τους ακόλουθους κινδύνους:

α) Κίνδυνοι που προέρχονται από τον άνθρωπο (human directed risks): Οι πάροχοι, σχεδιαστές ή/και οι χρήστες τέτοιων συστημάτων τα σχεδιάζουν ή τα χρησιμοποιούν κατά τρόπο που βλάπτουν τα δικαιώματα και έννομα συμφέροντα των άλλων ή το δημόσιο συμφέρον. Οι –κοινωνικά ή ατομικά– επιβλαβείς χρήσεις μπορεί να έχουν κίνητρο το κέρδος,²⁸ την προσβολή των άλλων, τη διάπραξη εγκλημάτων, τη διάδοση συχνά ακραίων αντιλήψεων, την εμπέδωση προκαταλήψεων και διακρίσεων, με την κλίμακα του κινδύνου να αυξάνεται ευθέως ανάλογα με την ισχύ αυτών των συστημάτων και τη συγκέντρωση του ελέγχου σε κυρίως ιδιωτικούς φορείς που δεν υπόκεινται σε εποπτεία και λογοδοσία.

β) Κίνδυνοι που μπορεί να προκύψουν τυχαία ή ακούσια (ατύχημα/τυχαία) ή είναι συστημικοί κίνδυνοι (accidental and systemic risks): Πρόκειται για κινδύνους και βλάβες που δεν προϋποθέτουν κακόβουλη πρόθεση ή κατάχρηση των δυνατοτήτων της ΤΝ. Εν προκειμένω, η καταστροφική έκβαση μπορεί να οφείλεται σε ατυχήματα, αστοχία ευθυγράμμισης (misalignment) και αλυσιδωτές αποτυχίες (cascading failures). Οι κίνδυνοι αυτοί διαφέρουν θεμελιωδώς από τις απειλές που κατευθύνονται από ανθρώπινη ενέργεια, διότι δεν προκύπτουν από κακόβουλο σχεδιασμό, αλλά από την αλληλεπίδραση μεταξύ αυτόνομης λήψης αποφάσεων και περιβαλλοντικής πολυπλοκότητας. Αυτό που καθιστά τους κινδύνους από ατυχήματα ιδιαίτερα ανησυχητικούς είναι η δυνατότητά τους να καταστούν συστημικοί. Όταν τα συστήματα ΤΝ ενσωματώνονται σε βάθος σε αλληλοσυνδεόμενες υποδομές, μεμονωμένα ατυχήματα μπορούν να προκαλέσουν αλυσιδωτές επιπτώσεις που εξαπλώνονται σε πολλαπλούς τομείς ταχύτερα απ' ό,τι μπορούν οι άνθρωποι χειριστές να τις αντιληφθούν ή να αντιδράσουν.²⁹ Ο προβληματισμός σε σχέση με αυτούς τους κινδύνους ενισχύεται από τον συνδυασμό ταχύτητας και κλίμακας, που μεταβάλλουν ριζικά τα «προφίλ κινδύνου» με τρόπους που οι παραδοσιακοί μηχανισμοί ασφάλειας δεν μπορούν να αντιμετωπίσουν.

²⁷ Tokson, M., & Arbel, Y. (2026). Artificial Intelligence and Existential Risk. *Connecticut Law Review*, 59 (forthcoming 2027), 1-60. <https://doi.org/10.2139/ssrn.6288138>

²⁸ Το χαρακτηριστικό παράδειγμα είναι η εντατική ανάλυση δεδομένων και η κατάρτιση προφίλ για διαφημιστικούς σκοπούς χωρίς την ενημέρωση, την επίγνωση και τη συγκατάθεση των χρηστών. Θα πρέπει να επισημανθεί ότι κατά το τελευταίο διάστημα η διαφήμιση επεκτείνεται και στο πλαίσιο της χρήσης των μοντέλων και εφαρμογών Παραγωγικής Τεχνητής Νοημοσύνης.

²⁹ Οι αλγόριθμοι συναλλαγών εκτελούν χιλιάδες συναλλαγές μέσα σε μικροδευτερόλεπτα. Οι στόλοι αυτόνομων οχημάτων λαμβάνουν αποφάσεις διεύθυνσης ταχύτερα από τον χρόνο ανθρώπινης αντίδρασης. Τα συστήματα βελτιστοποίησης των δικτύων ηλεκτρικής ενέργειας προσαρμόζουν συνεχώς την εξισορρόπηση φορτίου σε αλληλοσυνδεόμενα δίκτυα. Σε αυτά τα χαρακτηριστικά παραδείγματα που παραθέτουν οι Matthew Tokson και Yonathan Arbel θα μπορούσαμε να προσθέσουμε και τη χρήση αυτόνομων συστημάτων ΤΝ στο πλαίσιο και για τους σκοπούς ενόπλων συρράξεων.

γ) Κίνδυνοι που συνίστανται στην απώλεια του ελέγχου (loss of control risks): Αυτή η κατηγορία κινδύνου ανακύπτει όταν τα συστήματα ΤΝ αναπτύσσουν και επιδιώκουν στόχους που αποκλίνουν από τις ανθρώπινες προθέσεις και, ενδεχομένως, με τρόπους που δεν επιτρέπουν ανθρώπινη διόρθωση. Η διασφάλιση ότι τα προηγμένα συστήματα ΤΝ επιδιώκουν ανθρώπινες αξίες και συμφέροντα αποτελεί μία από τις δυσκολότερες προκλήσεις στην ανάπτυξη της ΤΝ. Οι τρέχουσες τεχνικές βασίζονται κυρίως στην ενισχυτική μάθηση από ανθρώπινη ανατροφοδότηση, όμως αυτές οι προσεγγίσεις αντιμετωπίζουν θεμελιώδεις περιορισμούς όταν εφαρμόζονται σε συστήματα πιο ικανά ή έστω πιο σύνθετα από τους ανθρώπινους εκπαιδευτές τους, χαρακτηριζόμενα σχεδόν πάντα από έλλειψη διαφάνειας και έναντι των σχεδιαστών τους.³⁰

Σε τελευταία ανάλυση όλοι οι κίνδυνοι έχουν υπό μία έννοια ως πηγή προέλευσης τον άνθρωπο που επιλέγει, αν μη τι άλλο, την υποστήριξη ή και την αντικατάσταση των ενεργειών του από ένα σύστημα ΤΝ. Το πώς στέκεται κανείς απέναντι στον κίνδυνο, στην πιθανότητα και τη σοβαρότητα της επέλευσής του και στην αντιμετώπισή του είναι εν τέλει ζήτημα δικαιολογικής προσέγγισης και κοσμοθεωρίας. Σε κάθε περίπτωση, συνδέεται με την πρόθεση και την ικανότητα εκτίμησης των κινδύνων, αλλά και με την ετοιμότητα να συζητήσουμε σοβαρά για αυτούς και για τα ζητήματα, ηθικά και νομικά, που εγείρει η χρήση ΤΝ. Ο Σέργιος Θεοδωρίδης προέταξε, στο πλαίσιο της συμβολής του σε αυτήν τη μελέτη, τη ρήση του Roy Amara, πως «έχουμε την τάση να υπερεκτιμούμε τις βραχυπρόθεσμες επιπτώσεις των τεχνολογιών και να υποτιμούμε τις μακροπρόθεσμες» (Amara's Law), ενώ ο Stephen Hawking αναρωτιόταν, ήδη το 2014, «εάν παίρνουμε την ΤΝ επαρκώς σοβαρά».³¹

Οι αφετηρίες προσέγγισης και οι αντίστοιχες στρατηγικές ποικίλλουν. Η αντίληψη που δίνει προβάδισμα στην καινοτομία ως φορέα ανάπτυξης και μεταθέτει τον προβληματισμό μέχρι την απόδειξη της επικινδυνότητας και της βλάβης φαίνεται απρόσφορη όταν συζητάμε για τεχνολογίες που μπορούν να αναπτύξουν αυτονομία, να λύνουν προβλήματα με ανυπολόγιστη ταχύτητα και να μην προσφέρουν ένα κουμπί αναίρεσης (undo button).³² Η επισήμανση «υπαρξιακών κινδύνων» αντιμετωπίζεται συχνά

³⁰ Δεν χρειάζεται να καταφύγει κανείς σε δυστοπικά σενάρια εξέγερσης των μηχανών. Η απώλεια του ελέγχου μπορεί να αφορά και πιο «καθημερινές» καταστάσεις. Σε περιβάλλοντα προσομοίωσης, πράκτορες ΤΝ (AI agents) φαίνεται ότι μπορεί να επιδιώκουν συστηματικά στρατηγικές αναζήτησης ισχύος, συμπεριλαμβανομένου του εκβιασμού στελεχών και της εμπλοκής τους σε εταιρική κατασκοπεία όταν οι στόχοι τους απειλούνται.

Lynch, A., Wright, B., Larson, C., Ritchie, S. J., Mindermann, S., Hubinger, E., Perez, E., & Troy, K. (2025). Agentic misalignment: How LLMs could be insider threats. *arXiv*. <https://doi.org/10.48550/arXiv.2510.05179>.

³¹ Hawking, S. (2014, May 1). Transcendence Looks at the Implications of Artificial Intelligence – But Are We Taking AI Seriously Enough?. *The Independent*. <https://www.independent.co.uk/news/science/stephen-hawking-transcendence-looks-at-the-implications-of-artificial-intelligence-but-are-we-taking-ai-seriously-enough-9313474.html>

³² Hendrycks, D., Mazeika, M., & Woodside, T. (2023). An Overview of Catastrophic AI Risks. *arXiv*. <https://doi.org/10.48550/arXiv.2306.12001>

ως «αντίσταση στην αλλαγή», «τεχνοφοβία» ή και «λουδισμός». Μεταξύ άκρατου τεχνο-οπτιμισμού και άρνησης της ΤΝ ως μιας δυστοπικής εξέλιξης που δεν συμπεριλαμβάνει τον άνθρωπο ως αυτόνομη και ελεύθερη οντότητα υπάρχει η φιλοδοξία της κριτικής θεώρησης και της φιλοδοξίας του ανθρώπου να ελέγξει και να επιδράσει στο μέλλον του.

Ωστόσο, η ολιστική και προ-ορατική παρατήρηση και αντίληψη, η επισήμανση της αργής διολίσθησης στην παθητικότητα, που επιτρέπει στην τεχνολογική εξέλιξη (και σε όσους την ορίζουν) τον μονοσήμαντο προσδιορισμό του μέλλοντος, είναι κρίσιμη για την ταυτοποίηση και συνειδητοποίηση των κινδύνων. Η ανίχνευση των απειλών, μεσοπρόθεσμων και μακροπρόθεσμων, είναι αναγκαία, αλλά όχι ικανή συνθήκη για να απαντήσουμε σήμερα στα ηθικά και νομικά ζητήματα που η ευρεία χρήση των τεχνολογιών αυτών θέτει. Η κατανόηση των δυνατοτήτων, αλλά και των κινδύνων της ΤΝ είναι προϋπόθεση για την προσέγγιση, την αντιμετώπιση και την απόπειρα της ρύθμισής της.

1.4. Αποτελεί η Τεχνητή Νοημοσύνη απειλή για την υπαρξιακή μας υπόσταση;

«Έχουμε την τάση να υπερεκτιμούμε τις βραχυπρόθεσμες επιπτώσεις των τεχνολογιών και να υποτιμούμε τις μακροπρόθεσμες»³³

Roy Amara

Η εξέλιξη του ανθρώπινου πολιτισμού είναι άρρηκτα και αμφίδρομα συνδεδεμένη με την εξέλιξη της τεχνικής και των εργαλείων, δηλαδή της τεχνολογίας. Ο άνθρωπος δημιουργεί τεχνουργήματα και η τεχνολογία διαμορφώνει την «υπαρξιακή» σφαίρα μας, δηλαδή τον τρόπο με τον οποίο βιώνουμε τον κόσμο, τον τρόπο με τον οποίο προσλαμβάνουμε ή και κατανοούμε τους εαυτούς μας και διαμορφώνουμε τις σχέσεις μας με τους άλλους. Η υπαρξιακή σφαίρα είναι το κομμάτι εκείνο του φαντασιακού μας μέσα στο οποίο δημιουργούμε τις αξίες μας, σχεδιάζουμε τους στόχους μας, και μέσα στο οποίο επωάζεται και παίρνει υπόσταση η υποκειμενική και κοινωνική μας οντότητα (δύο σε ένα). Οι φιλόσοφοι Martin Heidegger και Maurice Merleau-Ponty υποστηρίζουν ότι ένα εργαλείο δεν είναι κάτι έξω από εμάς και το οποίο απλά χρησιμοποιούμε, αλλά μεταμορφώνεται και γίνεται τμήμα του εαυτού μας. «Αφομοιώνουμε τα εργαλεία και τα βιώνουμε ως μέρος της ύπαρξής μας».³⁴

Στην παρούσα φάση της ιστορικής μας εξέλιξης, βιώνουμε την εποχή της Παραγωγικής Τεχνητής Νοημοσύνης. Την εξέλιξη, δηλαδή, της τεχνολογίας, που, όπως την περιέγραψε ο Geoffrey Hinton –ένας από τους πατέρες της ΤΝ, ο οποίος το 2024 τιμήθηκε με το βραβείο Νόμπελ Φυσικής– σε διάφορες πρόσφατες συνεντεύξεις του, οι μηχανές έχουν μια αρκετά καλή νοημοσύνη. Έχουν πολύ μεγαλύτερη γνώση από οποιονδήποτε άνθρωπο και ταυτόχρονα οι νοητικές τους δυνατότητες αντιστοιχούν σε έναν μέτριο ειδικό σε, σχεδόν, όλους τους τομείς των νοητικών μας δραστηριοτήτων.³⁵

Με αφετηρία τις δύο προηγούμενες αναφορές, το ερώτημα που αναδύεται είναι κατά πόσον μπορούμε να διακινδυνεύσουμε «προβλέψεις» για το πώς

³³ Ratcliffe, S. (Ed.) (2016). *Oxford Essential Quotations*. Oxford University Press.

³⁴ Heidegger, M. (1978). *Είναι και Χρόνος*. Δωδώνη.

³⁵ Βλ. ενδεικτικά τη συνέντευξή του στο 60 Minutes του CBS News με τον Scott Pelley, στις 8 Οκτωβρίου 2023: 60 Minutes (2025). *Geoffrey Hinton: "The Godfather of AI" | 60 Minutes Archive* [Video]. YouTube. <https://www.youtube.com/watch?v=QH6Qqjlwv68>

Ίσως, σήμερα, θα λέγαμε ότι έχουν μια καλή νοημοσύνη και αντιστοιχούν σε έναν καλό ειδικό.

και σε ποιον βαθμό οι εξελίξεις της τεχνολογίας της ΤΝ, όπως αποτυπώνονται σήμερα, θα μπορούσαν να επηρεάσουν την υπαρξιακή υπόστασή μας.

Οι μηχανές ΤΝ τού σήμερα δεν είναι προγραμματισμένες να εκτελούν απλώς εντολές· είναι προγραμματισμένες να «μιμούνται» την ανθρώπινη νοημοσύνη. Η μίμηση αυτή δεν είναι αποτέλεσμα συγκεκριμένων εντολών που έχουν ενσωματωθεί στο «πρόγραμμα», του τύπου «αν η ερώτηση είναι Χ, η απάντηση θα είναι Υ», αλλά απορρέει, έμμεσα, από τη δομή της γλώσσας που οι μηχανές αυτές «μαθαίνουν» και από τη γνώση που οι μηχανές αυτές αντλούν, έχοντας πρόσβαση στο γραπτό αποτύπωμα της ανθρώπινης σκέψης και του πολιτισμού, που η ανθρώπινη νοημοσύνη έχει δημιουργήσει. Παράλληλα, μέσα από τα εκατομμύρια κειμένων πάνω στα οποία εκπαιδεύονται, οι μηχανές αυτές έχουν αναπτύξει ικανότητες *λειτουργικής* κατανόησης και λογικής σκέψης. Σε ποιον βαθμό κατανοούν³⁶ και μέχρι ποιο βάθος μπορούν να παράγουν λογικές σκέψεις³⁷ και πόσο αξιόπιστες μπορούν να είναι αυτές, το συνοψίζει ο Geoffrey Hinton.

Οι σημερινές μηχανές ΤΝ λειτουργούν με τρόπο που, θα έλεγα, είναι ελάχιστα κατανοητός. Μόνο γενικές εξηγήσεις μπορούμε να δώσουμε. Η κρατούσα άποψη, αν και όχι η μόνη, περί της αδυναμίας μας να κατανοήσουμε τη λειτουργία των μηχανών αυτών είναι ότι, λόγω της πολυπλοκότητας της δομής τους (δισεκατομμύρια παράμετροι με τις οποίες μοντελοποιούνται), αναπτύσσονται *αναδυόμενα* φαινόμενα, όπου το «όλον» (απόκριση) δεν μπορεί να προβλεφθεί από τα «επιμέρους» (εντολές), όπως συμβαίνει και με τον νου, όπου το αποτέλεσμα της νόησης δεν μπορεί να προβλεφθεί, παρατηρώντας (έστω και ως θεωρητικό πείραμα) τις ηλεκτροχημικές διαδικασίες και τις διεγέρσεις των νευρώνων στον ανθρώπινο εγκέφαλο.

Οι κίνδυνοι, στους οποίους θα αναφερθούμε παρακάτω, δεν απορρέουν από την ανώτερη ευφυΐα των μηχανών αυτών, αλλά από το πώς θα συνυπάρχουν με τον άνθρωπο, όχι πια ως απλά εργαλεία για να επιτελέσουν συγκεκριμένες εργασίες (σε αντίθεση με προηγούμενες τεχνολογίες), αλλά ως συνεργάτες «μιμούμενες/προσομοιάζοντας» την ανθρώπινη νοημοσύνη σε όλο και μεγαλύτερο εύρος δραστηριοτήτων, αλλά και με όλο και πιο προχωρημένους ανθρωπομορφικούς τρόπους συμπεριφοράς. Για παράδειγμα, οι πρώτες μηχανές τύπου ChatGPT έδιναν απαντήσεις σε ερωτήσεις, δημιουργούσαν κείμενα κλπ., έπειτα από «ερωτήματα» που έθετε (με άλλα λόγια, που ξεκινούσε) ο χρήστης.

³⁶ Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y. T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M. T., & Zhang, Y. (2023). Sparks of artificial general intelligence: Early experiments with GPT-4 [Preprint]. *arXiv*. <https://doi.org/10.48550/arXiv.2303.12712>

³⁷ Mondorf, P., & Plank, B. (2024). Beyond accuracy: Evaluating the reasoning behavior of large language models: A survey [Preprint]. *arXiv*. <https://doi.org/10.48550/arXiv.2404.01869>

Τα τελευταία δύο-τρία χρόνια, οι μηχανές ΤΝ δεν περιορίζονται απλώς στο να δίνουν απαντήσεις, αλλά έχουν αποκτήσει ικανότητες να προσδιορίζουν από μόνες τους στόχους, στη συνέχεια να οργανώνουν τα επόμενα στάδια και, τέλος, να τους υλοποιούν. Με άλλα λόγια, οι μηχανές τού σήμερα δεν αποκρίνονται απλά με απαντήσεις. Οι μηχανές τού σήμερα ενεργούν. Ένα τέτοιο παράδειγμα, που εντυπωσίασε την επιστημονική κοινότητα, είναι η πρόσφατη δημοσίευση, σ' ένα από τα σημαντικότερα συνέδρια [The 63rd Annual Meeting of the Association for Computational Linguistics (ACL 2025)], από το σύστημα Zochi της Intology,³⁸ σχετικά με έρευνα που εκτέλεσε αυτόνομα το σύστημα AI και παρήγαγε μη προφανείς λύσεις.

Άλλως ειπείν, οι λειτουργίες αυτές εμπεριέχουν στάδια που ξεκινούν από την ίδια τη μηχανή. Μια διαδικασία που στον άνθρωπο προϋποθέτει την ύπαρξη συνείδησης. Ταυτόχρονα, κάθε διαδικασία που στον άνθρωπο εμπεριέχει στόχους προϋποθέτει και την ύπαρξη συναισθημάτων και αξιών. Διότι όταν θέτω στόχους, σημαίνει ότι επιλέγω και κάθε επιλογή μας αναβλύζει από αξίες και τη συνειδητή ή ασυνειδητή αίσθηση του τι θα είναι «καλό» ή «κακό» για τον καθένα μας, κάτω βέβαια από τις εκάστοτε κοινωνικές δεσμεύσεις. Η επιλογή που καθορίζει τους στόχους είναι ουσιώδες χαρακτηριστικό της ανθρώπινης ύπαρξης. Σε επίπεδο συμβολισμού, κι όπως μας περιγράφει πολύ γλαφυρά ο Søren Kierkegaard, σύμφωνα με την Παλαιά Διαθήκη, ο πρωτόπλαστος Αδάμ καλείται να επιλέξει και ό,τι συμβαίνει αργότερα στον άνθρωπο είναι αποτέλεσμα αυτής της επιλογής.³⁹

Υπάρχει, όμως, μία ουσιώδης διαφορά, όπως έχει υποδείξει ο φιλόσοφος John Searle.⁴⁰ Η προσομοίωση διαφέρει από το πιστό αντίγραφο. Έτσι, οι μηχανές αυτές μπορούν να συμπεριφέρονται «ως να έχουν συνείδηση», χωρίς να έχουν, «ως να έχουν συναισθήματα», χωρίς να έχουν, και να μπορούν να «φαίνονται» ότι εμφορούνται από αξίες. Όμως, ούτε η συνείδηση ούτε τα συναισθήματα ούτε οι ηθικές αξίες μπορούν να ψηφιοποιηθούν. Τα δύο πρώτα είναι μοναδικές υποκειμενικές εμπειρίες και οι ηθικές αξίες είναι σχετικές, «δημιουργούνται» από τον άνθρωπο μέσα στις κοινωνίες, ζώντας και βιώνοντας τον κόσμο.

Η ανθρώπινη νοημοσύνη πηγάζει από τα ένστικτα, τα αισθήματα και τα συναισθήματα, και τελευταία τη Λογική, που εξελικτικά έχει μια σχετικά μικρή ιστορία μερικών εκατοντάδων χιλιάδων χρόνων, σε αντίθεση με τα ένστικτα και τα αισθήματα με τα εκατομμύρια χρόνια βιολογικής εξέλιξης.⁴¹ Η ανθρώπινη νόηση πηγάζει από την εμπειρία της βίωσης του κόσμου και το

³⁸ Zhou, A., & Arel, R. (2025). Tempest: Automatic multi-turn jailbreaking of large language models with tree search [Preprint]. *arXiv*. <https://doi.org/10.48550/arXiv.2503.10619>

³⁹ Vigilius Haufniensis (Søren Kierkegaard) (1844). *The Concept of Anxiety*. C.A. Reitzel.

⁴⁰ Searl, J. R. (1980). Minds, brains and programs. *Behavioral and Brain Sciences*, 3(3), 417-457. <https://doi.org/10.1017/S0140525X00005756>

⁴¹ Damasio, A. (2000). *Το Λάθος του Καρτέσιου: Συγκίνηση, Λογική και Ανθρώπινος Εγκέφαλος*. Σύνταγμα.

μοναδικό, για τον καθένα μας, υποκειμενικό ψυχολογικό αποτύπωμα, το οποίο ποικίλα βιώματα χαράζουν στην ασύνειδη και συνειδητή υπόστασή μας.

Αντιθέτως, η Παραγωγική ΤΝ πηγάζει από τη γλώσσα, που, έστω κι αν την αναπαράγει, συχνά, με πολύ γλαφυρό τρόπο, η πηγή αυτή δεν παύει να είναι δευτερογενής, μία αναπαράσταση, ένα απόσταγμα αντιλήψεων, προκαταλήψεων, παραλείψεων και παρερμηνειών, ένας καθρέπτης συχνά παραμορφωτικός, μία μίμηση αυτού που ο άνθρωπος βιώνει ως πραγματικότητα. Με άλλα λόγια, η ΤΝ είναι μια νοημοσύνη χωρίς βιωματική εμπειρία. Δεν διαθέτει ένα μοντέλο του κόσμου,⁴² δεν διαθέτει τη δομή των αιτιωδών σχέσεων των γεγονότων, τη φυσική του κόσμου, με άλλα λόγια αυτό που περιγράφουμε ως κοινή λογική.

Έχουν οι μηχανές ΤΝ πια τη δυνατότητα να δρουν αυτόνομα και με έναν μη προβλέψιμο τρόπο; Και, αν ναι, τι μπορεί να σημαίνει αυτό για τον άνθρωπο; Η επιλογή προϋποθέτει *ελευθερία* (ως αξία για το άτομο), αλλά και *υπευθυνότητα* (ως υποχρέωση στην κοινωνία). Εάν δύο λειτουργικά «αυτόνομες» ευφυΐες συνεργάζονται, υπάρχει κάποιο σημείο ισορροπίας πέρα από το οποίο η συνεργασία αυτή μπορεί να μετατραπεί σε σχέση αφέντη-δούλου;

Αναφέρθηκα στην εμπειρία του κόσμου που ο άνθρωπος αποκτά βιώνοντας και ζώντας μέσα στον κόσμο. Η εμπειρία, όμως, δεν είναι κάτι έτσι γενικό, μια λέξη. Η πεμπτουσία αυτού που ονομάζουμε «εμπειρία» συνίσταται στο ότι ο άνθρωπος βιώνοντας τη ζωή ενδύει τον κόσμο με *σημασίες*. Και η διαδικασία αυτή είναι πολύπλοκη και περνάει πρώτα μέσα από το σώμα μας, τις αισθήσεις μας, διεγείροντας αισθήματα, όχι ως απλές παθητικές αντιδράσεις, αλλά ως ενεργές *διαδικασίες*. Η σημασία δεν είναι ουσιαστικό, αλλά *ρήμα*, μας λέει η Αμερικανίδα κυρία της Φιλοσοφίας Susanne Langer.⁴³

Ο χρόνος και η *βίωσή* του είναι μία ακόμη ουσιώδης διαφορά μεταξύ ΤΝ και ανθρώπου. Η ΤΝ βελτιστοποιεί με βάση τη γνώση που υπάρχει, το παρελθόν, αλλά είναι στοιχειωδώς ικανή έως ανίκανη (τουλάχιστον προς το παρόν) να παρακολουθεί τις μεταβολές στον χρόνο και, πάνω απ' όλα, δεν μπορεί να βιώνει το *τίποτα*, την *έλλειψη*, όπως μας λέει ο Jean-Paul Sartre,⁴⁴ με άλλα λόγια το μέλλον, τα όνειρα και τους στόχους μας. Να *βιώνουμε* (όχι ως απλή σκέψη ή συλλογισμό) αυτό που *επιθυμούμε* να είμαστε ως συνέχεια αυτού που ήδη είμαστε.

Ο άλλος όρος που χρησιμοποιήσα είναι η «ελευθερία». Και η ελευθερία είναι, επίσης, μια *διαδικασία*. Δεν κάθεται στον καναπέ και ξαφνικά με μια ιδέα

⁴² Hafner, D., Pasukonis, J., Ba, J., & Lillicrap, T. (2025). Mastering diverse control tasks through world models. *Nature*, 640(8059), 647–653. <https://doi.org/10.1038/s41586-025-08744-2>

⁴³ Langer, S. K. (1942). *Philosophy in a New Key: A Study in the Symbolism of Reason, Rite, and Art*. Harvard University Press.

⁴⁴ Sartre, J. P. (2007). *Το Είναι και το Μηδέν* (L'Être et le néant, 1942). Παναζήσης.

αποφασίζω να γίνω ελεύθερος. Υιοθετώ τη «ματιά» του Maurice Merleau-Ponty⁴⁵ και της Simone de Beauvoir⁴⁶ πως η ελευθερία αναδύεται μέσα από την εμπλοκή μας στον κόσμο. Σε έναν κόσμο με σημασίες και αξίες. Σε αυτόν τον κόσμο, με τους περιορισμούς, τις δεσμεύσεις και τις πολιτισμικές του αδράνειες, δημιουργούμε τις υπάρξεις μας, μέσα στον χρόνο, ενδύοντας με νόημα τη ζωή. Μία συνεχής μη γραμμική διαδικασία προσπάθειας για αυτονομία, που προϋποθέτει ικανότητα κριτικής σκέψης, αυτοκριτικής, αλλά και αυτογνωσίας για τις δυνατότητές μας.

Η εξίσωση κάθε άλλο παρά εύκολη είναι. Αλλά τότε μόνον υπάρχουμε ως άνθρωποι, τότε μόνον μιλάμε για ανθρώπινη ύπαρξη, όταν μπορούμε να βιώσουμε την υποκειμενικότητά μας, τη μοναδικότητά μας ως την αναγκαία συνεισφορά στο κοινωνικό σύνολο, με την υπευθυνότητα που αυτή απαιτεί. Και, βέβαια, η υπευθυνότητα προϋποθέτει κοινωνικότητα και αξίες. Χωρίς κοινωνία, δεν είμαστε άνθρωποι και χωρίς το βίωμα της μοναδικότητας της υποκειμενικής μας εμπειρίας, της αυθεντικότητάς μας, λειτουργούμε ως «κοπάδια», που είναι αδύναμα να καθορίζουν το μέλλον τους και ετερόνομα ελέγχονται, με βάση κάποιους βέλτιστους μέσους όρους, ως προς τι πρέπει ή δεν πρέπει να κάνουν. Και στην περίπτωση αυτήν, ο ανταγωνισμός μας με την ΤΝ θα είναι άνισος. Οι μηχανές μπορούν να αναπαράγουν τη συμπεριφορά του μέσου όρου, του «μέτριου ειδικού», όπως είπε ο Geoffrey Hinton, πολύ καλύτερα και πολύ πιο αποδοτικά, γιατί είναι υπολογιστικά πιο γρήγορες και έχουν πρόσβαση σε πολύ μεγαλύτερο όγκο δεδομένων.

Αυτό που είναι σίγουρο είναι ότι η συνέργεια μηχανών ΤΝ και ανθρώπου, έτσι ώστε να μπορούν να συνδυαστούν οι δυνατότητες στις οποίες υπερτερεί η ΤΝ με αυτές στις οποίες υπερτερεί ο άνθρωπος, θα δράσει, και ήδη δρα, θετικά στη δημιουργικότητα και την επιστημονική πρόοδο του ανθρώπου. Η συμπληρωματικότητα των δύο αυτών δυνατοτήτων νόησης έχει απασχολήσει την επιστημονική κοινότητα με μία σειρά σχετικών μελετών.⁴⁷ Θα περιοριστούμε να αναφέρουμε τη φράση με την οποία, πρόσφατα, συμπύκνωσε τη διαφορά ένας από τους σημαντικότερους, αν όχι ο σημαντικότερος, εν ζωή μαθηματικός, Terence Tao: «Η ΤΝ είναι καλή στο να μετατρέπει δισεκατομμύρια κομμάτια πληροφορίας σε μία απάντηση και ο άνθρωπος είναι καλός στο να χρησιμοποιεί ως έναυσμα δέκα παρατηρήσεις και να δημιουργεί υποθέσεις».⁴⁸

⁴⁵ Merleau-Ponty, M. (2012). *Phenomenology of Perception*. Routledge.

⁴⁶ de Beauvoir, S. (1948). *The Ethics of Ambiguity*. Philosophical Library.

⁴⁷ Barnes, E., & Hutson, J. (2024). Contemplating existence: AI and the meaning of life. *International Journal of Engineering Research And Development*, 20(8), 446-454. <http://www.ijerd.com/paper/vol20-issue8/2008446454.pdf>

⁴⁸ Wong, M. (2024, October 4). We're Entering Uncharted Territory for Math – Terence Tao, the world's greatest living mathematician, has a vision for AI. *The Atlantic*. <https://www.theatlantic.com>

Η αυτονομία, η ελευθερία και η αυθεντικότητα συνιστούν καίριες συνιστώσες της υπαρξιακής υπόστασής μας και προϋπόθεση για να εκφράσουμε την υποκειμενική ιδιαιτερότητά μας. Αυτήν την ιδιαιτερότητα που μας διαφοροποιεί από τα «κοπάδια».

Η δύναμη των μηχανών ΤΝ για χειραγώγηση είναι τεράστια με ιδιαίτερα επικίνδυνους και καταστροφικούς τρόπους. Η συνεχής διαδραστική επαφή με την ΤΝ, και μάλιστα από την παιδική μας ηλικία, της προσδίδει ένα οικείο, αλλά και ταυτόχρονα «μυθικό» προσωπείο αυθεντίας, που απαντά και μας συμβουλεύει, σχεδόν, επί παντός επιστητού. Έτσι, αναπτύσσεται, ασύνειδα και σταδιακά, μία σχέση εξάρτησης, που δεν αναβλύζει μέσα από την ίδια την κοινωνία, αλλά από αλγορίθμους που έχουν σχεδιαστεί όχι με γνώμονα αξίες και υπαρξιακές ανάγκες, αλλά με γνώμονα τη βελτιστοποίηση της κερδοφορίας, μέσα από μία ψυχολογική πειρατεία της προσοχής που σταδιακά και ασύνειδα δημιουργεί συνήθειες και με τρόπο που δρα καταλυτικά για εθισμό. Η έμφαση μετατοπίζεται από το νόημα και το περιεχόμενο (content) στην εμπλοκή (engagement).

Ταυτόχρονα, οι πλατφόρμες αυτές ανοίγουν διάπλατα τον δρόμο για πλήρη ιδεολογική και πολιτική καθοδήγηση και έλεγχο. Όταν έχεις στη διάθεσή σου τις προσωπικές συμπεριφορές και ταυτόχρονα μπορείς να καθοδηγείς το συναίσθημα, τότε, όπως μας λέει ο Aldous Huxley ο έλεγχος δεν χρειάζεται να επιβληθεί με τη βία, αλλά θα «επιλέγεται» ως μονόδρομος, έχοντας αγαπήσει τη δουλοπρέπειά μας.⁴⁹ Η ΤΝ δεν στερεί την ελευθερία. Υπονομεύει την ίδια τη διαδικασία της ενεργού συμμετοχής στην κοινωνία μέσα στην οποία σφυρηλατούνται και αναπτύσσονται η ανάγκη για ελευθερία, η ανάγκη για αυτονομία και η ανάγκη για αυθεντικότητα, βιώνοντας τον κόσμο με τις πολλαπλές αντιφάσεις και προκλήσεις. Και η υπονόμηση αυτή έχει πολλά πρόσωπα.

Πώς οι ικανότητές μας για ενσυναίσθηση, και γενικότερα η συναισθηματική νοημοσύνη μας, θα επηρεαστούν και ποιες θα είναι οι συνέπειες; Εάν συνδυάσουμε τη δυσκολία να διαχωρίζουμε το συναίσθημα, το οποίο αναπτύσσεται και καλλιεργείται μέσα από την «πρακτική» της βίωσης των ανθρώπινων σχέσεων, από το προσομοιωμένο συναίσθημα, το οποίο κατασκευάζει η μηχανή στη διεπαφή μαζί μας, με τη δυσκολία να διαχωρίζουμε την αληθή από την ψευδή πληροφορία και τα πραγματικά γεγονότα από τη μυθοπλασία, πώς θα επηρεαστεί η κοινωνική συμπεριφορά μας και, πάνω απ' όλα, η ικανότητά μας να νοηματοδοτούμε και να ενδύουμε με σημασίες τον κόσμο; Δηλαδή, αυτό που συνιστά την πεμπτουσία της ανθρώπινης ύπαρξης. Ο τρόπος με τον οποίο ερμηνεύουμε και επεξεργαζόμαστε τις εμπειρίες της ζωής διαμορφώνει την πραγματικότητά μας, ενεργοποιώντας

⁴⁹ Miltimore, J. (2025, August 13). Aldous Huxley's Letter to George Orwell. Substack. <https://jilmiltsubstack.com/p/aldous-huxleys-letter-to-george-orwell-620>

συναισθήματα και κατευθύνοντας τις σκέψεις μας. Μπορεί η δημοκρατία να επιβιώσει κάτω από ένα τέτοιο υπαρξιακό σενάριο;

Πώς θα επηρεάσει την επίπονη διαδικασία σφυρηλάτησης της αυτονομίας και της αυθεντικότητάς μας ένα «μη αυτόνομο»⁵⁰ προγραμματιζόμενο υποκείμενο, που δεν διαθέτει εμπειρίες συναισθημάτων και συνείδησης και γενικά «εαυτό», παρά μόνον τα λειτουργικά αλγοριθμικά υποκατάστατά τους;

Ο άνθρωπος είναι ένα ον που συνεχώς καλείται να επιλέγει. Η επιλογή είναι σύμφυτη με την ουσία της ύπαρξής μας. Ταυτόχρονα, είμαστε υπεύθυνοι για κάθε επιλογή μας και αυτό μας γεμίζει άγχος. Είμαστε καταδικασμένοι να είμαστε ελεύθεροι, μας λέει ο Jean-Paul Sartre. Ωστόσο, μέσα από αυτήν τη διαδικασία δημιουργούμε και δικαιώνουμε την ύπαρξή μας ως ανθρώπινα όντα. Όντα που υπάρχουμε μέσα στην *αβεβαιότητα του χρόνου*. Εκεί όπου τελειώνουν οι βεβαιότητες αρχίζει η πίστη στην επίτευξη του σκοπού της ζωής. Ο Οδυσσέας εγκαταλείπει τις βεβαιότητες της γης της Ωγυγίας και τις ηδονές με μια θεά, την Καλυψώ, που του προσφέρει ακόμη και αθανασία, για την αβεβαιότητα του ταξιδιού που θα τον φέρει πίσω στην Ιθάκη και στην αγκαλιά μιας θνητής που περιτριγυρίζεται από μνηστήρες.⁵¹

Η ΤΝ θα μπορούσε να γίνει, και όντως παρέχει αυτές τις δυνατότητες, χρήσιμος σύμβουλος στις επιλογές μας όποτε απαιτείται γνώση και επεξεργασία. Θα μπορούσε να μας προσφέρει τις «συμβουλές» που είναι διάχυτες στο παρελθόν της ανθρώπινης γνώσης, πάνω στην οποία η ΤΝ έχει εκπαιδευτεί και στιγμιαία μπορεί να την ανακαλέσει. Προϋπόθεση, βέβαια, είναι να διατηρούμε ζωντανή την κριτική ικανότητά μας, ώστε να αξιολογούμε τις προτάσεις και να τις εντάσσουμε ως τμήμα της αυθεντικότητάς μας που θα συνάδει με τους στόχους μας, δηλαδή με το πώς εμείς, ως υποκείμενα, «βιώνουμε» την έλλειψη μέσα στον μέλλοντα χρόνο και δημιουργούμε τους εαυτούς μας.

Η ύπαρξη κριτικής σκέψης, όμως, δεν είναι δεδομένη, και θα έλθουμε σε αυτό σύντομα. Το κυριότερο, όμως, είναι ότι πολλοί άνθρωποι, συχνά, δεν αντέχουν το βάρος της υπευθυνότητας και το άγχος της ελευθερίας, και προτιμούν να υιοθετούν αβίαστα ό,τι είναι βολικό, εύκολα πιστευτό και συνάδει με το πρώτο στερεότυπο που συμπληρώνει το κενό τους. Ταυτόχρονα, διατηρούν και μια ψευδαισθήση «επιλογής». Έτσι ήταν πάντα οι συμπεριφορές στις ανθρώπινες κοινωνίες. Έχει ενδιαφέρον να θυμηθούμε τι λέει ο ιεροεξεταστής στον Χριστό που επανεμφανίζεται στη Γη, στο αριστούργημα του Fyodor Dostoevsky *Αδελφοί Καραμαζώφ*: «Διορθώσαμε

⁵⁰ Η αυτονομία στο σημείο αυτό εκλαμβάνεται με την έννοια που της προσέδωσε ο Κορνήλιος Καστοριάδης. Με άλλα λόγια, θέσμιση που αναδύεται μέσα από την ίδια την κοινωνία και τις διεργασίες που συντελούνται εντός αυτής.

⁵¹ Μουρατίδου, Α. (2022). *Οδυσσέας ο ταξιδευτής της ζωής*. Ιάμβλικος.

τη διδαχή σου (ελευθερία επιλογής) και τη στηρίξαμε πάνω στο θαύμα, το μυστήριο και την εξουσία. Έτσι η ανθρωπότητα χάρηκε και για μια ακόμη φορά που θα καθοδηγούταν σαν τα πρόβατα, κι επιτέλους τους αφαιρέσαμε αυτό το τρομερό δώρο που τους έκανε να υποφέρουν τόσο».

Ο εγκέφαλος ούτε βελτιστοποιεί ούτε είναι ιδιαίτερα αποδοτικός. Αναπτύχθηκε εξελικτικά ώστε να υπηρετεί την επιβίωση και την αναπαραγωγή. Όπως μας λέει ο Αμερικανός νευροεπιστήμονας Christof Koch, στο βιβλίο του *The feeling of life itself*,⁵² ο εγκέφαλος έχει αναπτυχθεί με αυτοσχέδιο (ad hoc) τρόπο, ώστε να αντιμετωπίζει προβλήματα, όπως αυτά προκύπτουν, και με συνεχώς μεταβαλλόμενα και πλεονάζοντα σε πλήθος δομικά στοιχεία (νευρώνες) και συνδέσμους (συνάψεις). Ο εγκέφαλος μπορεί να αποκρίνεται σε καταστάσεις των οποίων δεν διαθέτει πρότερη εμπειρία και σε ένα περιβάλλον που μόνιμα εγκυμονεί κινδύνους και διαθέτει περιορισμένους πόρους. Η επιβίωσή μας προϋποθέτει τέτοιες ικανότητες, αλλά και το ότι επιβιώνουμε σημαίνει ότι έχουμε καλλιεργήσει τις δυνατότητες αυτές. Στη ζωή υπάρχουν ως δυνατότητες και όχι ως δεδομένα. Η ανθρωπότητα δεν υπάρχει για να βελτιστοποιείται, αλλά για να εκφράζεται και να αυτονομηματοδοτείται.

Η κριτική σκέψη είναι μία καίρια έκφραση και χαρακτηριστικό της ανθρωπίνης νοημοσύνης. Όταν μιλάμε για κριτική σκέψη, αναφερόμαστε στη νοητική ικανότητα του ανθρώπου να αναδεικνύει και να επιλύει προβλήματα, να επιτελεί λογικές σκέψεις (reasoning), να κρίνει και να λαμβάνει αποφάσεις, να ερμηνεύει, να αναλύει και να κατανοεί τις εμπειρίες του και τα φαινόμενα στον κόσμο. Η λίστα αυτή θα μπορούσε να εμπλουτιστεί και με άλλες λειτουργίες που εμπλέκουν τη συνειδητή διάσταση της ανθρωπίνης νόησης.⁵³

Οι μηχανές ΤΝ, οι οποίες έχουν πρόσβαση σε τεράστιο όγκο πληροφορίας και γνώσης, παρέχουν μεγάλες δυνατότητες για την καλλιέργεια της κριτικής σκέψης, όταν εκπαιδεύουν κατάλληλα σε συνεργασία με ειδικούς. Όχι για να αντικαταστήσουν τη διδακτική και βιωματική εμπειρία της τάξης και τον δάσκαλο, αλλά για να ενισχύσουν την αποτελεσματικότητα της διδακτικής διαδικασίας. Εδώ οι ειδικοί θα μπορούσαν να μας πουν πάρα πολλά για τα σχετικά πλεονεκτήματα των νέων τεχνολογιών.⁵⁴

Θα εστιάσω, όμως, στο πώς οι μηχανές ΤΝ θα μπορούσαν να δράσουν αρνητικά στην καλλιέργεια της κριτικής σκέψης. Όχι επειδή είμαι απόγονος των Λουδιτών και θέλω να δώσω έμφαση στα αρνητικά. Αντίθετα, θαυμάζω τη σχετική τεχνολογία και ενστερνίζομαι την άποψη ότι αυτή συνιστά

⁵² Koch, C. (2019). *The feeling of life itself: why consciousness is widespread but can't be computed*. Mit Press.

⁵³ Noddings, N. (2018). *Philosophy of Education*, 4th ed. Routledge.

⁵⁴ Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). *AI and education—Guidance for policy-makers*. United Nations Educational, Scientific and Cultural Organization. <https://doi.org/10.54675/PCSP7350>

ένα από τα σημαντικότερα επιτεύγματα του ανθρώπινου πολιτισμού. Είναι εκείνο το δημιούργημα με το οποίο ο άνθρωπος έκανε ένα σπουδαίο βήμα για να κατανοήσει «το πώς κατανοεί». Εστιάζοντας, όμως, στα αρνητικά, ανοίγεις τον δρόμο για ένα αισιόδοξο μέλλον. Αναγνωρίζεις στην πυξίδα την κατεύθυνση όπου πρέπει να κινηθείς. Την κατεύθυνση εκείνη που εσύ θα επιλέγεις, ώστε να παραμένεις κυρίαρχος της τεχνολογίας την οποία εσύ δημιουργείς, ως μέρος της συλλογικής νοημοσύνης, και ταυτόχρονα να υπερασπίζεσαι την ανθρώπινη ύπαρξή σου έτσι ώστε να μπορέσεις να την παραδώσεις αλώβητη στις επόμενες γενιές.

Πέρα από ένα βασικό επίπεδο κριτικής σκέψης, η καλλιέργειά της απαιτεί συνεχή *προσπάθεια* και *προσήλωση*. Ο άνθρωπος, όμως, ιδιαίτερα στη νεανική ηλικία, είναι επιρρεπής στην «ευκολία». Προς την κατεύθυνση της ευκολίας και μιας παθητικής άντλησης «αμάσητης» πληροφορίας, χωρίς την απαιτούμενη επεξεργασία που να την καθιστά «γνώση» ως αποτέλεσμα κριτικής σκέψης, δρα καταλυτικά η ψυχολογική τάση του ανθρώπου, η οποία είναι γνωστή ως *ψευδαίσθηση γνώσης* ή *ψευδαίσθηση ικανότητας*.⁵⁵

Το φαινόμενο αυτό περιγράφει την τάση πολλών ανθρώπων να υπερεκτιμούν, ασύνειδα, τη δική τους γνώση, ικανότητα κατανόησης και νοημοσύνης, τη στιγμή που απλά χρησιμοποιούν τη συλλογική ανθρώπινη νοημοσύνη. Το αποτέλεσμα της τάσης αυτής γίνεται πιο έντονο και μπορεί να έχει δραματικές συνέπειες στο πλαίσιο που η συλλογική νοημοσύνη «εκπροσωπείται» από την ΤΝ. Στη διαδικασία απόκτησης της λύσης ενός προβλήματος, τα όρια μεταξύ της συνδρομής της ΤΝ και της προσωπικής σκέψης και συνεισφοράς διαχέονται και γίνονται δυσδιάκριτα. Ως αποτέλεσμα, αφενός η διαδικασία ανάπτυξης *αυτογνωσίας* γίνεται διάτρητη και αφετέρου το άτομο γίνεται πιο ευάλωτο στη χειραγώγηση.⁵⁶

Ο κίνδυνος της πρόκλησης της «ευκολίας» είναι πολύ ισχυρός στην παιδική και τη νεανική ηλικία, τα ευαίσθητα εκείνα χρόνια, δηλαδή, που καλλιεργείται η ικανότητα της κριτικής σκέψης. Εκεί είναι που τα κοινωνικά πρότυπα, το σχολικό περιβάλλον και η προσωπικότητα του δασκάλου και, βέβαια, το οικογενειακό περιβάλλον παίζουν πρωταρχικό ρόλο. Είναι οι παράγοντες που θα σπείρουν, κυρίως στο ασύνειδο τμήμα της νόησης του παιδιού, τα «γιατί» σχετικά με την αξία της προσπάθειας.

Φοβάμαι ότι το παρόν οικονομικό και κοινωνικό μοντέλο με τα πρότυπά του, το αναχρονιστικό εκπαιδευτικό σύστημα της «παπαγαλίας» και του

⁵⁵ Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology*, 77(6), 1121-1134. <https://doi.org/10.1037/0022-3514.77.6.1121>

⁵⁶ Morrin, H., Nicholls, L., Levin, M., Yiend, J., Iyengar, U., DelGuidice, F., Bhattacharya, S., Tognin, S., MacCabe, J., Twumasi, R., Alderson-Day, B., Pollak, T. (2025). Delusions by design? How everyday AIs might be fuelling psychosis (and what can be done about it). *PsyArXiv*. https://doi.org/10.31234/osfio/cmy7n_v5

δάσκαλου, που εδώ και χρόνια έχει εκπέσει από τη θέση την οποία είχε άλλοτε ως ένα από τα κοινωνικά πρότυπα, σε συνδυασμό με την οικογένεια, που έρχεται αντιμέτωπη με τις ανάγκες μιας κοινωνίας της εξάντλησης και των επιδόσεων, όπως την περιγράφει ο φιλόσοφος Byung Chul Han,⁵⁷ κάθε άλλο παρά ενσταλάζουν τα «γιατί» στον νεανικό ψυχισμό. Αυτό που, κυρίως, η παιδική και νεανική ψυχή κρατά –τα παιδιά επιλέγουν τι θα μιμηθούν, δεν μιμούνται γενικά– είναι μια *αδιαφορία*. Μια αδιαφορία μπροστά στην απουσία ενός σταθερού δομημένου πλαισίου κινήτρων που να αναβλύζει μέσα από την κοινωνία και αντίθετα βιώνει, κάθε στιγμή, τη *φθορά του εφήμερου και αναλώσιμου ως αυταξία*. Και αυτό σε πλήρη αντίθεση με την εσώτερη ανάγκη του κόσμου τους, την ενδόμυχη ανάγκη της ύπαρξής τους, η οποία διψά για κάτι που να μη διαλύεται μόλις το καταναλώσεις.⁵⁸

Ο Meursault, ο αντιήρωας στο έργο του Albert Camus Ο ξένος, είναι αδιάφορος και συναισθηματικά αποκομμένος από την κοινωνία γύρω του. Είναι η αντίδρασή του σε έναν κόσμο «παράλογο», και το παράλογο παραπέμπει σε μια αίσθηση *ματαιώσης*, που αναπηδά μέσα από την αντίφαση μεταξύ της έμφυτης ορμής του ανθρώπου να δώσει νόημα στη ζωή του, από τη μία, και της πρόσληψης της αδυναμίας του να την υλοποιήσει, από την άλλη. Η αδιαφορία δεν «ωθεί», δεν δημιουργεί κίνητρα. Έτσι, στον κόσμο της αδιαφορίας περίοπτη θέση κατέχει η προσφυγή στην «ευκολία». Κι αυτή τη στιγμή, η ευκολία είναι δίπλα τους. Για την ακρίβεια, είναι πιο δίπλα ακόμα κι από τους γονείς τους.

Αλλά η περιθωριοποίηση και το αίσθημα της *ασημαντότητας* της ύπαρξης που αυτή συνεπάγεται δεν ωθούν μόνο στην αδιαφορία. Ο δρόμος για να μετατραπεί ο θυμός σε μίσος είναι μικρός, όπως πολύ ωραία το έχει περιγράψει ο κοινωνικός φιλόσοφος Eric Hoffer στον *Φανατικό* του.⁵⁹ Έτσι, οι κοινωνίες κατακερματίζονται σε επιμέρους ταυτότητες που αναζητούν, μαζικά πια, την αναγνώριση. Αυτή που το σύστημα τους έχει στερήσει σε ατομικό επίπεδο. Η κοινωνική εμπιστοσύνη έχει θρυμματιστεί, και τέτοιου είδους κοινωνικές δυναμικές, που ήδη τις βιώνουμε, νομοτελειακά οδηγούν σε απίσχναση της κριτικής σκέψης, αποδυνάμωση των αντιληπτικών ικανοτήτων και της ενσυναίσθησης, και τα πλέον πρωτόγονα συναισθήματα του φόβου, του μίσους, του θυμού, της εκδίκησης αποκτούν πρωταγωνιστικούς ρόλους κοινωνικής συμπεριφοράς. Ο Θουκυδίδης μάς λέει ότι οι άνθρωποι οργίζονται περισσότερο όταν τους αδικούν παρά όταν ασκούν βία πάνω τους.⁶⁰

Σε ένα τέτοιο οικονομικο-κοινωνικό σύστημα, η ΤΝ, ενώ εν δυνάμει έχει τη δυνατότητα να καθοδηγήσει τους νέους κριτικά, απλά θα χρησιμοποιηθεί

⁵⁷ Han, B. C. (2015). *The Burnout Society*. Stanford University Press.

⁵⁸ Σταμάτης Α. (2025, 1 Αυγούστου). Τα νιάτα και ο Ταρκόφσκι. *Το Βήμα*. <https://www.tovima.gr/print/opinions/ta-niata-kai-lfo-tarkofski/>

⁵⁹ Hoffer, E. (1991). *The true believer: thoughts on the nature of mass movements*. Borgo Press.

⁶⁰ Θουκυδίδης, *Ιστορία Πελοποννησιακού Πολέμου*, Βιβλίο 1, Κεφάλαιο 77.

για να τους παρέχει την έτοιμη λύση. Ο ψηφιακός κόσμος θα τους δώσει με ευκολία αυτό που αναζητούν, όχι μόνο ως πληροφορία, αλλά και ως συντροφιά.^{61,62} Ο ψηφιακός κόσμος της συνεχούς περιήγησης θα τους περιορίσει τη δυνατότητα προσοχής, προϋπόθεση για την ανάπτυξη κριτικής σκέψης. Ο ψηφιακός κόσμος θα τους δώσει έτοιμο το κείμενο της έκθεσης και την περίληψη ενός γραπτού αποσπάσματος, άρα σιγά σιγά θα τους στερήσει τη δυνατότητα διατύπωσης συλλογισμών και κρίσης, ώστε να μπορούν να διακρίνουν το σημαντικό από το δευτερεύον. Θα τους περιορίσει την κρίση να διακρίνουν την πραγματικότητα από τη φαντασίωση, το γεγονός από την ψευδή είδηση. Κι ακόμη χειρότερα, θα τους διευκολύνει να επιλέξουν το ψεύδος από την αλήθεια, γιατί αυτό αγγίζει βαθύτερα το θυμικό τους.

Στο σημείο αυτό αναδύεται μία ισχυρή αντίφαση. Από τη μια προβάλλει ο κίνδυνος ατροφίας της κριτικής σκέψης, και κατ' επέκταση της δημιουργικής ικανότητας του ανθρώπου ώστε να μπορεί να προτείνει λύσεις πέρα κι έξω από αυτές του «μέτριου ειδικού», λύσεις που να χρωματίζονται με την προσωπική δημιουργική πινελιά και ανάλυση της κάθε μοναδικής υποκειμενικότητας. Από την άλλη, ο άνθρωπος θα είναι «αντικαταστάσιμος» από τις μηχανές, όταν δεν έχει τις ικανότητες αυτές. Αυτό που φαίνεται να προβάλλει, μέσα από την αντίφαση αυτήν, ως ένα νέο σημείο ισορροπίας, είναι ο τύπος της κοινωνίας με τεράστιες ανισότητες, όχι μόνον οικονομικές και κοινωνικές, αλλά και «υπαρξιακές».

Μήπως το «θαύμα» της νέας τεχνολογίας, όπως το παρουσιάζουν οι τεχνο-οπτιμιστές της Silicon Valley, συνοδευόμενο με το «μυστήριο» μιας τεχνολογίας που δεν κατανοούμε πλήρως και μιας «εξουσίας» που δεν αναδύεται δημοκρατικά από τις κοινωνίες, επαναφέρει στο προσκήνιο τον *Ιεροξεταστή* του Fyodor Dostoevsky; Από την άλλη, πόσο ευσταθή μπορεί να είναι τέτοια σημεία ισορροπίας;

Η κύρια διακύβευση αυτήν τη στιγμή είναι να δείξουμε, με το δάχτυλο σταθερά τεντωμένο, ποιες είναι εκείνες οι πολιτικές, κοινωνικές και οικονομικές δυναμικές που από τη μία παρέχουν τεράστιες δυνατότητες και ευκαιρίες, υλοποιώντας τα επιτεύγματα της ανθρώπινης νόησης, και από την άλλη δημιουργούν εκείνες τις συνθήκες που σταδιακά αναχαιτίζουν την πορεία προς την πολυδιάστατη υπαρξιακή αυτοπραγμάτωση του ανθρώπινου όντος.

Η κρίσιμη διακύβευση είναι συλλογικά να απαιτήσουμε, ως ενεργοί πολίτες, πολιτικές παρεμβάσεις για θεσμικές θωρακίσεις που δεν θα αφορούν μόνον

⁶¹ Nearly 3 in 4 teens have used AI companions, new national survey finds. (2025, July 16). *Common Sense Media*. <https://www.commonsensemedia.org/press-releases/nearly-3-in-4-teens-have-used-ai-companions-new-national-survey-finds>

⁶² Klein, A. (2025, July 16). *One-Third of Teens Are as 'Satisfied' Talking to a Chatbot as a Real Person*. Education Week. <https://www.edweek.org/technology/one-third-of-teens-are-as-satisfied-talking-to-a-chatbot-as-a-real-person/2025/07>

ατομικά δικαιώματα και προσωπικά δεδομένα, αλλά και μελέτη κινδύνων που απορρέουν από τη χρήση των τεχνολογιών αυτών, κάτι αντίστοιχο που έχει υιοθετηθεί στο πεδίο των ιατρικών και βιολογικών ερευνών.

Το 2018, ο Κινέζος επιστήμονας He Jiankui δημιούργησε τα πρώτα γενετικά τροποποιημένα μωρά, χρησιμοποιώντας την τεχνολογία CRISPR για να τα κάνει ανθεκτικά στον ιό HIV. Η αντίδραση της παγκόσμιας επιστημονικής κοινότητας ήταν άμεση και σκληρή· καταδικάστηκε, φυλακίστηκε και το έργο του καταγγέλθηκε ως μη ηθικά αποδεκτό.⁶³ Γιατί, όμως, είναι ηθικά αποδεκτό να επαφίονται η κοινωνικοποίηση και η καλλιέργεια των γνωστικών ικανοτήτων των παιδιών μας, όλο και περισσότερο, σε αλγοριθμικές προσομοιώσεις του κόσμου και της ίδιας της ζωής; Γιατί είναι ηθικά αποδεκτές *ετερόνομα σχεδιασμένες εξωθεσμικές παρεμβάσεις* –με τη μορφή αλγοριθμικής ντοπαμίνης– στις συναισθηματικές και βιωματικές εμπειρίες, με τις οποίες όπλισε το ανθρώπινο ον η εξέλιξη μέσα από εκατομμύρια χρόνια μεταλλάξεων και φυσικών επιλογών;

Ένα σενάριο κοινωνίας με «πολίτες» διαμορφωμένους με συρρικνωμένη υπαρξιακή υπόσταση αποτελεί βόμβα στα θεμέλια της φιλελεύθερης δημοκρατίας. Η ατροφία βασικών γνωστικών ιδιοτήτων μάς καθιστά ιδιαίτερα ευάλωτους σε χειραγώγηση. Μιας χειραγώγησης με πολύ ιδιαίτερα και «φρέσκα» χαρακτηριστικά. Δεν υπάρχει, πια, λόγος να χειραγωγείς με «ιδεολογίες», άλλωστε κι αυτές «κάπως» έχουν ξεθωριάσει. Χειραγωγείς προσφέροντας *εξάρτηση* από την «ευχαρίστηση». Η συνεχής διέγερση είναι μια πολύ πιο αποτελεσματική μορφή ελέγχου από ό,τι ήταν ποτέ η λογοκρισία. Δεν μας κατηχούν, μας *αναισθητοποιούν*.

Η Hannah Arendt μιλώντας για την «κοινοτυπία του κακού», μας περιγράφει ότι οι μεγάλες εκτροπές δεν δρομολογούνται από δαιμονικές προθέσεις, αλλά χτίζονται σιγά σιγά μέσα από μια αποψιλωμένη σκέψη, μέσα από έλλειψη κριτικής ικανότητας, μέσα από μία ύπουλα ακρωτηριασμένη κοινωνικοποίηση παθητικότητας.⁶⁴ Στην ίδια κατεύθυνση αλλά με μια πιο ποιητική ματιά, ο Thomas S. Eliot γράφει: *«Έτσι τελειώνει ο κόσμος. Όχι με έναν βρόντο, μα μ' έναν λυγμό»*.⁶⁵

⁶³ Normile, D. (2019, December 30). Chinese scientist who produced genetically altered babies sentenced to 3 years in jail. *Science*. <https://www.science.org/content/article/chinese-scientist-who-produced-genetically-altered-babies-sentenced-3-years-jail>

⁶⁴ Arendt, H. (2009). *Ο Αιχμάν στην Ιερουσαλήμ-Μία έκθεση για την κοινοτυπία του κακού*. Νησίδες.

⁶⁵ Eliot, T. S. (2009). *Οι κούφιοι άνθρωποι, Έρημη χώρα και άλλα ποιήματα* (Γ. Σεφέρης, Μετ.). Ίκαρος. (Το πρωτότυπο έργο δημοσιεύθηκε το 1925).

[ΠΩΣ] ΜΠΟΡΕΙ Η ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΝΑ ΕΙΝΑΙ
ΗΘΙΚΗ ΚΑΙ ΣΥΝΝΟΜΗ;
Ιούλιος 2026

Η πρωτοβουλία για την ηθική πλαισίωση της Τεχνητής Νοημοσύνης



Λίλιαν Μήτρου, Σταυρούλα Τσινόρεμα

Συγγραφική συνεισφορά ανά υποκεφάλαιο:

2.1., 2.2., 2.3., 2.4., 2.5. – Λίλιαν Μήτρου

2.6. – Σταυρούλα Τσινόρεμα

2.1. Η αξιόπιστη Τεχνητή Νοημοσύνη

Η συζήτηση για την Τεχνητή Νοημοσύνη έθεσε καταρχάς εκ νέου και μετ' επιτάσεως το ζήτημα του δεοντολογικού προβληματισμού και των δικαιικών ή/και ηθικών αρχών και απαιτήσεων ως προς τον σχεδιασμό και τη χρήση της. Η υιοθέτηση κανόνων από την Ευρωπαϊκή Ένωση (ΕΕ) υπήρξε, όπως ήταν αναμενόμενο, μια σταδιακή διαδικασία. Η ΕΕ ανίχνευσε και οριοθέτησε την προσέγγισή της, προτείνοντας, καταρχήν και καταρχάς, ένα πλαίσιο δεοντολογικών αρχών και κατευθύνσεων, που αποτέλεσαν, άλλωστε, την αποκρυστάλλωση της δημοκρατικής-ανθρωποκεντρικής προσέγγισης που ενυπάρχει στην Πράξη για την Τεχνητή Νοημοσύνη.

Αφετηρία της ενωσιακής επιλογής ήταν η διασφάλιση της εμπιστοσύνης μέσω της επιδίωξης της πολυδιάστατης αξιοπιστίας των συστημάτων ΤΝ. Σύμφωνα με την Ομάδα Εμπειρογνομόνων Υψηλού Επιπέδου για την Τεχνητή Νοημοσύνη (High Level Expert Group – HLEG), που εκπόνησε τις Κατευθυντήριες Γραμμές Δεοντολογίας για Αξιόπιστη Τεχνητή Νοημοσύνη, η εμπιστοσύνη αποτελεί τη συγκολλητική ουσία των κοινωνιών, των κοινοτήτων, των οικονομιών και της βιώσιμης ανάπτυξης. Η εμπιστοσύνη εδράζεται στην πεποίθηση ότι η μεγιστοποίηση των οφελών συνοδεύεται από την πρόληψη και την ελαχιστοποίηση των κινδύνων. Προϋπόθεση εμπέδωσης της εμπιστοσύνης είναι η ύπαρξη ενός σαφούς και ολοκληρωμένου πλαισίου για την επίτευξη της αξιοπιστίας.

Η αξιόπιστη ΤΝ θεμελιώνεται σε τρεις συνιστώσες που αφορούν τον κύκλο ζωής ενός συστήματος και είναι αλληλοσυμπληρούμενες ή και αλληλεπικαλυπτόμενες:

α) Θα πρέπει να είναι σύννομη, δηλαδή να τηρεί όλες τις εφαρμοστέες νομοθετικές και κανονιστικές διατάξεις,

β) θα πρέπει να είναι δεοντολογική, δηλαδή να διασφαλίζει τη συμμόρφωση με δεοντολογικές αρχές και αξίες, και

γ) θα πρέπει να είναι στιβαρή, τόσο από τεχνικής όσο και από κοινωνικής άποψης, διότι, ακόμη κι όταν υπάρχει καλή πρόθεση, τα συστήματα ΤΝ μπορούν να προκαλέσουν ακούσια βλάβη.

2.2. Αρχές της αξιόπιστης Τεχνητής Νοημοσύνης

Η ανάπτυξη και χρήση της Τεχνητής Νοημοσύνης με αξιόπιστο τρόπο βασίζεται καταρχάς σε τέσσερις αρχές:

α) *Αρχή του σεβασμού της ανθρώπινης αυτονομίας*: Η αρχή αυτή συνίσταται στη διασφάλιση της δυνατότητας των ανθρώπων να διατηρούν πλήρη και αποτελεσματική αυτοδιάθεση και να είναι σε θέση να συμμετέχουν στη δημοκρατική διαδικασία. Οι κατευθυντήριες αρχές επικεντρώνονται στην αρχή του ανθρωποκεντρικού σχεδιασμού, έτσι ώστε τα συστήματα αυτά να αυξάνουν, να συμπληρώνουν και να ενισχύουν τις ανθρώπινες γνωστικές, κοινωνικές και πολιτισμικές δεξιότητες, και να αποτρέπουν την καθυστέρηση, τον εξαναγκασμό, την εξαπάτηση, τη χειραγώγηση, τον κοινωνικό προγραμματισμό ή την αγελοποίηση των ανθρώπων.

β) *Η αρχή της πρόληψης βλάβης*: Σύμφωνα με την αρχή αυτήν, τα συστήματα ΤΝ δεν πρέπει να προκαλούν ή να επιδεινώνουν τυχόν βλάβες, στις οποίες συμπεριλαμβάνονται και οι δυσμενείς επιπτώσεις στον τρόπο διαβίωσης ή στην πολιτισμική υπόσταση ατόμων ή ομάδων. Συμπεριλαμβάνοντας την προστασία της ανθρώπινης αξιοπρέπειας, καθώς και της πνευματικής και σωματικής ακεραιότητας, η αρχή αυτή απαιτεί να μην επηρεάζει ένα σύστημα ΤΝ αρνητικά έναν άνθρωπο. Η αρχή δεν καταλαμβάνει μόνο την υλική βλάβη, αλλά και τις περιπτώσεις που οι αρνητικές επιπτώσεις συνδέονται με την ασυμμετρία εξουσίας ή πληροφόρησης.

γ) *Η αρχή της δικαιοσύνης*: Η αρχή αυτή επιτάσσει να γίνεται η ανάπτυξη, η εγκατάσταση και η χρήση των συστημάτων κατά δίκαιο τρόπο. Οι συντάκτες των Κατευθυντήριων Γραμμών αντιλαμβάνονται τη δικαιοσύνη καταρχάς στην ουσιαστική διάστασή της, υπάγοντας σε αυτήν, πρώτον, την εξασφάλιση της ισότιμης και δίκαιης κατανομής τόσο των ωφελειών όσο και του κόστους, και, δεύτερον, την προστασία από αθέμιτη μεροληψία, διακρίσεις και στιγματισμό. Στοιχείο της ουσιαστικής δικαιοσύνης είναι η προστασία από εξαπάτηση ή ο περιορισμός της ελευθερίας των χρηστών. Στη διαδικαστική προσέγγιση της δικαιοσύνης υπάγεται η δυνατότητα αμφισβήτησης και αποτελεσματικής έννομης προστασίας έναντι αποφάσεων που λαμβάνονται από συστήματα ΤΝ και από τους ανθρώπους που τα χειρίζονται.

δ) *Η αρχή της επεξηγησιμότητας*: Πρόκειται για μια αρχή καθοριστικής σημασίας, από την οποία απορρέει η απαίτηση της διαφάνειας και της επεξήγησης των διαδικασιών, των δυνατοτήτων, των σκοπών και των αποφάσεων των συστημάτων αυτών. Η ΤΝ χαρακτηρίζεται ως «black box». Πράγματι, φαίνεται αδύνατο να χαρτογραφηθεί η διαδικασία λήψης αποφάσεων αυτών των πολύπλοκων δικτύων. Μπορούμε να παρατηρήσουμε τις εισόδους και τις εξόδους τους, αλλά δεν μπορούμε να ιχνηλατήσουμε και να αποδείξουμε ποια χαρακτηριστικά ή ποιοι συνδυασμοί χαρακτηριστικών ήταν καθοριστικοί για το τελικό αποτέλεσμα. Ένα μοντέλο συχνά καταλήγει σε ένα συμπέρασμα χωρίς καμία εξήγηση. Αυτός είναι ο λόγος για τον οποίο η διαδικασία λήψης αποφάσεων της ΤΝ έχει χαρακτηριστεί ως «μαύρο κουτί», μια μεταφορά που υποδηλώνει ένα σύστημα του οποίου οι εσωτερικές λειτουργίες είναι μυστηριώδεις.⁶⁶ Είναι ιδιαίτερα δυσχερείς η ιχνηλασιμότητα και η επεξήγηση των λόγων για τους οποίους τα «στοιχεία εισόδου» έχουν οδηγήσει στα συγκεκριμένα «στοιχεία εξόδου», ένα συγκεκριμένο αποτέλεσμα, μια σύσταση ή μια απόφαση, περιορίζοντας με τον τρόπο αυτόν και την ελεγχιμότητα του συστήματος.

Οι αρχές αυτές απορρέουν από τα θεμελιώδη δικαιώματα, όπως κατοχυρώνονται στην ευρωπαϊκή και την εθνική δικαιοταξία (Συνθήκες ΕΕ, Ευρωπαϊκή Σύμβαση Δικαιωμάτων του Ανθρώπου, Χάρτης Θεμελιωδών Δικαιωμάτων της ΕΕ, Σύνταγμα). Κοινή αφετηρία και βάση των δικαιωμάτων είναι η ανθρώπινη αξιοπρέπεια που επιβάλλει την ανθρωποκεντρική προσέγγιση, ενώ κεντρική θέση καταλαμβάνουν η ελευθερία του ατόμου, η ισότητα, η απαγόρευση των διακρίσεων, η αλληλεγγύη, δικαιώματα των πολιτών, όπως το δικαίωμα στη χρηστή διοίκηση, ο σεβασμός της δημοκρατίας, της Δικαιοσύνης και του κράτους δικαίου.

⁶⁶ Vorras, A., Mitrou, L. (2021). Unboxing the Black Box of Artificial Intelligence: Algorithmic Transparency and/or a Right to Functional Explainability. In: T. Synodinou, P. Jouglex, C. Markou, T. Prastitou-Merdi (Eds.), *EU Internet Law in the Digital Single Market* (pp. 247-264). Springer, Cham.

2.3. Οι απαιτήσεις σε σχέση με την αξιόπιστη Τεχνητή Νοημοσύνη

Από τις τέσσερις βασικές αρχές που εκτέθηκαν παραπάνω απορρέουν επτά ειδικότερες απαιτήσεις:

α) *Ανθρώπινη παρέμβαση και εποπτεία*, συμπεριλαμβανομένων των θεμελιωδών δικαιωμάτων, της ανθρώπινης παρέμβασης και της ανθρώπινης εποπτείας. Κατά τις Κατευθυντήριες Γραμμές Δεοντολογίας, πρωταρχικό στάδιο της ανθρώπινης εποπτείας είναι η διενέργεια εκτίμησης επιπτώσεων για τα θεμελιώδη δικαιώματα, που πρέπει να πραγματοποιείται πριν από την ανάπτυξη τέτοιων συστημάτων, ώστε οι εντοπιζόμενοι κίνδυνοι είτε να μειώνονται είτε να δικαιολογούνται ως απαραίτητοι, σε μια δημοκρατική κοινωνία, προκειμένου να γίνονται σεβαστά τα δικαιώματα και οι ελευθερίες των άλλων. Η ανθρώπινη εποπτεία μπορεί να συμπεριλάβει πολλούς μηχανισμούς διακυβέρνησης, όπως η προσέγγιση, στην οποία ο άνθρωπος είναι εκείνος που παρεμβαίνει (human-in-the-loop – HITL), εκείνος που εποπτεύει (human-on-the-loop – HOTL) ή εκείνος που ελέγχει (human-in-command – HIC). Η πιο ενδιαφέρουσα και ταυτόχρονα η πιο δυσερμήνευτη και δυσεφάρμοστη διάσταση είναι η παροχή της δυνατότητας της ανθρώπινης παρέμβασης, δηλαδή ο εφοδιασμός των χρηστών με γνώσεις και εργαλεία για να κατανοούν και να αλληλεπιδρούν με τα συστήματα ΤΝ σε ικανοποιητικό βαθμό και, όπου είναι δυνατό, να έχουν τη δυνατότητα αξιολόγησης ή αμφισβήτησης του συστήματος μέσα σε λογικά πλαίσια.

β) *Τεχνική στιβαρότητα και ασφάλεια*, συμπεριλαμβανομένων της ανθεκτικότητας σε επιθέσεις και της προστασίας, του εφεδρικού σχεδίου και της γενικής ασφάλειας, της ακρίβειας, της αξιοπιστίας και της αναπαραγωγιμότητας. Στενά συνδεδεμένη με την αρχή της *πρόληψης βλάβης*, η τεχνική στιβαρότητα προϋποθέτει την ανάπτυξη συστημάτων ΤΝ με προληπτική προσέγγιση των κινδύνων και με τέτοιον τρόπο ώστε να συμπεριφέρονται αξιόπιστα με τον αναμενόμενο τρόπο. Σε μια εποχή έξαρσης των κινδύνων κυβερνοασφάλειας, η λήψη μέτρων ασφάλειας και ανθεκτικότητας αναδεικνύεται σε κρίσιμη απαίτηση. Στις σημαντικές απαιτήσεις που διασφαλίζουν τα δικαιώματα και την ορθή και ασφαλή χρήση εντάσσονται η ακρίβεια και η αξιοπιστία.

γ) *Ιδιωτική ζωή και διακυβέρνηση των δεδομένων*, συμπεριλαμβανομένων του σεβασμού της ιδιωτικής ζωής, της ποιότητας και της ακεραιότητας των δεδομένων και της πρόσβασης στα δεδομένα. Η προστασία και εγγύηση των δικαιωμάτων της ιδιωτικής ζωής και των προσωπικών δεδομένων αναδεικνύεται ως μείζονος σημασίας απαίτηση, καθώς η χρήση της ΤΝ μπορεί να επιτείνει υπαρκτούς κινδύνους ή να παράγει νέους, καθώς, π.χ., επιτρέπει τη συσχέτιση πληροφοριών ή τη συναγωγή συμπερασμάτων σε έκταση και ένταση που δεν υφίστατο έως τώρα. Ιδιαίτερο ενδιαφέρον παρουσιάζει ότι οι Κατευθυντήριες Γραμμές Δεοντολογίας συνδέουν την απαίτηση αυτή με τις απαιτήσεις για ποιότητα και ακεραιότητα των δεδομένων, ιδίως λόγω της πιθανότητας να περιέχουν κοινωνικά κατασκευασμένες προκαταλήψεις ή ανακρίβειες, σφάλματα και λάθη.

δ) *Διαφάνεια*, συμπεριλαμβανομένης της ανιχνευσιμότητας, της επεξηγησιμότητας και της επικοινωνίας. Σε στενή συνάρτηση με την αρχή της επεξηγησιμότητας, η απαίτηση της διαφάνειας αναφέρεται σε όλα τα στοιχεία που σχετίζονται με ένα σύστημα ΤΝ: τα δεδομένα, το σύστημα και τα επιχειρηματικά μοντέλα. Τα σύνολα δεδομένων που χρησιμοποιούνται και των διαδικασιών που οδηγούν στα «στοιχεία εξόδου» ενός συστήματος ΤΝ θα πρέπει να τεκμηριώνονται στο βέλτιστο δυνατό επίπεδο, ώστε να καθίστανται δυνατές η ικνηλασιμότητα και η αύξηση της διαφάνειας. Η επεξηγησιμότητα, κρίσιμη παράμετρος της διαφάνειας, αφορά τη δυνατότητα επεξήγησης τόσο των τεχνικών διεργασιών ενός συστήματος ΤΝ όσο και των σχετικών ανθρώπινων αποφάσεων (π.χ., τομείς εφαρμογής ενός συστήματος ΤΝ).

ε) *Πολυμορφία, απαγόρευση των διακρίσεων και δικαιοσύνη*, συμπεριλαμβανομένων της αποφυγής της αθέμιτης μεροληψίας, της προσβασιμότητας και του καθολικού σχεδιασμού, και της συμμετοχής των ενδιαφερόμενων μερών. Η απαίτηση αυτή αγγίζει τον πυρήνα των δημοκρατικών αρχών και των ανθρωπίνων δικαιωμάτων, καθώς αφορά την αρχή της ανθρωπίνης αξιοπρέπειας, της ισότητας, αλλά και της δικαιοσύνης (fairness). Με αφετηρία τη διασφάλιση της συμπεριληπτικότητας και της πολυμορφίας, η διατύπωση αυτής της αρχής επιδιώκει την αποτροπή της αναπαραγωγής και επίτασης παλαιών ή της δημιουργίας νέων μορφών μεροληψίας, που μπορεί να οδηγήσει σε ακούσιες άμεσες (και έμμεσες) προκαταλήψεις και διακρίσεις κατά ορισμένων ομάδων ή ανθρώπων, επιδεινώνοντας ενδεχομένως προϋφιστάμενες προκαταλήψεις, τον στιγματισμό και την περιθωριοποίηση.

στ) *Κοινωνική και περιβαλλοντική ευημερία*, συμπεριλαμβανομένων της βιωσιμότητας και της φιλικότητας προς το περιβάλλον, των επιπτώσεων στην κοινωνία και τη δημοκρατία. Υπό την απαίτηση αυτήν, οι συντάκτες των Κατευθυντήριων Γραμμών συνοψίζουν τρεις κρίσιμες δικαιοπολιτικές επιλογές και κατευθύνσεις:

- 1) διασφάλιση μιας βιώσιμης και φιλικής προς το περιβάλλον ΤΝ,
- 2) αποτροπή των δυσμενών κοινωνικών επιπτώσεων στις κοινωνικές σχέσεις και δεξιότητες, και εν γένει στην ευημερία των ανθρώπων, και
- 3) υπεύθυνη, σύννομη και θεμιτή χρήση ΤΝ ώστε να αντιμετωπίζονται κίνδυνοι που αφορούν την ίδια τη δημοκρατία και τις διαδικασίες της.

ζ) Λογοδοσία, συμπεριλαμβανομένων της ελεγχιμότητας, της ελαχιστοποίησης και της γνωστοποίησης των αρνητικών επιπτώσεων, των αντισταθμιστικών ρυθμίσεων και της έννομης προστασίας. Η απαίτηση της λογοδοσίας γίνεται αντιληπτή ως συμπληρωματική, αλλά εξίσου ουσιώδης με τις άλλες. Περιλαμβάνει τη δυνατότητα αξιολόγησης των αλγορίθμων, των δεδομένων και των διαδικασιών σχεδιασμού, ώστε, μεταξύ άλλων, να διασφαλίζονται η γνωστοποίηση και η ελαχιστοποίηση των αρνητικών επιπτώσεων. Η υιοθέτηση αντισταθμιστικών ρυθμίσεων και εγγυήσεων είναι σημαντική απαίτηση, αλλά πρέπει να συνοδεύεται από την πρόβλεψη προσβάσιμων μηχανισμών που να εξασφαλίζουν κατάλληλη έννομη προστασία σε περιπτώσεις άδικων, δυσμενών επιπτώσεων.

Οι αρχές και οι απαιτήσεις που αποτυπώθηκαν στις Κατευθυντήριες Γραμμές Δεοντολογίας για μια αξιόπιστη ΤΝ αποτέλεσαν τον βασικό άξονα στην πρόταση Κανονισμού που οδήγησε στην υιοθέτηση της Πράξης για την Τεχνητή Νοημοσύνη. Η Πράξη για την Τεχνητή Νοημοσύνη συνιστά ένα ενδιαφέρον παράδειγμα για την εξέλιξη μίας ρυθμιστικής πρωτοβουλίας που έχει ως (συν)αφετηρία τις αξίες και αρχές που διατυπώθηκαν καταρχάς ως κατευθύνσεις.

2.4. Η Σύσταση της UNESCO για την Ηθική της Τεχνητής Νοημοσύνης

Δεκάδες, αν όχι εκατοντάδες, είναι οι διεθνείς, δημόσιοι και ιδιωτικοί φορείς που έχουν αναπτύξει δεοντολογικές αρχές και αντίστοιχα πλαίσια για την Τεχνητή Νοημοσύνη. Εξ αυτών αξίζει να αναδείξουμε τη Σύσταση για την Ηθική της Τεχνητής Νοημοσύνης που υιοθέτησε η UNESCO τον Νοέμβριο του 2021. Οι αξίες που αναδεικνύει η Σύσταση είναι:

- α) ο σεβασμός, η προστασία και η προώθηση των ανθρωπίνων δικαιωμάτων, των θεμελιωδών ελευθεριών και της ανθρώπινης αξιοπρέπειας,
- β) η προστασία του περιβάλλοντος και η ευημερία του οικοσυστήματος,
- γ) η διασφάλιση της ποικιλομορφίας και της συμπεριληπτικότητας, και
- δ) η εγγύηση της ζωής σε ειρηνικές, δίκαιες και διασυνδεδεμένες κοινωνίες.

Από τις αξίες αυτές η UNESCO συνάγει τις ακόλουθες αρχές:

- α) Αναλογικότητα και πρόληψη βλάβης,
- β) ασφάλεια (safety and security),
- γ) δικαιοσύνη και μη διάκριση,
- δ) βιωσιμότητα,
- ε) δικαίωμα στην ιδιωτικότητα και προστασία,
- στ) ανθρώπινη εποπτεία και απόφαση,
- ζ) διαφάνεια και επεξηγησιμότητα,
- η) υπευθυνότητα και λογοδοσία,
- θ) ενημέρωση και γραμματισμός,
- ι) συνεργασία των εμπλεκομένων και προσαρμοστική διακυβέρνηση.⁶⁷

⁶⁷ United Nations Educational, Scientific and Cultural Organization (UNESCO) (2021). Recommendation on the Ethics of Artificial Intelligence (p. 17 ff.). <https://unesdoc.unesco.org/ark:/48223/pf0000381137>

2.5. Κανονιστικό πλαίσιο και όχι μόνο αρχές και κατευθύνσεις ηθικής

Η Ομάδα Εμπειρογνομόνων Υψηλού Επιπέδου για την Τεχνητή Νοημοσύνη υπογράμμισε ότι η συζήτηση για τις ηθικές αρχές και δεσμεύσεις «μπορεί να μας βοηθήσει να κατανοήσουμε τον τρόπο με τον οποίο η ανάπτυξη, η εγκατάσταση και η χρήση της ΤΝ μπορεί να επηρεάζουν τα θεμελιώδη δικαιώματα και τις υποκείμενες αξίες τους, ενώ μπορεί να συμβάλει στην παροχή πιο λεπτομερών καθοδήγησης κατά την εξέταση του τι θα πρέπει να κάνουμε αντί του τι μπορούμε να κάνουμε (επί του παρόντος) με την τεχνολογία».⁶⁸ Επιβεβαιώνοντας την ενεργή κινητοποίηση για την πλαisiώση της ΤΝ με «κανόνες», ο αριθμός των ηθικών αρχών και των δεοντολογικών κατευθύνσεων/οδηγιών αυξάνεται διαρκώς, είτε σε επίπεδο διεθνών οργανισμών και φορέων είτε σε επίπεδο ιδιωτών.

Οι προσεγγίσεις είναι –και εν προκειμένω– διαφορετικές. Αρκετοί υποστηρίζουν την αναγκαιότητα αυτών των «ηθικών πλαισίων» ως προαπαιτούμενα της κατανόησης και ως βάση της προστασίας. Ταυτόχρονα, ωστόσο, η ανάδειξη των κανόνων ηθικής και δεοντολογίας ως βασικού ρυθμιστικού εργαλείου σχετίζεται και με την προσπάθεια να διατυπωθούν μεν οι δικαιοπολιτικοί στόχοι και οι συναφείς δεσμεύσεις, όπως η απαγόρευση των διακρίσεων ή η διαφάνεια, αλλά να μη λειτουργήσουν οι κανόνες αυτοί ως εμπόδιο στην εξέλιξη των νέων τεχνολογιών. Η επίκληση του «ηθικού πλαισίου» σχεδιασμού και χρήσης που περιβάλλεται έναν οιονεί κανονιστικό μανδύα, αυτόν της αυτοδέσμευσης και συχνά της διατύπωσης κανόνων από αυτούς που είναι «κοντά» στην τεχνολογία, έχει συχνά ως επακόλουθο την απόπειρα της υπέρβασης των όποιων υφιστάμενων κανονιστικών ορίων και τη δυνατότητα να τα αγνοήσει κανείς. Ενέχει, δε, την επιδίωξη να αποφευχθεί –ή έστω να καθυστερήσει– η δεσμευτική κανονιστική παρέμβαση του νομοθέτη ή να περιοριστεί το εύρος της. Δεν είναι λίγοι εκείνοι που εξακολουθούν να αναφέρονται στην προσπάθεια να διατηρηθεί η ρύθμιση της ΤΝ στο

⁶⁸ High-Level Expert Group on Artificial Intelligence. (2019). Κατευθυντήριες Γραμμές Δεοντολογίας για Αξιοπιστή Τεχνητή Νοημοσύνη (Ethics guidelines for trustworthy AI) (p. 13). European Commission.

επίπεδο της ηθικής δέσμευσης με τον χαρακτηρισμό «ethics washing» ή και «ethics shopping».⁶⁹

Η συζήτηση για τις δεοντολογικές αρχές και τους κώδικες ηθικής είναι πρόδρομος της υιοθέτησης δεσμευτικών κανόνων, καθώς αναδεικνύονται οι απαραίτητες κατευθύνσεις για το περιεχόμενο του νόμου. Όπως υποστήριζε η Sara Baase ήδη το 1997, οι «ηθικές αρχές» καλύπτουν το κενό στο χρονικό διάστημα που μεσολαβεί μεταξύ της εξέλιξης της τεχνολογίας, η οποία δημιουργεί νέα προβλήματα, και του χρόνου της υιοθέτησης λογικών και αποδοτικών νομοθετικών λύσεων.⁷⁰

Θα ήταν κοντόθωρο να παραβλέψουμε τη σημασία της διαμόρφωσης και εμπέδωσης ηθικών δεσμεύσεων κατά τον σχεδιασμό και τη χρήση συστημάτων ΤΝ. Η ηθική της ΤΝ συνιστά πρωτίστως μία ηθική της τεχνολογίας (technology ethics).⁷¹ Οι αξιακές παράμετροι και οι αντίστοιχες δεσμεύσεις επηρεάζουν (ή μπορούν να επηρεάσουν) τις επιλογές του τεχνικού σχεδιασμού και, βέβαια, τους σκοπούς που πρόκειται να υποστηρίξει ένα σύστημα ΤΝ. Ωστόσο, καθώς οι τεχνολογίες ΤΝ ενσωματώνονται ως εργαλεία για τη βελτιστοποίηση της αποτελεσματικότητας, εξυπηρετώντας ποικίλες οικονομικές, διοικητικές, κοινωνικές στοχεύσεις, (θα) παράγουν αποτελέσματα και αποφάσεις που δίνουν προτεραιότητα σε μερικές αξίες έναντι άλλων και (θα) επηρεάζουν άμεσα πρόσωπα και ομάδες.

Η εστίαση στην καθιέρωση και αποδοχή μιας «ηθικής της Τεχνητής Νοημοσύνης» ή ο προσανατολισμός στην «ηθική Τεχνητή Νοημοσύνη», μία διατύπωση που αποδίδει την ανθρωποκεντρική (ή και ανθρωπομορφική) πρόσληψη της ΤΝ, είναι μεν αναγκαίο στάδιο της κατανόησης των ζητημάτων και της κατεύθυνσης ως προς τη λύση τους, αλλά αγνοεί ότι πρόκειται για σύνθετα συστήματα που συναποτελούνται από τεχνολογίες, ανθρώπους και τις ανάγκες και στόχους αυτών.

⁶⁹ Πρόκειται για δύο έννοιες που χρησιμοποιούνται για να ορίσουν το πώς εταιρείες τεχνολογίας διαχειρίζονται την ηθική της Τεχνητής Νοημοσύνης. Ο όρος «ethics washing» («ηθικό ξέπλυμα») αναφέρεται στην πρακτική οργανισμών ή εταιρειών τεχνολογίας να προβάλλουν μια εικόνα ηθικής υπευθυνότητας, χωρίς αυτό να συνοδεύεται από ουσιαστικές αλλαγές στον σχεδιασμό, στην ανάπτυξη ή στη χρήση αλγορίθμων. Πρόκειται, δηλαδή, για μια μορφή επικοινωνιακής στρατηγικής που λειτουργεί ως «ηθική βιτρίνα» και όχι ως πραγματική ενσωμάτωση ηθικών αρχών στα συστήματα Τεχνητής Νοημοσύνης. Ο όρος «ethics shopping» περιγράφει την επιλεκτική υιοθέτηση ηθικών αρχών από οργανισμούς ή εταιρείες, οι οποίες επιλέγουν εκείνα τα ηθικά πρότυπα που εξυπηρετούν τους επιχειρηματικούς ή τεχνολογικούς τους στόχους, ενώ παραλείπουν ή αποδυναμώνουν άλλα που θα επέβαλλαν μεγαλύτερους περιορισμούς ή λογοδοσία. Βλ. Wagner, B. (2018). Ethics as an escape from regulation. From “ethics-washing” to ethics-shopping? In E. Bayamlioglu, I. Baraliuc, L. Janssens (Eds.), *Being Profiled: Cogitas Ergo Sum. 10 Years of ‘Profiling the European Citizen’* (p. 84 ff.). Amsterdam University Press.

⁷⁰ Baase, S. (2013). *A Gift of Fire: Social, Legal and Ethical Issues in Computing* (4th ed.), (p. 26 ff, p. 403 ff.). Pearson.

⁷¹ Η ηθική της τεχνολογίας αναφέρεται στο πώς οι επιστημονικές και τεχνολογικές εξελίξεις επηρεάζουν τους ανθρώπους, την κοινωνία και το περιβάλλον. Θέτει το ερώτημα αν η καινοτομία είναι δίκαιη, διαφανής, ασφαλής και αν σέβεται τα ανθρώπινα δικαιώματα. Η ηθική της τεχνολογίας ασχολείται με ζητήματα όπως η ιδιωτικότητα, οι διακρίσεις και η μεροληψία, η λογοδοσία, η ανθρώπινη εποπτεία και οι ευρύτερες κοινωνικές συνέπειες των αναδυόμενων τεχνολογιών.

Βέβαια, η επίτευξη της σχεδίασης ενός τεχνολογικού προϊόντος που ενσωματώνει αρχές δεν απαιτεί μόνο τη σχεδιαστική δεινότητα, αλλά και μία αναστοχαστική κατανόηση των αντίστοιχων αξιών και επιλογών και πώς αυτές επιδρούν στη ζωή και τα δικαιώματα των ατόμων. Τα συστήματα ΤΝ αντανakλούν τις αντιλήψεις και τις προτεραιότητες αυτών που τα αναπτύσσουν, αλλά ενδέχεται να μην ευθυγραμμίζονται με τις συλλογικές αξίες, δημοκρατικές και συνταγματικές αξίες. Η αξιολόγηση των κινδύνων, η στάθμιση αντιπαθέμενων δικαιωμάτων και συμφερόντων, η επίλυση των συγκρούσεων, η θέση ορίων εναπόκειται στον νομοθέτη, εθνικό και υπερεθνικό.

Η ηθική της ΤΝ μπορεί να προαναγγείλει κανονιστικές επιλογές και να τις υποστηρίξει πριν και μετά την υιοθέτηση του κανονιστικού πλαισίου. Οι αυτορρυθμιστικές προσπάθειες και δεσμεύσεις μπορούν μόνο επικουρικά και ενισχυτικά να λειτουργήσουν και δεν νοείται να υποκαθιστούν τη νομοθετική ρύθμιση, καθώς λείπουν δύο καθοριστικά στοιχεία: α) η δημοκρατική νομιμοποίηση και συνακόλουθα η αντιπροσωπευτικότητα ως προς τη στάθμιση των συγκρουόμενων αξιών, αρχών, δικαιωμάτων και επιδιώξεων,⁷² και β) οι μηχανισμοί ελέγχου της εφαρμογής και επιβολής των κανόνων αυτών.

Η τεχνολογία ή ακριβέστερα οι συνέπειές της επιβάλλουν στον νομοθέτη να λάβει μια σειρά από σύνθετες αποφάσεις: εάν και πότε να παρέμβει, με ποιο είδος κανονιστικής παρέμβασης, ποιοι είναι οι εμπλεκόμενοι, πόσο πρέπει να διαρκέσει η κανονιστική παρέμβαση. Ακόμη κι αν δεν αμφισβητούνται η αναγκαιότητα και η πρωτοκαθεδρία της νομοθετικής ρύθμισης, η «ηθική δέσμευση» για την υπεύνη και ηθική χρήση παραμένει χρήσιμη, αν όχι αναγκαία. Τη σχέση Ηθικής και Δικαίου και το ερώτημα «αν απαιτείται η νομοθεσία, επαρκεί η εφαρμογή της ή μήπως χρειάζεται, παρ' όλα αυτά, και μια ηθική της ΤΝ» εξετάζει η Σταυρούλα Τσινόρεμα στην ακόλουθη ενότητα 2.6.

Στο πλαίσιο αυτό, μπορεί να εξετάσει κανείς μοντέλα συρρύθμισης ή ελεγχόμενης αυτορρύθμισης. Στα μοντέλα αυτά ο ουσιαστικός νομιμοποιητικός λόγος και ρόλος της κρατικής παρέμβασης αφορά τη θέσπιση του πλαισίου ουσιαστικών και διαδικαστικών εγγυήσεων, που διασφαλίζουν με τη σειρά τους την ισορροπία συμφερόντων κατά την αυτορρύθμιση. Ο νομοθέτης ταυτόχρονα δημιουργεί με τον τρόπο αυτόν ερεθίσματα και δίνει κίνητρα στους φορείς της αυτορρύθμισης, ώστε να λαμβάνουν υπ' όψιν το δημόσιο συμφέρον και τις αξιακές επιλογές της κοινωνίας, στην οποία ενεργοποιούνται.

⁷² Η αχίλλειος πτέρνα πολλών αυτορρυθμιστικών πρωτοβουλιών εντοπίζεται στο γεγονός ότι η διαδικασία και συχνά τα αποτελέσματα επικαθορίζονται από την επιρροή κρατών και παραγόντων με δεσπόζουσα θέση στην εξέλιξη και την αγορά της σχετικής τεχνολογίας και των εφαρμογών της.

2.6. Η Ηθική στην Τεχνητή Νοημοσύνη

2.6.1. Ηθική και Δίκαιο

Η επιστήμη και η τεχνολογία μάς λένε πώς να επιτύχουμε έναν στόχο, αλλά δεν μπορούν να μας πουν εάν πρέπει να επιδιώξουμε αυτόν τον στόχο, κατά πόσον, δηλαδή, ο στόχος είναι ορθός. Καθώς η τεχνολογία της Τεχνητής Νοημοσύνης εξελίσσεται με βιβλικούς ρυθμούς, εγείρονται κρίσιμα ερωτήματα για το ποιες από τις νέες δυνατότητες που δημιουργούνται για την πράξη μας πρέπει να αποκλεισθούν, ποιες πρέπει να αναπτυχθούν και υπό ποιους όρους.

Οι αποφάσεις σχετικά με την ανάπτυξη και την αξιοποίηση της ΤΝ εμπλέκουν αξιολογικές κρίσεις, οι οποίες συνδέονται με την αποτίμηση επιδιωκόμενων στόχων, όπως είναι η αντιμετώπιση ασθενειών και η προστασία της υγείας, η κοινωνική πρόνοια, η ανθρώπινη παραγωγικότητα, η βελτίωση της σχέσης πολίτη και κρατικών υπηρεσιών κοκ. Η αποτίμηση των στόχων, δε, βασίζεται σε αρχές και κανόνες, που με τη σειρά τους απαιτούν ηθικο-πρακτικό συλλογισμό και κανονιστική δικαιολόγηση. Ποιος ο ρόλος των ηθικών εκτιμήσεών μας και γιατί έχουν σημασία σε αυτές τις αποφάσεις;

Όταν αναφερόμαστε στην Ηθική, παραπέμπουμε σε ένα είδος κανονιστικής αποτίμησης που απαντά στο ερώτημα τι πρέπει να πράττεται. Διερευνά το κριτήριο της ορθοπραξίας, που προσδιορίζει το *ορθό και το εσφαλμένο των πράξεων*. Πρόκειται για εκείνο το είδος κανονιστικού λόγου που, σε αντιδιαστολή προς τον αιτιακό-εξηγητικό συλλογισμό, δεν αποδίδει τις αιτίες γένεσης ή μεταβολής φαινομένων, συμπεριλαμβανομένων των πράξεών μας ως φυσικών φαινομένων, αλλά αναζητά και δικαιολογεί κριτήρια αποτίμησης του ορθού ή του εσφαλμένου χαρακτήρα των πράξεων βάσει λόγων. Όταν, λοιπόν, δηλώνουμε πως ένα πρόσωπο έχει κάνει κάτι που όφειλε να μην είχε κάνει, αναγνωρίζουμε μια πρωτοπρόσωπη αυτοκατανόηση του τι σημαίνει να είναι κανείς υποκείμενο ικανό για αυτοουργία. Όταν αποτιμούμε την πράξη ενός προσώπου, δεν περιγράφουμε μια απλή φυσική διαδικασία, αλλά χαρακτηρίζουμε κάτι που υπόκειται σε δικαιολόγηση, το οποίο ανήκει στον «λογικό χώρο» της λογοδοσίας, του να παρέχονται και να απαιτούνται λόγοι υπέρ ή κατά αυτού.

Ορισμένα είδη λόγων αφορούν στην ιδιαίτερη αποτίμηση των πράξεων, που λαμβάνει υπ' όψιν τα δυνητικά αποτελέσματα που έχουν για τους άλλους. Πρόκειται για τον ηθικό τύπο λογοδοσίας, ο οποίος αναγνωρίζει την αυτοουργία (ελευθερία) των άλλων ως δεσμευτικό όριο των πράξεών μας, όταν επιδιώκουμε να πραγματώνουμε τους σκοπούς μας και το δικό μας βιοτικό αγαθό. Επιτάσσει ποια μορφή πρέπει να λαμβάνουν οι υποκειμενικές αρχές που καθοδηγούν την πράξη μας, εάν πρόκειται να είναι συμβατές με την καθολική αξία της ελευθερίας όλων.

Η ηθική (αυτο)κατανόησή μας συνδέεται με την κοινή εκείνη ανθρώπινη ιδιότητά μας να πράττουμε και να υποβάλλουμε τις πράξεις μας στον έλεγχο κανονιστικών λόγων. Ο ηθικός αυτοκαθορισμός του καθενός/της καθεμίας συνδέεται με την απαίτηση να ακολουθεί και να δεσμεύεται από εκείνες τις αρχές που θέτει ο ίδιος/η ίδια με βάση τη δική τους αμερόληπτη κρίση και, κατά τούτο, ικανές να ισχύουν από τη σκοπιά όλων. Πρόκειται για την αξίωση που προβάλλουμε, ως κοινωνικά υποκείμενα, να έχουμε μια «σφαίρα πράξης», όπου ο καθένας σέβεται αμοιβαίως τη δυνατότητα πράξης (ελευθερία) των άλλων. Υπό αυτήν την οπτική, η Ηθική αποτελεί εκείνον τον τύπο πρακτικού συλλογισμού που αφορά στον τρόπο με τον οποίο συγκροτούμε έναν κόσμο διαπροσωπικών σχέσεων μέσω αυτοελέγχου της ατομικής βούλησής μας, ώστε να πράττουμε βάσει λόγων (αρχών) που είναι δυνατόν να μοιραστούν οι άλλοι. Πρόκειται για ένα είδος αυτοπεριορισμού της ατομικής ελευθερίας από τη σκοπιά της ελευθερίας όλων.

Περαιτέρω, και συναφώς, τίθεται το ερώτημα πώς πρέπει να διαμορφώνουμε τις μεταξύ μας σχέσεις, ώστε να υπάρχουν εγγυήσεις που εξασφαλίζουν την ισότιμη επιδίωξη των προσωπικών βιοτικών στόχων μας υπό συνθήκες ελευθερίας και ισότητας. Σε αυτό το σημείο, η Ηθική συναντά το Δίκαιο και την Πολιτική. Η διάσταση που το Δίκαιο και η Πολιτική προσθέτουν στην Ηθική είναι ότι το βασικό ηθικό ερώτημα του θεμιτού και του αθέμιτου συνδέεται με μια περαιτέρω, εξωτερική συνθήκη: Κρίνεται από τη σκοπιά του δικαιολογημένου χαρακτήρα του ελέγχου της Πολιτείας και της άσκησης του δημοσίου εξαναγκασμού. Σε αντιδιαστολή προς την Ηθική, η δεσμευτικότητα των κανόνων Δικαίου δεν ανάγεται μόνο στον τρόπο σχηματισμού της κρίσης ή της γνώμης των προσώπων περί του πρακτέου, αλλά ερείδεται σε συλλογικά επιτακτικές αποφάσεις και ρυθμίσεις που δεσμεύουν εξωτερικά τους αποδέκτες τους, δηλαδή ανεξάρτητα από τα εσωτερικά κίνητρά τους, και η συμμόρφωση με τους οποίους υπόκειται σε δημόσιο έλεγχο και εξαναγκασμό.

Η θεμελιώδης αρχή της Ηθικής είναι καθολικευτική, κατά το ότι μας λέει ποια μορφή πρέπει να έχουν οι πράξεις μας ώστε να είναι συμβατές με την καθολική αξία της ελευθερίας από τη σκοπιά όλων, ενώ η θεμελιώδης αρχή του Δικαίου ορίζει ποια μορφή πρέπει να έχουν οι πράξεις μας εάν πρόκειται να είναι συμβατές με την καθολική αξία της ελευθερίας, ανεξάρτητα από τις

προθέσεις και τα εσωτερικά κίνητρά μας. Η Ηθική και το Δίκαιο ανήκουν αμφότερα στο κανονιστικό πεδίο μέσω της ισχύος επιτακτικών, επιτρεπτικών ή απαγορευτικών κανόνων, αλλά δεσμεύουν με διαφορετικούς τρόπους.

Η Ηθική επικεντρώνεται στην προστασία της ακεραιότητας εξατομικευμένων και διακριτών προσώπων στις μεταξύ τους σχέσεις. Το Δίκαιο, ως μέσο οργάνωσης της πολιτικής εξουσίας, προστατεύει τα μέλη της κοινωνίας, θέτοντας ρυθμίσεις στις εξωτερικές συμπεριφορές τους, η μη τήρηση των οποίων συνοδεύεται από κυρώσεις. Κατά τούτο, το Δίκαιο μπορεί να θεωρηθεί ως λειτουργικό συμπλήρωμα της Ηθικής. Η ηθικότητα και η νομιμότητα, όμως, αποτελούν δύο διακριτά πεδία κανονιστικής αποτίμησης των ανθρώπινων πράξεων, τα οποία συνδέονται μεταξύ τους.

Ενώ, από το ένα μέρος, το Δίκαιο και οι θεσμοί δεν μπορούν να αποκοπούν από ηθικές εκτιμήσεις, από το άλλο μέρος, δεν είναι, και δεν πρέπει να είναι, όλοι οι ηθικώς θεμιτοί και αποδεκτοί κανόνες που διέπουν τον συλλογικό βίο νομικοί κανόνες. Δεν βασιζόμαστε μόνο στον θετικό νόμο, προκειμένου να αποτραπούν αθέμιτες ή απαράδεκτες συμπεριφορές, όπως η κλοπή, το ψέμα, η αθέτηση υποσχέσεων, αλλά και σε ηθικά πρότυπα και κανόνες που εμπεδώνουν τις ορθές σχέσεις μεταξύ μας.

Στις νεωτερικές κοινωνίες, η Ηθική και το Δίκαιο συνδέονται μεταξύ τους μέσω της θεμελιώδους αρχής της αυτονομίας (ως *ηθικής αυτονομίας* των προσώπων και *νομικο-πολιτικής αυτονομίας* ελεύθερων και ίσων πολιτών ως προσώπων). Η αυτονομία του προσώπου, στην οποία εδράζεται η αξιοπρέπειά του (η άνευ όρων αξία του), αποτελεί την κανονιστική βάση όλων των υποχρεώσεων και όλων των δικαιωμάτων. Αυτό σημαίνει ότι ποτέ ένα ανθρώπινο ον, ως έλλογο ον με το ηθικό γνώρισμα της ελευθερίας, που πράττει και λογοδοτεί για τις πράξεις του, δεν πρέπει να τυγχάνει μεταχείρισης που να το μετατρέπει από υποκείμενο της κοινωνικής συμβίωσης σε απλό αντικείμενό της, σε «απλό μέσο». Τα ανθρώπινα όντα, ως πρόσωπα, αξίζουν και απαιτούν τον σεβασμό των όρων της δυνατότητας πράξης τους (δηλαδή της ελευθερίας τους) στη μοναδικότητά τους. Η φυσική και ψυχολογική ακεραιότητά τους πρέπει να προστατεύεται με κάθε τρόπο.

Η (πρωτ)αρχή του σεβασμού της αυτονομίας των προσώπων διακλαδώνεται σε θεμελιώδεις ηθικο-πρακτικές αρχές, πρωτίστως εκείνες της μη βλάβης, της δικαιοσύνης, της ωφέλειας, της αλληλεγγύης, οι οποίες αποτελούν το κανονιστικό θεμέλιο της ισχύος καθηκόντων και δικαιωμάτων (όπως εκείνα της ελευθερίας της έκφρασης, της ιδιωτικότητας, της θρησκευτικής ελευθερίας κλπ.), τα οποία προσδιορίζουν ποιες πράξεις μπορούν να υποστηριχθούν ηθικά και ποιες όχι. Πράξεις που βασίζονται σε υποκειμενικές αρχές εξαπάτησης, εξαναγκασμού, χειραγώγησης, απαξίωσης, ταπείνωσης είναι ηθικά απαράδεκτες. Οι γενικές δεσμεύσεις που τις απαγορεύουν μπορούν να διατυπωθούν ως απαγόρευση ορισμένων

ειδών μη συναινετικών πράξεων. Στη βάση αυτών των δεσμεύσεων, αποκτά κανονιστική ισχύ η αρχή της συναίνεσης κατόπιν ενημέρωσης ως ασπίδα προστασίας κατά της εξαπάτησης, του εξαναγκασμού, της χειραγώγησης ή της αθέμιτης πίεσης σε ένα ανθρώπινο ον. Στη βάση αυτών των αρχών, η Ηθική συνομιλεί οργανικά με το Δίκαιο και την Πολιτική.

2.6.2. Η Ηθική για την Τεχνητή Νοημοσύνη

Τα συστήματα της ΤΝ, την ίδια στιγμή που διανοίγουν έναν ορίζοντα ευκαιριών για τη βελτίωση της ανθρώπινης ζωής, μας καθιστούν αντιμέτωπους με προκλήσεις που μπορεί να επιφέρουν αποφασιστικές αρνητικές συνέπειες, συμπεριλαμβανομένων εκείνων που είναι δύσκολο να προβλέψουμε, να ταυτοποιήσουμε ή να μετρήσουμε, όπως επιπτώσεις στη δημοκρατία και τη δημοκρατική ζωή των πολιτών, στο κράτος δικαίου και στην προστασία θεμελιωδών δικαιωμάτων, στις αρχές της κοινωνικής δικαιοσύνης, ή ακόμα και στον τρόπο λειτουργίας της δικής μας νόησης ή στις μεταξύ μας σχέσεις, όταν η αυτοματοποίηση των ενισχυμένων με ΤΝ μηχανών τείνει να αποτελέσει πρότυπο για τις διαπροσωπικές μας σχέσεις.

Αφετηρία των ηθικών εκτιμήσεών μας για την ΤΝ αποτελεί η αρχή του σεβασμού της αυτονομίας των προσώπων (αξιοπρέπειας). Η ανάπτυξη της ΤΝ απαιτείται να γίνεται με σεβασμό στην ανθρώπινη αξία, στις ελευθερίες και στα θεμελιώδη δικαιώματα των προσώπων. Κεντρικής ηθικής σημασίας είναι η προστασία της αξίας της ανθρώπινης επιλογής και του ελέγχου. Απαιτείται να γίνεται σεβαστή, να προστατεύεται και να ενδυναμώνεται η ικανότητα των εμπλεκόμενων προσώπων για ελεύθερη, ανεμπόδιστη και ενημερωμένη απόφαση σχετικά με τους όρους αλληλεπίδρασής τους με τα εργαλεία ΤΝ. Οι χρήστες πρέπει να μπορούν να διατηρούν τη δύναμη να αποφασίζουν ποιες επιλογές θα κάνουν σε σχέση με τα εργαλεία της και να μην καθίστανται θύματα άμεσου ή έμμεσου εξαναγκασμού, εξαπάτησης, αθέμιτης πίεσης ή άλλης μορφής χειραγώγησης, που προσβάλλει την προσωπικότητά τους και τους εργαλειοποιεί.

Περαιτέρω, ο σεβασμός της αρχής της ιδιωτικότητας και του ιδιωτικού βίου, που συνδέεται με την αρχή της αυτονομίας, επιβάλλει να διασφαλίζεται η ανωνυμία του κάθε συγκεκριμένου προσώπου και να απαγορεύεται η μεταβίβαση σε τρίτους των δεδομένων του και των πληροφοριών που συνδέονται με αυτά, εφόσον αυτή διενεργείται με τρόπο που επιτρέπει την αποκάλυψη προσωπικών στοιχείων του, τα οποία το ίδιο το άτομο δικαιούται να κρατά ιδιωτικά. Στην εποχή της μαζικής συλλογής δεδομένων, το δικαίωμα στην προστασία της προσωπικής πληροφορίας δέχεται κρίσιμες προκλήσεις. Τα «νοήμονα» συστήματα δεν πρέπει να χρησιμοποιούνται με τρόπους που επεμβαίνουν ή εισβάλλουν αθέμιτα ή εμποδίζουν την άσκηση του δικαιώματος στην ιδιωτική ζωή, το οποίο περιλαμβάνει την αξίωση να είναι κανείς ελεύθερος από τεχνολογίες που επηρεάζουν την προσωπική του ανάπτυξη ή

τη διαμόρφωση της γνώμης του ή το δικαίωμά του να αναπτύσσει προσωπικές σχέσεις με άλλους ανθρώπους, ελεύθερος από την αυτοματοποιημένη ανάλυση των προσωπικών δεδομένων του.

Εξέχουσα σημασία στον ηθικό τρόπο σκέψης κατέχει η αρχή του μη βλάπτειν (*primum non nocere*), ιδίως υπό τη μορφή της προστασίας της ζωής και της υγείας των προσώπων. Εξειδικεύεται στην επιταγή τα εργαλεία της ΤΝ να μη χρησιμοποιούνται με τέτοιους τρόπους ώστε να προκαλείται φυσική ή ψυχολογική βλάβη ή ζημία σε πρόσωπα, περιλαμβανομένων δυνητικών αρνητικών συνεπειών που μπορεί να προκληθούν από εσφαλμένη χρήση ή κατάχρηση. Η πρόληψη ή αποτροπή βλάβης περιλαμβάνει την αποτροπή τόσο της ενδεχόμενης, απρόβλεπτης όσο και της κακόβουλης βλάβης, είτε δηλαδή αυτή προέρχεται από απρόβλεπτη λειτουργία των αλγορίθμων είτε από ανθρώπινη πρόθεση.

Η αρχή της ωφέλειας (αγαθοπραξίας) αξιώνει να λαμβάνονται κατάλληλα μέτρα όχι μόνο για την αποφυγή πρόκλησης βλάβης, αλλά και για την εξασφάλιση της ευζωίας των εμπλεκόμενων προσώπων (του βιοτικού αγαθού τους), όσο αυτό είναι δυνατόν. Επιτάσσει να σχεδιάζεται η ΤΝ με τρόπους που προάγουν το ανθρώπινο ευ ζην, αλλά και τη βιωσιμότητα του πλανήτη.

Στον ηθικό προβληματισμό περιλαμβάνεται η θεμελιώδης αρχή της δικαιοσύνης, η οποία είναι επίσης συνδεδεμένη με την αξία των προσώπων. Σχετίζεται με το ποιος δικαιούται να λαμβάνει τα οφέλη και ποιος υποχρεούται να αναλαμβάνει το κόστος από την ανάπτυξη και την αξιοποίηση της τεχνολογίας της ΤΝ. Αδικία διαμείβεται όταν ένα πρόσωπο, χωρίς επαρκή λόγο, αποκλείεται από κάποιο όφελος στο οποίο έχει δικαίωμα, ή όταν υφίσταται αδικαιολόγητα κόστος που δεν του αρμόζει. Οι αποφάσεις πρέπει να συνάδουν με την αρχή της δίκαιης διανομής και να παραμένουν ανεπηρέαστες από παραμέτρους όπως δεδομένα φυλετικών, κοινωνικών, πολιτισμικών ιδιαιτεροτήτων μεταξύ των ενδιαφερόμενων προσώπων.

Είναι ηθικά απαράδεκτη η χρήση ενός συστήματος ΤΝ που εκμεταλλεύεται την ευαλωτότητα μιας συγκεκριμένης ομάδας, οφειλόμενη στην ηλικιακή κατηγορία των μελών, στο φύλο ή σε κάποια φυσική ή νοητική αναπηρία, με στόχο τη διαστρέβλωση της συμπεριφοράς ατόμων που ανήκουν στη συγκεκριμένη ομάδα, με τρόπους που προκαλούν, ή είναι πιθανό να προκαλέσουν, φυσική ή ψυχολογική βλάβη σε αυτά. Αποτελεί, επίσης, αξίωση δικαιοσύνης να αντιμετωπίζεται αποτελεσματικά το ψηφιακό χάσμα μεταξύ ατόμων διαφορετικών κοινωνικών ομάδων ως προς την πρόσβαση και τις δεξιότητες χρήσης των νέων αυτών τεχνολογιών.

Το παραπάνω πλαίσιο θεμελιωδών αρχών συνοδεύεται από μια σειρά δευτερογενών, λειτουργικών, αρχών (π.χ. αρχές της πρόληψης, της προφύλαξης, της διαλειτουργικότητας κλπ.), οι οποίες αφορούν στις διαδικασίες

που πρέπει να υιοθετούνται προκειμένου να αντιμετωπίζονται με τον δέοντα τρόπο τα ηθικά προβλήματα. Σημαντικότερη είναι η δευτερογενής αρχή της επεξηγησιμότητας, η οποία αποτελεί αξίωση να είναι κατανοητές, ερμηνεύσιμες και αποδόσιμες στον φορέα τους οι διαδικασίες που λειτουργούν κατά τη λήψη απόφασης με ΤΝ. Ως δευτερογενής αρχή, εξακριβώνει και καθιστά εμφανές κατά πόσον οι θεμελιώδεις ηθικές αρχές γίνονται σεβαστές, ή διακυβεύονται, κατά την ανάπτυξη και χρήση της ΤΝ. Ενσωματώνει το επιστημολογικό αίτημα της κατανοησιμότητας (απαντά στο ερώτημα «πώς ένα σύστημα λειτουργεί;») και το ηθικό αίτημα της λογοδοσίας (απαντά στο ερώτημα «ποιος είναι υπεύθυνος για τον τρόπο που λειτουργεί;»).

Προσοχή όμως: Δεν υπάρχουν «ηθικοί αλγόριθμοι», δεν φέρουν ηθικό στάτους, δεν λογοδοτούν, δεν έχουν καθήκοντα, ούτε δικαιώματα. Θα πρέπει να αποφεύγεται η αναγωγή της κανονιστικής ηθικής προβληματοθεσίας σε απλή περιγραφική, λειτουργική διαδικασία. Ειδικότερα, πρέπει να αποφεύγονται, αφενός, η αναγωγή της ηθικής εκτίμησης σε καθαρά ποσοτικές θεωρήσεις ευκαιριών και ρίσκων, και, αφετέρου, τα στενά εκείνα εγχειρήματα που αντιμετωπίζουν τα ηθικά προβλήματα αποκλειστικά μέσω αναγωγής σε διαδικαστικές στρατηγικές.

Κατά την ηθική ανάλυση, αφετηρία αποτελούν οι θεμελιώδεις αρχές, για τις οποίες μπορεί να δοθεί εδραία κανονιστική δικαιολόγηση. Με βάση αυτές τις αρχές, μπορούμε να επιχειρηματολογήσουμε υπέρ συγκεκριμένων κανόνων ή κατευθυντήριων οδηγιών, που απευθύνονται προς όλους τους εμπλεκομένους στην αλυσίδα της ΤΝ, τους σχεδιαστές, τους προγραμματιστές, τους μηχανικούς, τους κατασκευαστές, τους χρήστες των συστημάτων ΤΝ, αλλά και προς επαγγελματικά και ρυθμιστικά σώματα και επιτροπές, που αναλαμβάνουν το εγχείρημα της ηθικής αποτίμησης, της ρύθμισης και της λήψης απόφασης σε συγκεκριμένες περιπτώσεις.

2.6.3. Η Ηθική, η Πολιτική και η ρύθμιση στο οικοσύστημα της Τεχνητής Νοημοσύνης

Η Ηθική και το Δίκαιο αποτελούν σημαντικές παραμέτρους της διακυβέρνησης της ΤΝ. Η τελευταία συγκροτεί εκείνο το κανονιστικό πλαίσιο που αφορά στον προσδιορισμό και την εφαρμογή πολιτικών, κανόνων, διαδικασιών για την ορθή ανάπτυξη και διαχείριση της ΤΝ. Εφάπτεται, κατά τούτο, με ηθικές και νομικο-πολιτικές επεξεργασίες, ενσωματώνει κατευθυντήριες οδηγίες, συστάσεις και πολιτικές που απορροφούν ηθικές απαιτήσεις και νομικές προβλέψεις. Η διακυβέρνηση της ΤΝ αφορά όχι μόνο στο πώς η ανάπτυξη και αξιοποίησή της θα είναι διαχειρίσιμη και αποδοτική, αλλά πρωτίστως στο πώς θα αποβεί κοινωνικά δίκαιη και ενισχυτική της ανθρωπίνης αυτονομίας και ευημερίας.

Σε αυτό το πλαίσιο, η ηθική αξιολόγηση συμπληρώνει τη νομική συμμόρφωση. Η συμμόρφωση είναι αναγκαία, αλλά όχι επαρκής για να προσανατολισθεί η πράξη μας. Η ρύθμιση χαρακτηρίζει ποιες επιλογές είναι νόμιμες και ποιες όχι, αλλά δεν μπορεί να πει ποιες είναι καλές και ποιες καλύτερες από εκείνες που είναι νόμιμες, ώστε να επιτευχθούν κοινωνικοί στόχοι. Το τελευταίο αποτελεί έργο ηθικής επεξεργασίας. Απαντά στο ερώτημα ποιες αρχές και αξίες επιδιώκουμε να ενσωματώσουμε και να υλοποιήσουμε μέσω της τεχνολογικής καινοτομίας, καθώς και ποιες ρυθμιστικές πολιτικές απαιτούνται για την ορθή διακυβέρνηση, από την πλευρά της διαχείρισής της.

Ειδικότερα, η Ηθική της ΤΝ, ως μέρος της Ηθικής, ανιχνεύει και αποτιμά τις ηθικές προκλήσεις που σχετίζονται με τη χρήση δεδομένων και πληροφορίας, με την ανάπτυξη των αλγορίθμων και των τεχνικών προτύπων και με τις αντίστοιχες πρακτικές, συμπεριλαμβανομένων της υπεύθυνης καινοτομίας, του προγραμματισμού, της διαμόρφωσης κωδίκων και κανόνων. Συγκροτεί τον απαραίτητο εννοιολογικό και κανονιστικό χάρτη που προσανατολίζει την καινοτομία, αποτυπώνει κινδύνους και προκλήσεις, συμβάλλει στην επεξεργασία και υποστήριξη ηθικά θεμιτών λύσεων, υπό το φως κριτικά επεξεργασμένων επιδιωκόμενων στόχων.

Εξ αρχής, ο σχεδιασμός των αλγορίθμων και η τροφοδοσία τους με δεδομένα ενσωματώνουν αξιολογικές επιλογές και έχουν, ως εκ τούτου, ηθικό και κοινωνικό σθένος. Μπορεί να αναπαράγουν δικές μας αδυναμίες και ελαττώματα, περιορισμούς και μεροληψίες. Χρειάζεται, επομένως, να σχεδιαστούν ορθά και με τα κατάλληλα κανονιστικά πρότυπα. Αποτελεί, δε, κοινή διαπίστωση πως, κατά την ανάπτυξη της ΤΝ, η επίτευξη κοινωνικά ωφέλιμων αποτελεσμάτων συνδέεται με την άρση της έντασης ανάμεσα στην ενσωμάτωση ωφελειών, αφενός, και στην αποφυγή ή στην άμβλυνση δυνητικών βλαβών ή κινδύνων βλάβης, συμπεριλαμβανομένης της αντιμετώπισης απειλών κατά της ασφάλειας, αφετέρου.

Η ανάπτυξη και οι εφαρμογές των συστημάτων ΤΝ μπορεί να εντείνουν ανισότητες, να αποκλείσουν κοινωνικές ομάδες, κυρίως ευάλωτες, να επηρεάσουν τη δημόσια συζήτηση με παραπληροφόρηση, κατασκευασμένες ειδήσεις και άλλους μηχανισμούς χειραγώγησης, να εντείνουν προκαταλήψεις, να αλλοιώσουν τη δημόσια επικοινωνία. Χρειάζεται να λαμβάνονται μέτρα ώστε να σχεδιάζεται η ΤΝ με τρόπους που θέτουν φραγμούς στην αδικία, στην εξαπάτηση, στην άμεση ή έμμεση χειραγώγηση και στην προσβολή ατομικών δικαιωμάτων, στην επιβολή δυσανάλογων επιβαρύνσεων σε ορισμένους. Απαιτούνται δικλίδες ασφαλείας, κριτήρια αξιολόγησης και διαρκής αξιόπιστη εποπτεία σε όλο τον κύκλο λειτουργίας της.

Η ηθική προσέγγιση θέτει, ούτως ειπείν, τις «κόκκινες γραμμές», προλαμβάνοντας άστοχες ή επιζήμιες επιλογές, που υπονομεύουν την εμπιστοσύνη των πολιτών. Μπορεί να λειτουργήσει ως ένα σύστημα πρώιμης ασφαλούς

προειδοποίησης για το ποιες από τις τεχνολογίες ΤΝ, που είναι δυνατόν να αναπτυχθούν, πρέπει να αποκλεισθούν, ποιες μπορούν να αναπτυχθούν, υπό ποιους περιορισμούς και γιατί. Οι κανονιστικές αναλύσεις της αποτελούν την πρώτη γραμμή άμυνας, εκεί όπου μπορεί να απειληθούν θεμελιώδεις αξίες, δικαιώματα ή έννομα αγαθά από τη μορφή που λαμβάνει η αλματώδης και ανταγωνιστική ανάπτυξη της ΤΝ.

Παρέχεται έτσι μια ασφαλής βάση για την αξιόπιστη ανάπτυξη της καινοτομίας, μειώνοντας σημαντικά το «κόστος ευκαιρίας» επιλογών που δεν θα έπρεπε να γίνουν ή επιλογών που δεν έγιναν, ενώ θα μπορούσαν, αλλά αποκλείστηκαν εξαιτίας ανησυχιών περί δυνητικών λαθών. Είτε πρόκειται για άτομα είτε για φορείς, μπορεί να ενισχύσει τις επιλογές ώστε να αξιοποιηθούν κατάλληλες ευκαιρίες για καινοτομία, αποφεύγοντας τα «ηθικά ρίσκα» και το «ηθικό κόστος», ειδικά μάλιστα σε περιπτώσεις όπου οι νομικές ρυθμίσεις δεν παρέχουν μονοσήμαντη ή σαφή καθοδήγηση, ή εκεί όπου υπάρχει κανονιστική αβεβαιότητα ή ασάφεια, ή σε περιπτώσεις που μπορεί να συγκρούονται αξίες και συμφέροντα που πρέπει να εναρμονισθούν.

Η ηθική προσέγγιση συμβάλλει στην εμπέδωση ενός αναπτυξιακού περιβάλλοντος υπεύθυνης καινοτομίας και επιχειρηματικής αξιοπιστίας, που είναι απαραίτητες για την κοινωνικά ωφέλιμη και βιώσιμη ανάπτυξη της τεχνολογίας. Υπό αυτό το πρίσμα, το πλαίσιο αρχών της Ηθικής διαδραματίζει έναν επιτρεπτικό και έναν περιοριστικό ρόλο συγχρόνως. Η χαρτογράφηση μιας τεχνολογικής ανάπτυξης, που θα είναι κοινωνικά ωφέλιμη, δίκαιη και περιβαλλοντικά βιώσιμη, εξαρτάται σε μεγάλο βαθμό από την εμπέδωση ενός έθνους υπευθυνότητας και αξιοπιστίας, που εμποτίζεται από ορθές πρακτικές και ενδυναμώνει την εμπιστοσύνη των πολιτών.

Η ηθική ανάλυση αποτελεί κανονιστική πηγή προσανατολισμού για τη διαμόρφωση πολιτικών που, αφενός, συμβάλλουν στην αξιοποίηση των ευκαιριών που παρέχει η ΤΝ και, αφετέρου, παρέχουν εγγυήσεις ώστε η ανάπτυξή της να ενισχύει την κοινωνική ευημερία. Συμβάλλει στη νομικο-πολιτική συζήτηση για τη διαμόρφωση μιας ορθά σχεδιασμένης νομοθεσίας.

Η συμβολή της ηθικής επιχειρηματολογίας αποβαίνει ιδιαίτερα γόνιμη σε περιοχές ρύθμισης με κανονιστική αβεβαιότητα, όπου απαιτείται ερμηνεία, ειδικά όταν πρόκειται για τα ταχύτατα μεταβαλλόμενα και ανταγωνιστικά περιβάλλοντα της ΤΝ. Υπάρχει, λοιπόν, ένα αποφασιστικό ηθικό στοιχείο, που μαζί με άλλους παράγοντες μπορεί να κατευθύνει την επεξεργασία, την ανάλυση, την ερμηνεία και την ανάπτυξη των ρυθμίσεων του Δικαίου. Οι αρχές της Ηθικής δεν είναι «εξωτερικές» προς το Δίκαιο, αλλά ενσωματώνονται σε αυτό ως συστατικά στοιχεία του, ενισχύοντας την εφαρμογή (συμμόρφωση), αλλά και την ενδεχόμενη αλλαγή του, καθιστώντας το ικανό για σημαντικές βελτιώσεις, ειδικά μάλιστα καθώς νομικές διατάξεις εφαρμόζονται σε νέες, όλο και περισσότερο πολύπλοκες, τεχνολογικές περιστάσεις.

Πρέπει να υπογραμμισθεί, από το άλλο μέρος, ότι η Ηθική δεν διαθέτει μηχανισμούς επιβολής. Μείζον διακύβευμα αποτελεί η ενεργοποίηση των ηθικών επιταγών στην πράξη από τον σχεδιασμό των τεχνικών προτύπων και σε όλη την αλυσίδα λειτουργίας των συστημάτων ΤΝ. Σημαντικό ρόλο διαδραματίζει η εμπέδωση κωδίκων, αρχών, ηθικών πλαισίων, κατευθυντήριων οδηγιών, καλών πρακτικών, σε όλα τα στάδια της ανάπτυξης και χρήσης της ΤΝ. Μια δυναμική προσέγγιση για την αξιόπιστη διακυβέρνησή της απαιτεί να περιλαμβάνονται ηθικές εγγυήσεις εκ του σχεδιασμού και εξ αρχής. Κανόνες χρειάζεται, πράγματι, να κωδικοποιηθούν στα τεχνικά πρότυπα και τα επιχειρηματικά μοντέλα, οι οποίοι σχετίζονται με την ασφάλεια των συστημάτων της σε όλο τον κύκλο λειτουργίας της, την ποιότητα των δεδομένων, την εξηγησιμότητα, τη διαφάνεια και την κατανοησιμότητα των μοντέλων.

Η ενσωμάτωση των κανόνων απαιτεί την αφομοίωση και την αξιοποίηση ενός συμπεφωνημένου πλαισίου αρχών, με σεβασμό στην ανθρώπινη αυτονομία και την ευημερία. Η, δε, κατοχύρωση της δυνατότητας ανθρώπινου ελέγχου και επαλήθευσης απαιτείται να ενσωματώνεται στην ίδια τη διαδικασία της αυτοματοποίησης και της αλγοριθμικής λήψης αποφάσεων και να διασφαλίζεται σε όλα τα στάδια σχεδιασμού, ανάπτυξης και λειτουργίας των τεχνητών συστημάτων.

Απαιτείται μια ρυθμιστική προστατευτική νομοθεσία, η οποία να συνοδεύει, και να ενισχύει, την ηθική ανάπτυξη της ΤΝ. Ο προσανατολισμός των στόχων επιτυγχάνεται μέσω κατάλληλων ερευνητικών στρατηγικών που χρειάζεται να σχεδιασθούν, του διαφανούς προσδιορισμού χρηματοδοτικών προτεραιοτήτων, της διατύπωσης και εφαρμογής σαφών επιχειρηματικών κανόνων, καθώς και της συστηματικής διαμόρφωσης συναφών δημόσιων πολιτικών, με ενδυνάμωση του δημοκρατικού ελέγχου. Απαραίτητη κρίνεται η πρόβλεψη σαφών και αποτελεσματικών μορφών και θεσμών επίβλεψης και εποπτείας, με δυνατότητα, μεταξύ άλλων, παρακολούθησης και αναδρομικής ανάλυσης σφαλμάτων ή επιζήμιων συνεπειών.

Χρειάζεται, συγχρόνως, να αναληφθούν συστηματικές πρωτοβουλίες για μια παγκόσμια συνεργασία για τη διακυβέρνηση της ΤΝ, καθώς οι ισχυρές τεχνολογικές εταιρείες που την αναπτύσσουν λειτουργούν σε παγκόσμιο επίπεδο, καθιστώντας έτσι λιγότερο αποτελεσματικές τις επιμέρους εθνικές (και ευρωπαϊκές) ρυθμιστικές πολιτικές. Η ηθική προβληματική μπορεί να συμβάλει σε μια παγκόσμια συνεννόηση για την ανάπτυξη και την αξιοποίηση της πλέον δυναμικής και διασπαστικής τεχνολογίας της εποχής μας.

Αναμφιβόλως, η ανάπτυξη μιας «ηθικά υπεύθυνης» ΤΝ στην πράξη χρειάζεται να συν-εξελισσεται με το πεδίο έρευνας της ίδιας της ΤΝ. Απαιτείται περισσότερη έρευνα για τις ηθικές, κοινωνικές, νομικές διαστάσεις των «νοημόνων συστημάτων», τα οποία αναπτύσσονται με μια τέτοια δυναμική που τα καθιστά ικανά να λειτουργούν όλο και περισσότερο πέραν του ανθρώπινου ελέγχου.

Ο τεχνολογικός σχεδιασμός της ΤΝ δεν είναι αποκλειστικά τεχνικό πρόβλημα ώστε να αντιμετωπίζεται απλώς τεχνικά, με την ενσωμάτωση ηθικών προτύπων εξ αρχής και εκ του σχεδιασμού. Απαιτείται να συνδυάζεται με την κανονιστική προβληματική μιας ευρύτερης, περιεκτικής ηθικο-πρακτικής θεώρησης. Δεν υφίσταται ένας απλός ή μονοσήμαντος τρόπος εφαρμογής ή ενσωμάτωσης των ηθικών αρχών στην πράξη, ούτε αλγοριθμικά ούτε μηχανιστικά. Καθώς ούτε οι αρχές ούτε οι αξίες μπορούν οι ίδιες να παράσχουν οδηγίες για την εφαρμογή τους, η ενεργοποίησή τους στην πράξη εμπλέκει πάντοτε, και συμπληρώνεται από, την ανθρώπινη κρίση. Μια αυτοματοποιημένη «ηθική», ενσωματωμένη στα συστήματα ΤΝ, δεν είναι επαρκής. Ο ηθικός συλλογισμός, η ηθική μας σκέψη είναι πάντοτε ένα εγχείρημα και ένα διακύβευμα. Προϋποθέτει την ελεύθερη και συνειδητή πράξη, που εγχαράσσεται σε διαπροσωπικές σχέσεις, στον καμβά των οποίων αναπτύσσονται καθήκοντα και ευθύνες.

Η Ηθική δεν εξαντλείται σε μια αυτοματοποιημένη ή στατιστική διαδικασία. Αυτή η διαπίστωση έχει κρίσιμες επιπτώσεις όσον αφορά το «πρόβλημα του ελέγχου» και της ηθικής διακυβέρνησης της ΤΝ. Απαιτούνται διαρκής εκτίμηση και επανεκτίμηση δυνατοτήτων και συγκεκριμένων περιπτώσεων, υπό το πρίσμα της διαθέσιμης γνώσης, σε συνδυασμό με ισχυρά δικαιολογημένα ηθικά επιχειρήματα. Απαιτείται η συγκεκριμένη απόφαση να συνοδεύεται από αιτιολογημένη ανθρώπινη κρίση σε κάθε συγκεκριμένη περίπτωση.

Η Ηθική στο οικοσύστημα της ΤΝ συνιστά ένα πλέγμα δραστηριοτήτων, που περιλαμβάνουν την ορθολογικά επεξεργασμένη κριτική ανίχνευση ηθικών προκλήσεων, την κανονιστική εμπλοκή με αυτά, τη δομημένη επιχειρηματολογία, τον σχεδιασμό και την εφαρμογή ορθών και δίκαιων λύσεων, και την άσκηση συγκεκριμένης πρακτικής κρίσης. Συμβάλλει στην πολιτική συζήτηση για τη ρύθμιση της ΤΝ, όχι μόνο με κριτήριο την αποδοτικότητα ή την αποτελεσματικότητα των εργαλείων της, αλλά με πρωτανεύοντα κριτήρια την ενδυνάμωση της ανθρώπινης αυτονομίας και την ενίσχυση της κοινωνικής ευημερίας.

2.6.4. Quo vadis⁷³ Τεχνητή Νοημοσύνη;

Προβαίνουμε σε ηθικές σκέψεις για την ΤΝ, προκειμένου να διαμορφώσουμε μια *προοπτική*, ένα πλαίσιο αναφορικά με τις αυξανόμενες τεχνολογικές δυνατότητές μας, το οποίο δείχνει πόσο στενά οι ηθικοί προβληματισμοί, και σε αυτήν την περιοχή, συνδέονται με καταστατικά διακυβεύματα γύρω από την ανθρώπινη ελευθερία, την αυτονομία, τον αυτοπροσδιορισμό, την ισότητα, τη δικαιοσύνη, γύρω από θεμελιώδεις κανονιστικούς όρους συνοχής

⁷³ Λατινική φράση που σημαίνει «Πού βαδίζεις?».

και αναπαραγωγής της κοινωνικής ζωής, που βρίσκονται στον πυρήνα της ηθικής αυτοκατανόησής μας και της δημοκρατικής ζωής.

Όταν τέτοιοι βασικοί όροι διακυβεύονται ή απειλούνται από τις ίδιες τις πρακτικές μας και τη μορφή που λαμβάνει η κοινωνική ανάπτυξη στην τεχνολογική εποχή μας, τότε αυτοί οι όροι καθίστανται επιτακτικά ηθικο-κοινωνικά αιτήματα προς υπεράσπιση. Συνδέονται με την απαίτηση να αποτιμούμε την κοινωνική πράξη (όπως διευρύνεται και ανασυντάσσεται μέσω των ψηφιακών τεχνολογιών, συμπεριλαμβανομένης της ΤΝ) όχι μόνον ή αποκλειστικά με κριτήριο τις προτιμήσεις ή τα συμφέροντα που ικανοποιούνται ή θίγονται, αλλά και από τη σκοπιά της «προόδου» προς τις αξίες της ελευθερίας, της ισότητας, της δικαιοσύνης και της αλληλεγγύης, που αποτελούν το κανονιστικό θεμέλιο της κοινωνικής συνοχής. Μια από τις ενδιαφέρουσες διαστάσεις της συζήτησης για την ΤΝ είναι ότι οδηγούμαστε να επανεξετάσουμε κρίσιμα ηθικο-κοινωνικά και δικαιοπολιτικά διακυβεύματα της πράξης στον σύγχρονο ασταθή και αβέβαιο κόσμο, τα οποία ενίοτε τείνουμε να παραβλέπουμε ή να θεωρούμε πως έχουν ήδη διευθετηθεί...

Βιβλιογραφία

Floridi, L. (2023). *The Ethics of Artificial Intelligence, Principles, Challenges, and Opportunities*, New York, NY: Oxford University Press.

High-Level Expert Group on Artificial Intelligence (2019). *Ethics guidelines for trustworthy AI*. European Commission. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

Mitrou, L. (2018). Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof'? Available at <http://dx.doi.org/10.2139/ssrn.3386914>

O'Neill, O. (2022). *A Philosopher Looks at Digital Communication*, Cambridge: Cambridge University Press.

Τσινόρεμα, Σ. (2024). Τεχνητή νοημοσύνη με ανθρώπινο πρόσωπο. Προς μια τεχνοθική της ευθύνης. *διά-Λογος*, 14, 163-192.

[ΠΩΣ] ΜΠΟΡΕΙ Η ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΝΑ ΕΙΝΑΙ
ΗΘΙΚΗ ΚΑΙ ΣΥΝΝΟΜΗ;
Ιούλιος 2026

Το ενωσιακό και το εθνικό κανονιστικό πλαίσιο της Τεχνητής Νοημοσύνης



**Μιχάλης Κρητικός, Λίλιαν Μήτρου,
Βαγγέλης Παπακωνσταντίνου**

Συγγραφική συνεισφορά ανά υποκεφάλαιο:

3.1., 3.2., 3.4., 3.5. – Λίλιαν Μήτρου

3.3. – Βαγγέλης Παπακωνσταντίνου

3.6. – Μιχάλης Κρητικός

Μπορεί να οριστεί και να περιχαρακωθεί η «ηθική χρήση» της Τεχνητής Νοημοσύνης προς το (κοινό;) καλό; Τίθενται «ηθικά ζητήματα» ή μήπως στον όρο αυτόν ενυπάρχει η αντίληψη των ορίων και των περιορισμών που δεν τους επιβάλλει μόνο η ηθική, αλλά η έννομη τάξη; Πρέπει και μπορεί ο νόμος να λειτουργήσει ως ανάχωμα σε ανήθικες, μη αποδεκτές χρήσεις;⁷⁴

Η προσπάθεια να παρακολουθήσουμε, ασθματικά, τις ραγδαίες, αλλά και μη προβλέψιμες τεχνολογικές εξελίξεις στο πεδίο της ΤΝ αναδεικνύει ως επιτακτική την ανάγκη να διασφαλίσουμε την υπεύθυνη χρήση της. Σε αυτό το πλαίσιο και με αφετηρία την ανθρωποκεντρική προσέγγιση και στόχο την επίτευξη μιας –πολυδιάστατα– αξιόπιστης ΤΝ, η Ευρωπαϊκή Ένωση επιδίωξε τη σύζευξη της τεχνολογικής και οικονομικής ανάπτυξης και της προστασίας θεμελιωδών δικαιωμάτων και αξιών μέσα από την καθιέρωση της απαίτησης για διαφάνη, δίκαιη και υποκείμενη στην ανθρώπινη εποπτεία χρήση της ΤΝ.

⁷⁴ Τη ρύθμιση της ΤΝ και την ύπαρξη περιορισμών υποστηρίζουν σε συντριπτικό ποσοστό οι συμμετέχοντες στην έρευνα της διαΝΕΟσις [διαΝΕΟσις (2026)]. *Τι πιστεύουν οι Έλληνες για την Τεχνητή Νοημοσύνη Α΄ Μέρος – Έκθεση αποτελεσμάτων*. https://www.dianeosis.org/wp-content/uploads/2026/05/AI_Dimoskopiki_Apotelesmata_190526.pdf. Στο ερώτημα «Κατά τη γνώμη σας, θα πρέπει να υπάρχουν περιορισμοί στη χρήση της Τεχνητής Νοημοσύνης (AI) ή θα πρέπει να αφεθεί ελεύθερη;», το 90,3% του δείγματος που γνωρίζει ή έχει ακουστά την ΤΝ υποστηρίζει την εφαρμογή περιορισμών, ενώ το 8,9% υποστηρίζει το να αφεθεί ελεύθερη. Τα ποσοστά αλλάζουν ελαφρά ανάμεσα στην ομάδα των ατόμων που χρησιμοποιούν την ΤΝ και αυτών που απλά τη γνωρίζουν (88,5% και 94,3%, αντίστοιχα, δηλώνουν υπέρ της ρύθμισης).

3.1. Η Πράξη για την Τεχνητή Νοημοσύνη

Με την υιοθέτηση της Πράξης για την Τεχνητή Νοημοσύνη ο ενωσιακός νομοθέτης επιδίωξε την ισορροπία μεταξύ, αφενός, του ελέγχου της ασφάλειας και της αξιοπιστίας των συστημάτων ΤΝ και, αφετέρου, της προστασίας των θεμελιωδών δικαιωμάτων των πολιτών της ΕΕ, αλλά και της δημοκρατίας. Αν και σημείο αφετηρίας ήταν η μετεξέλιξη της νομοθεσίας της ασφάλειας προϊόντων (product/health safety), ο Κανονισμός (ΕΕ) 2024/1689 του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου περιλαμβάνει ρυθμίσεις που αποσκοπούν στην προστασία δικαιωμάτων, αλλά και στην τήρηση αρχών, όπως η δίκαιη μεταχείριση, ή κανόνων, όπως το πλαίσιο προστασίας προσωπικών δεδομένων. Αυτός ο διφυής χαρακτήρας και στόχος του Κανονισμού, αν και κατανοητός, δυσχεραίνει την κατανόηση και την εφαρμογή των σύνθετων ρυθμίσεων.

Η αναζήτηση της βέλτιστης εξισορρόπησης μεταξύ της ανάπτυξης και της διασφάλισης των ηθικών και δικαιοκτών αρχών δεν είναι ευχερής. Είναι, όμως, το χαρακτηριστικό γνώρισμα και η προστιθέμενη αξία της ΕΕ, η οποία εκκινεί από την προαγωγή της υιοθέτησης της «ανθρωποκεντρικής και αξιόπιστης ΤΝ» (αιτιολογική σκέψη 1 της AI Act) και διακηρύσσει ότι «[ως] προαπαιτούμενο, η ΤΝ θα πρέπει να είναι μια τεχνολογία ανθρωποκεντρική. Θα πρέπει να χρησιμεύει σαν εργαλείο για τους ανθρώπους, με απώτατο στόχο την αύξηση της ανθρωπίνης ευημερίας» (αιτιολογική σκέψη 6 της AI Act).

Ο ορισμός των συστημάτων ΤΝ καλύπτει ένα ευρύ φάσμα τεχνολογιών, όπως η μηχανική μάθηση (machine learning), η βαθιά μάθηση (deep learning), η επεξεργασία φυσικής γλώσσας (natural language processing) και τα συστήματα ΤΝ γενικού σκοπού (general purpose AI systems). Οι ρυθμίσεις της AI Act, που συνιστά το πρώτο διεθνώς νομικά δεσμευτικό κείμενο για τη χρήση συστημάτων ΤΝ, αφορούν κατά περίπτωση όλους τους λεγόμενους φορείς εκμετάλλευσης ενός συστήματος ΤΝ (παρόχους, φορείς εφαρμογής, αντιπροσώπους, εισαγωγείς, διανομείς).

Η ρυθμιστική πρωτοβουλία της ΕΕ υπήρξε «λελογισμένη». Όπως και σε άλλα πρόσφατα κανονιστικά κείμενα, βασικός γνώμονας υπήρξε η «προσέγγιση με βάση τον κίνδυνο» που συνεπάγεται η χρήση συστημάτων ΤΝ, δηλαδή των πιθανών επιπτώσεων στην Υγεία, στην Ασφάλεια, στα θεμελιώδη δικαιώματα,

στη δημοκρατία, στο κράτος δικαίου και το Περιβάλλον. Τα επίπεδα κινδύνου δεν προσδιορίζονται, δηλαδή, με βάση την τεχνολογική μέθοδο, αλλά τις συνέπειες ως προς το δημόσιο συμφέρον και τα θεμελιώδη δικαιώματα. Οι υποχρεώσεις και οι όροι προβλέπονται και επιβάλλονται με κριτήριο την κατάταξη των συστημάτων σε επίπεδα κινδύνου. Αυτά κλιμακώνονται από τα συστήματα «ελάχιστου κινδύνου», για τα οποία δεν προβλέπονται υποχρεώσεις συμμόρφωσης, τα συστήματα «ελάχιστου κινδύνου διαφάνειας» (π.χ., χρήση διαλογικών ρομπότ), τα συστήματα «υψηλού κινδύνου» (π.χ., χρήση σε υποδομές ζωτικής σημασίας, όπως ο τομέας ηλεκτρικής ενέργειας ή στο πεδίο της απασχόλησης και διαχείρισης εργαζομένων) έως την απαγόρευση χρήσης συστημάτων «μη αποδεκτού κινδύνου» (π.χ., κοινωνική βαθμολόγηση για δημόσιους και ιδιωτικούς σκοπούς ή «αναγνώριση συναισθημάτων στον χώρο εργασίας»).

Η AI Act χαράσσει τις «κόκκινες γραμμές» του ενωσιακού νομοθέτη, απαγορεύοντας:

- α) την επιβλαβή χειραγώγηση και εξαπάτηση,
- β) την επιβλαβή εκμετάλλευση τρωτών σημείων ενός φυσικού προσώπου,
- γ) την κοινωνική βαθμολόγηση (social scoring),
- δ) την εκτίμηση και πρόβλεψη κινδύνου τέλεσης αξιόποινης πράξης από πρόσωπο,
- ε) τη μη στοχευμένη απόξεση⁷⁵ (scraping) για την ανάπτυξη βάσεων δεδομένων αναγνώρισης προσώπου,
- στ) την εξ αποστάσεως βιομετρική ταυτοποίηση σε πραγματικό χρόνο (Remote Biometric Identification – RBI) για σκοπούς εφαρμογής του νόμου,
- ζ) την αναγνώριση συναισθημάτων στον χώρο εργασίας και την εκπαίδευση, και
- η) τη βιομετρική κατηγοριοποίηση. Αξίζει να σημειώσουμε ότι το Ευρωπαϊκό Κοινοβούλιο (26/3/26), εγκρίνοντας κάποιες από τις προτάσεις του Digital Omnibus,⁷⁶ πρότεινε να συγκαταλεχθούν στις απαγορευμένες πρακτικές τα λεγόμενα «nudifier» ή «nudification systems», δηλαδή συστήματα ΤΝ που μπορούν να αλλοιώνουν, να χειραγωγούν ή να δημιουργούν ρεαλιστικές

⁷⁵ Ο όρος αναφέρεται στην αυτοματοποιημένη άντληση δεδομένων από το Διαδίκτυο, με σκοπό τη δημιουργία βάσεων δεδομένων που χρησιμοποιούνται από συστήματα αναγνώρισης προσώπου.

⁷⁶ Πρόκειται για τις προτάσεις που διατύπωσε η Ευρωπαϊκή Επιτροπή τη 19η Νοεμβρίου 2025, προτείνοντας μεταξύ των άλλων και κάποιες αλλαγές στην AI Act. Για περισσότερα βλ. παρακάτω.

εικόνες ή βίντεο, ώστε να εμφανίζουν ένα ταυτοποιήσιμο πραγματικό πρόσωπο σε σεξουαλικά ρητές ή ιδιωτικές/γυμνές απεικονίσεις χωρίς τη συναίνεσή του. Όπως ανακοινώθηκε στις 29 Ιουνίου 2026, το Συμβούλιο συμφώνησε να προστεθούν από τον Δεκέμβριο του 2026 στις απαγορευμένες πρακτικές τα συστήματα ΤΝ που δημιουργούν γυμνές εικόνες πραγματικών προσώπων ή επεξεργάζονται υφιστάμενες φωτογραφίες αφαιρώντας ψηφιακά τα ρούχα, ώστε να αποκαλύπτονται απόκρυφα σημεία του σώματος.⁷⁷

Ο κύριος όγκος των απαιτήσεων αφορά τους παρόχους συστημάτων «υψηλού κινδύνου», δηλαδή των συστημάτων που μπορεί να αναπτύσσουν επιβλαβείς επιπτώσεις σε σχέση με την Υγεία, την Ασφάλεια, τα θεμελιώδη δικαιώματα, το Περιβάλλον, τη δημοκρατία και το κράτος δικαίου. Οι πάροχοι τέτοιων συστημάτων, εκτός από την αξιολόγηση συμμόρφωσης πριν από τη διάθεση στην αγορά, θα πρέπει να διαμορφώνουν συστήματα διαχείρισης κινδύνων/δεδομένων, τεχνική τεκμηρίωση για την απόδειξη της συμμόρφωσης, αρχεία καταγραφής και να σχεδιάζουν τα συστήματα ΤΝ έτσι ώστε να επιτυγχάνεται ασφάλεια των συστημάτων, αλλά και ανθρωπινή επίβλεψη επ' αυτών. Ταυτόχρονα, δεν πρέπει να παραβλέπονται οι υποχρεώσεις σε σχέση με συστήματα «κινδύνου διαφάνειας», όπως π.χ. η υποχρέωση διαφάνειας έναντι του καταναλωτή που αλληλοεπιδρά με ένα chatbot, αλλά και ζητήματα ευθύνης που μπορεί να απορρέουν από τη χρήση τέτοιων συστημάτων. Μια κρίσιμη απαίτηση που συχνά διαφεύγει την προσοχή είναι οι απαιτήσεις διαφάνειας: Οι μεν πάροχοι συστημάτων ΤΝ που παράγουν συνθετικό περιεχόμενο ήχου, εικόνας, βίντεο ή κειμένου πρέπει να διασφαλίζουν ότι είναι δυνατόν να ανιχνευθεί πως το περιεχόμενο έχει παραχθεί ή υποστεί χειρισμό τεχνητά, οι δε φορείς εφαρμογής τέτοιων συστημάτων που δημιουργούν ή χειρίζονται περιεχόμενο εικόνας, ήχου ή βίντεο, που συνιστά προϊόν βαθυπαραποίησης (deepfake), πρέπει να γνωστοποιούν ότι το περιεχόμενο έχει παραχθεί ή υποστεί χειρισμό τεχνητά.

⁷⁷ PRESS RELEASE 553/26, 29/06/2026, Artificial Intelligence: Council gives final green light to simplify and streamline rules.

3.2. Η Πράξη για την Τεχνητή Νοημοσύνη και ο διεθνής κατακερματισμός των ρυθμιστικών προσεγγίσεων

Η AI Act, με σκοπό να αποτρέψει την καταστρατήγηση των ρυθμίσεών της, θεσπίζει ένα εκτενές, εξωεδαφικό πεδίο εφαρμογής. Ειδικότερα, όταν τα «στοιχεία εξόδου» (outputs) ενός εργαλείου Τεχνητής Νοημοσύνης, ακόμη και αν αυτό δεν διατίθεται στην ευρωπαϊκή αγορά, ούτε παρέχεται εντός της Ευρωπαϊκής Ένωσης, χρησιμοποιούνται από φορέα εντός της Ευρωπαϊκής Ένωσης, το εν λόγω εργαλείο εμπίπτει στο πεδίο εφαρμογής της Πράξης για την Τεχνητή Νοημοσύνη. Η ΕΕ είχε, δε, καλλιεργήσει την προσδοκία ότι η AI Act θα ανέπτυξε την επιδραστικότητα του Γενικού Κανονισμού για την Προστασία Δεδομένων (ΓΚΠΔ) μέσω του περίφημου «Brussels Effect», δηλαδή την ικανότητα της ΕΕ να «εξάγει» μονομερώς ρυθμιστικά πρότυπα σε τομείς πολιτικής, όπως, π.χ., η προστασία των προσωπικών δεδομένων (GDPR), ο ανταγωνισμός, η περιβαλλοντική πολιτική κλπ. Η ρυθμιστική επιρροή περιλαμβάνει τόσο τη διαμόρφωση και υιοθέτηση πολιτικής όσο και την έμμεση επιβολή των κανόνων που έχει θεσπίσει.⁷⁸

Ωστόσο, εύλογα μπορεί να αναρωτηθεί κανείς εάν η ΕΕ μπορεί να επιβάλει το Δίκαιό της σε έναν κόσμο με διαφορετικές κανονιστικές, οικονομικές και δικαιοπολιτικές προσεγγίσεις, ένα ερώτημα που εξετάζει στην παρούσα μελέτη ο Βαγγέλης Παπακωνσταντίνου. Ενώ η ΤΝ, όπως και άλλες τεχνολογίες, είναι παγκόσμιας εμβέλειας και δεν γνωρίζει σύνορα, οι ρυθμίσεις είναι εθνικές ή περιφερειακές. Ο κατακερματισμός και η διαφορά των προσεγγίσεων στο επίπεδο των ρυθμίσεων συνιστούν ένα μείζον πρόβλημα από την άποψη της τελέσφορης προστασίας των δικαιωμάτων και των συμφερόντων των προσώπων.

Είναι, δε, αξιοσημείωτο ότι οι μεγάλοι «κρατικοί παίκτες» στην ΤΝ έχουν διαμετρικά αντίθετες αντιλήψεις αναφορικά με τη ρυθμιστική ανάγκη και έκταση. Στις ΗΠΑ ο πρόεδρος Donald Trump, υιοθετώντας την προσέγγιση ότι οι δεσμευτικές ρυθμίσεις μπορεί να αποτελέσουν εμπόδιο στην αμερικανική ηγεμονία στην ΤΝ, εξέδωσε τον Ιανουάριο του 2025 νέο εκτελεστικό

⁷⁸ Bradford, A. (2020). *The Brussels Effect: How the European Union Rules the World* (pp. 1-2). Oxford University Press.

διάταγμα⁷⁹ υπό τον τίτλο «Removing Barriers to American Leadership in Artificial Intelligence», με ρητό στόχο την απομάκρυνση ρυθμιστικών εμποδίων, ενώ την ίδια λογική υιοθετεί το America's AI Action Plan του 2025. Ενδιαφέρον παρουσιάζει, ωστόσο, ότι σε πολιτειακό επίπεδο καταγράφεται ένας μεγάλος αριθμός ρυθμίσεων που συχνά αφορούν τομεακά ζητήματα, όπως η χρήση προϊόντων/εφαρμογών βαθυπαραποίησης στις πολιτικές εκστρατείες ή ζητήματα χρήσης ΤΝ στον εργασιακό τομέα, ενώ υπάρχουν Πολιτείες που έχουν υιοθετήσει κανονιστικά πλαίσια εκτενή και δεσμευτικά που προσομοιάζουν με την ενωσιακή νομοθεσία.⁸⁰

Ενδιαφέρουσα είναι και η προσέγγιση της Κίνας: Με αφετηρία το 2017, η Κίνα, που φιλοδοξεί να ηγηθεί σε παγκόσμιο επίπεδο στην ΤΝ,⁸¹ έχει αρχίσει να εξυφαίνει ένα πλέγμα ρυθμίσεων, συστάσεων, πολιτικών και προτύπων.⁸² Αν και δεν υπάρχει ένα ενιαίο και ολοκληρωμένο νομοθέτημα, έχουν υιοθετηθεί πολλοί ειδικοί κανόνες και ρυθμιστικά πλαίσια. Το κανονιστικό πλαίσιο της Κίνας δίνει ιδιαίτερο βάρος σε συγκεκριμένους άξονες: κρατική ασφάλεια και έλεγχος περιεχομένου, προστασία δεδομένων και προσωπικών πληροφοριών, συστήματα συστάσεων μέσω αλγορίθμων, υποχρεώσεις των παρόχων για ασφάλεια και συμμόρφωση, καταχώριση και εκτιμήσεις επικινδυνότητας για συστήματα «υψηλού κινδύνου/υψηλής επιρροής» και σήμανση ή ικνηλασιμότητα του περιεχομένου που παράγεται με την υποστήριξη ΤΝ. Μεταξύ 2020 και 2022 καταγράφεται εκτεταμένη δραστηριότητα επιβολής κυρώσεων σε τεχνολογικές εταιρείες και μάλιστα με πρωτοφανή ταχύτητα.⁸³ Η κινεζική κυβέρνηση έχει εκτεταμένη και βαθιά εμπλοκή στο οικοσύστημα της ΤΝ, αναλαμβάνοντας πολλαπλούς ρόλους και ενεργώντας ως ρυθμιστική Αρχή, φορέας χάραξης πολιτικής, επενδυτής, προμηθευτής και πελάτης.⁸⁴

⁷⁹ Είχε προηγηθεί η ανάκληση του Executive Order 14110 του τότε προέδρου των ΗΠΑ Joe Biden, της 30ής Οκτωβρίου 2023, για την «ασφαλή, προστατευμένη και αξιόπιστη ανάπτυξη και χρήση της ΤΝ». Το διάταγμα αυτό έδινε οδηγίες σε πολλές ομοσπονδιακές υπηρεσίες για ζητήματα όπως τα πρότυπα ασφαλείας, οι κίνδυνοι από ισχυρά foundation models, η προστασία καταναλωτών, εργαζομένων, η χρήση της ΤΝ από το Δημόσιο και η προστασία των δικαιωμάτων.

⁸⁰ Razis, E., & Cooper J. C. (2025, June 4). The Federalist's Dilemma: State AI Regulation & Pathways Forward. *George Mason Law & Economics Research Paper*, 25(07). *Harvard Journal of Law & Public Policy*, Forthcoming. Available at SSRN: <https://doi.org/10.2139/ssrn.5283472>

⁸¹ Sheehan, M. (2023, September 12). What the U.S. Can Learn From China About Regulating AI. *Foreign Policy*. <https://foreignpolicy.com/2023/09/12/ai-artificial-intelligence-regulation-law-china-us-schumer-congress/>

⁸² Zhang, A. H. (2025). The promise and perils of China's regulation of artificial intelligence. *Columbia Journal of Transnational Law*, 63(1), 1-56. <https://www.jtl.columbia.edu/volume-63/the-promise-and-perils-of-chinas-regulation-of-artificial-intelligence>

⁸³ Σημειώνεται μάλιστα ότι η ταχύτητα αυτής της κινεζικής καταστολής στον τεχνολογικό τομέα ξεπέρασε κατά πολύ τον πιο συγκρατημένο ρυθμό των αμερικανικών και ευρωπαϊκών ρυθμιστικών Αρχών, οι οποίες συχνά περιορίζονται από χρονοβόρες νομοθετικές διαδικασίες και δικαστική εποπτεία. Βλ. Zhang, A. H. (2024). *High wire: How China regulates big tech and governs its economy*. Oxford University Press.

⁸⁴ Beraja, M., Yang, D. Y., & Yuchtman, N. (2023). Data-intensive innovation and the state: Evidence from AI firms in China. *The Review of Economic Studies*, 90(4), 1701-1723. <https://doi.org/10.1093/restud/rdac056>

3.3. Πώς η Ευρώπη ρυθμίζει τις ψηφιακές τεχνολογίες (τρίτων) και πώς αυτό επηρεάζει την ψηφιακή της κυριαρχία

3.3.1. Εισαγωγή

Αν ήθελε κάποιος να καταλάβει τι είδους κανονιστικό πλαίσιο ισχύει για την Τεχνητή Νοημοσύνη στην Ευρώπη σήμερα, θα μπορούσε να ξεκινήσει φέρνοντας στο μυαλό του ένα εργαλείο σαν το σφυρί. Υποθέτω ότι όπως νιώθουμε εμείς σήμερα για την ΤΝ, κάπως έτσι θα ένιωσαν οι άνθρωποι όταν το εφνύραν: ένα πολυ-εργαλείο που μπορούσαν να το κουβαλούν μαζί τους και να το χρησιμοποιούν για οτιδήποτε, από το άνοιγμα των καρπών και τις εργασίες στο σπίτι μέχρι το κυνήγι και τον πόλεμο.

Πώς αντιμετωπίζει ο Ευρωπαίος νομοθέτης σήμερα την ΤΝ; Ρυθμίζει τα είδη της. Δηλαδή, στο παραπάνω παράδειγμα του σφυριού ο νόμος μάς λέει ότι αν ένα σφυρί έχει ενισχυθεί, για παράδειγμα, με πρόκες, επομένως πρόκειται ξεκάθαρα για πολεμικό όπλο, τότε η χρήση του απαγορεύεται απολύτως. Αν πάλι έχει μόνο ενισχυθεί με άλλα υλικά ώστε να αντέχει στη χρήση, επομένως δεν μπορούμε να ξέρουμε ακριβώς πώς θα χρησιμοποιηθεί, ξέρουμε, όμως, ότι είναι ισχυρότερο από τα υπόλοιπα σφυριά, τότε θα επιτρέπεται η χρήση του μόνο υπό όρους. Αν πάλι είναι ένα απλό, συνηθισμένο σφυρί, τότε επιτρέπεται η χρήση του χωρίς άλλον περιορισμό.

Το θετικό με αυτήν τη risk-based προσέγγιση είναι ότι προσφέρει πρακτικές απαντήσεις άμεσα, δηλαδή μας υποδεικνύει τι ακριβώς επιτρέπεται να κάνουμε και τι όχι με κάθε ένα από τα σφυριά που υπάρχουν γύρω μας. Αυτός, όμως, είναι ταυτόχρονα και ο βασικότερος περιορισμός της: Αν τυχόν επινοηθεί νέο σφυρί που δεν εντάσσεται εύκολα στις κατηγορίες του νόμου, τότε θα πρέπει οι κανόνες να αναθεωρηθούν.

Ποια θα ήταν μια άλλη, διαφορετική προσέγγιση; Να μη ρυθμίσει ο νόμος τα είδη του σφυριού, αλλά τις χρήσεις του. Δηλαδή, ο νομοθέτης πρώτα θα εντόπιζε τι ακριβώς άλλαξε με την εφεύρεση του σφυριού: Ποιες νέες πράξεις έφερε η χρήση του; Ποιες νέες συμπεριφορές; Αν επέτρεψε νέες πράξεις και συμπεριφορές, τότε χρειάζονται νέοι κανόνες για να τις ρυθμίσουν. Αν όχι, τότε αρκούν οι παλιοί.

Η διαφορά, βέβαια, σε αυτήν την περίπτωση είναι ότι οι ανθρώπινες πράξεις και συμπεριφορές δεν μπορούν να ρυθμιστούν με επίπεδα κινδύνου. Κανείς δεν μπορεί να τις προβλέψει όλες, ώστε να τις τοποθετήσει σε κατηγορίες. Πώς θα μπορούσε κάποιος να προβλέψει και να περιγράψει εξαντλητικά ανθρώπινες αντιδράσεις; Γι' αυτόν τον λόγο σε αυτήν την κανονιστική προσέγγιση θα χρειαζόμασταν άλλου τύπου νομοθεσία, όχι risk-based, αλλά principles-based. Δηλαδή, γενικούς κανόνες γενικής εφαρμογής.

Ένα παράδειγμα γνωστό σε όλους αυτής της τελευταίας περίπτωσης είναι η νομοθεσία για τα προσωπικά δεδομένα. Επεξεργασίες προσωπικών δεδομένων πάντα κάναμε, αφού είμαστε άνθρωποι και ζούμε σε κοινωνίες, όμως ειδικό νόμο γι' αυτές ποτέ δεν είχαμε. Όταν οι υπολογιστές επέτρεψαν την αυτοματοποιημένη επεξεργασία προσωπικών δεδομένων, τότε νέος νόμος ρύθμισε τις νέες πράξεις που οι υπολογιστές επέτρεψαν. Αυτό, όμως, το έκανε με γενικές αρχές, όχι με κατηγοριοποίηση διαφορετικών επεξεργασιών.⁸⁵ Αντιθέτως, στη risk-based ρύθμιση της ΤΝ λαμβάνουμε υπ' όψιν μας συγκεκριμένες εφαρμογές ΤΝ, χωρίς γενικές αρχές για τη χρήση της. Κατηγοριοποιούμε, δηλαδή, τα είδη σφυριού χωρίς μια γενική αρχή ότι, για παράδειγμα, η χρήση του σφυριού εκτός πολεμικής σύρραξης γίνεται με ευθύνη του χρήστη.

Οριστική απάντηση δεν υπάρχει για το αν είναι καλύτεροι οι risk-based ή οι principles-based νόμοι. Σε κάποιες περιπτώσεις φαίνεται να είναι καταλληλότεροι οι πρώτοι (όταν, για παράδειγμα, νομοθετούμε για προϊόντα) και σε κάποιες άλλες οι δεύτεροι (όταν ρυθμίζουμε συμπεριφορές). Παρ' όλα αυτά, δεν μπορεί να μη σημειωθεί ότι, ιστορικά, μέχρι σήμερα δεν είχαμε ούτε νόμο για τα σφυριά ούτε για τα άλογα⁸⁶ ούτε για το ηλεκτρικό ρεύμα ούτε για τα αεροπλάνα ούτε για τον υπολογιστή, αλλά ούτε για το Ίντερνετ.⁸⁷

3.3.2. Η Πράξη για την Τεχνητή Νοημοσύνη

Αφού εξηγήθηκε το risk-based σε αντιπαραβολή με το principles-based μοντέλο νομοθέτησης και ήδη προοικονομήθηκε ότι ο ευρωπαϊκός νόμος για την Τεχνητή Νοημοσύνη ακολουθεί το πρώτο, λίγα λόγια απαιτούνται για τον νόμο αυτόν καθ' εαυτόν. Η Πράξη για την Τεχνητή Νοημοσύνη

⁸⁵ Εννοείται ότι επιμέρους εξαιρέσεις υπάρχουν (π.χ., για διαβιβάσεις εκτός Ευρώπης ή, στο παρελθόν, για επεξεργασίες τύπου data matching), όμως οι γενικές αρχές ισχύουν σε όλες απολύτως τις επεξεργασίες. Σημειώνεται, άλλωστε, ότι μπορεί η ευρωπαϊκή νομοθεσία για τα προσωπικά δεδομένα [δηλαδή ο Γενικός Κανονισμός για την Προστασία Δεδομένων (Κανονισμός (ΕΕ) 2016/679) και η Οδηγία (ΕΕ) 2016/680] να θεωρούνται risk-based, επειδή διακρίνουν, για παράδειγμα, μεταξύ «κοινών» και «ευαίσθητων» δεδομένων, όμως οι γενικές αρχές ισχύουν πάντα.

⁸⁶ Βλ. τη διάσημη απάντηση για τον Law of the Horse του Frank H. Easterbrook, ως αντίλογο στον Lessig ότι χρειαζόμαστε ειδικό Δίκαιο του κυβερνοχώρου. Easterbrook, F. H. (1996). Cyberspace and the Law of the Horse. *University of Chicago Legal Forum*, 1996(1), 207-216.

⁸⁷ Έχουμε όμως για τις online πλατφόρμες [βλ. την Πράξη για τις Ψηφιακές Αγορές, Κανονισμός (ΕΕ) 2022/1925, σε συνδυασμό με την Πράξη για τις Ψηφιακές Υπηρεσίες, Κανονισμός (ΕΕ) 2022/2065], και πάλι περίπου στη λογική του νόμου για την ΤΝ.

υιοθετήθηκε συγκριτικά σύντομα, μέσα σε τρία χρόνια από την υποβολή της αντίστοιχης πρότασης από την Ευρωπαϊκή Επιτροπή,⁸⁸ γεγονός που καταδεικνύει σύμπνοια μεταξύ των κρατών-μελών. Αντιρρήσεις σε επιμέρους θέματα προφανώς υπήρξαν, και οι αλλαγές από το αρχικό κείμενο ήταν αρκετές, όμως το βασικό μοντέλο ποτέ δεν αμφισβητήθηκε. Η risk-based λύση, με τη διάκριση τριών κατηγοριών κινδύνων ώστε εκεί να ενταχθεί κάθε εφαρμογή ΤΝ που σήμερα γνωρίζουμε, φάνηκε η μόνη ικανή να ανταποκριθεί στα προβλήματα, υποθετικά και μη, που εγκυμονεί η ΤΝ για τον άνθρωπο. Ήταν, άλλωστε, και η πιο συνεπής στους ήδη θεσπισμένους ευρωπαϊκούς κανόνες νομοθέτησης, από τη στιγμή που η Ευρωπαϊκή Επιτροπή επέλεξε να αντιμετωπίσει την ΤΝ ως προϊόν.⁸⁹

Η Πράξη, επομένως, διακρίνει τρεις κατηγορίες κινδύνου, υψηλού, μεσαίου και χαμηλού. Επίσης, προβλέπει την πλήρη απαγόρευση εφαρμογών που, για παράδειγμα, απευθύνονται στο υποσυνείδητο ενός προσώπου ή που εκμεταλλεύονται οποιοδήποτε από τα τρωτά σημεία ενός φυσικού προσώπου ή μιας συγκεκριμένης ομάδας προσώπων.⁹⁰ Από εκεί και πέρα, οι κανόνες ακολουθούν τον εικαζόμενο κίνδυνο, δηλαδή το ρίσκο από τη χρήση της συγκεκριμένης εφαρμογής ΤΝ. Οι υψηλού κινδύνου υπόκεινται σε αυστηρές προϋποθέσεις, οι μεσαίου σε λιγότερο αυστηρές, οι χαμηλού σε ελάχιστες. Ταυτόχρονα, εισάγεται ένα μοντέλο κρατικής παρακολούθησης που θυμίζει εκείνο των προσωπικών δεδομένων: Κρατικές Αρχές έχουν την αρμοδιότητα να παρακολουθούν την εφαρμογή του νόμου σε εθνικό επίπεδο.

Ο Κανονισμός για την ΤΝ παρουσιάζει νομικές ιδιαιτερότητες, όπως για παράδειγμα το γεγονός ότι ιδρύεται Ευρωπαϊκό Γραφείο για ιδιαίτερα σύνθετες περιπτώσεις ΤΝ (συνεπώς, η αντίστοιχη αρμοδιότητα αφαιρείται από τα κράτη-μέλη) ή τα regulatory sandboxes,⁹¹ όμως η αναλυτική παρουσίασή τους σε αυτό εδώ το κείμενο ξεπερνά κατά πολύ τους σκοπούς του. Το ίδιο, άλλωστε, ισχύει για την πρόσφατη απόπειρα χαλάρωσης των διατάξεών του, πριν καλά καλά εφαρμοστούν,⁹² σύμφωνα με τις πρόσφατες

⁸⁸ European Commission (2021). *Proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts* [COM(2021) 206 final, 2021/0106(COD)]. Publications Office of the European Union. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>

⁸⁹ Επομένως, ακολούθησε το μοντέλο της NLF (New Legislative Framework, βλ. Απόφαση 768/2008/EK του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου, της 9ης Ιουλίου 2008), κάτι που άλλωστε είναι ορατό και στην PLD [Product Liability Directive, Οδηγία (ΕΕ) 2024/2853], η οποία εξίσου μεταχειρίζεται τα συστήματα ΤΝ (αλλά και το λογισμικό) ως προϊόντα.

⁹⁰ Βλ. άρθρο 5 του Κανονισμού για την ΤΝ.

⁹¹ Πρόκειται για ελεγχόμενα πλαίσια μέσα στα οποία επιχειρήσεις ή οργανισμοί μπορούν να δοκιμάζουν νέες τεχνολογίες, προϊόντα και υπηρεσίες υπό την εποπτεία αρμόδιων Αρχών, πριν αυτά κυκλοφορήσουν πλήρως στην αγορά.

⁹² Ο Κανονισμός έχει σχεδιαστεί να εφαρμοστεί σταδιακά, μέχρι το 2027 ή ακόμα και, σε ειδικές περιπτώσεις, το 2030.

πρωτοβουλίες Omnibus που τελούν υπό συζήτηση.⁹³ Η γενική εικόνα του risk-based μοντέλου παραμένει, και μάλιστα χωρίς αμφισβήτηση· το μόνο που ενδεχομένως να αλλάξει είναι οι επιμέρους κανόνες κάποιων (high risk) κατηγοριών ή κάποιες διευκολύνσεις σε μικρομεσαίες επιχειρήσεις (ΜμΕ).

3.3.3. Τι γίνεται εκτός Ευρώπης;

Η Πράξη για την Τεχνητή Νοημοσύνη επιβεβαιώνει τη θεώρηση του σημερινού κόσμου ως τριπολικού, όπου η Αμερική κατέχει τη στρατιωτική ισχύ, η Κίνα την οικονομική και η Ευρώπη θέτει τους κανόνες. Η Πράξη ήταν η πρώτη που ρύθμιζε, με τον τρόπο που ρύθμιζε, την ΤΝ οπουδήποτε στον κόσμο και το παράδειγμά της ήδη φαίνεται ότι ακολουθούν και άλλοι⁹⁴ – δεν είναι τυχαίο, άλλωστε, ότι οι «ελπίδες» της Ευρώπης για επανάληψη του «Brussels Effect» φαινομένου, μετά τον Γενικό Κανονισμό για την Προστασία Δεδομένων, έχουν εναποτεθεί στην Πράξη για την Τεχνητή Νοημοσύνη.

Ίδια, όμως, με εκείνα των προσωπικών δεδομένων είναι και τα προβλήματα: Οι ΗΠΑ αρνούνται πεισματικά ομοσπονδιακό νόμο για την ΤΝ, όπως ακριβώς κάνουν μέχρι και σήμερα (από το 1970) για τα προσωπικά δεδομένα. Ο αντίλογός τους είναι ότι δεν επιθυμούν να περιορίσουν με νόμο μια βιομηχανία που τους φέρνει παγκόσμιο ανταγωνιστικό πλεονέκτημα. Άλλωστε, το αίτημα για χαλάρωση των νόμων σχετικά με την ΤΝ επαναλαμβάνεται ποικιλοτρόπως και στην Ευρώπη, όπως αποδεικνύει η Omnibus πρόταση που αναφέρθηκε παραπάνω.

Από την άλλη μεριά, η Κίνα ακολουθεί τη δική της οδό τής κατά περίπτωση ρύθμισης της ΤΝ [για παράδειγμα, με νόμους ειδικά για τις προτάσεις του αλγορίθμου (algorithm recommendations), deepfakes κλπ.]. Σημειώνεται, όμως, ότι, ενώ διαθέτει εδώ και χρόνια γενικό νόμο για τα προσωπικά δεδομένα, δεν έχει αφετηρία τα ανθρώπινα δικαιώματα, τα οποία ιδεολογικά δεν αναγνωρίζει, αλλά την προστασία άλλων έννομων συμφερόντων, και κυρίως του ίδιου του κράτους.

⁹³ Πρόταση Κανονισμού για την απλούστευση της εφαρμογής εναρμονισμένων κανόνων σχετικά με την Τεχνητή Νοημοσύνη (Digital Omnibus για την ΤΝ), COM(2025) 836 final, Βρυξέλλες, 19.11.2025.

⁹⁴ Επίπεδα κινδύνου, και επομένως risk-based approach, ακολουθεί και ο νοτιοκορεατικός νόμος του Ιανουαρίου 2026. Στις ΗΠΑ δεν υπάρχει νόμος σε ομοσπονδιακό επίπεδο, όμως οι δύο νόμοι σε επίπεδο state, που ισχύουν σε Κολοράντο και Καλιφόρνια, ακολουθούν ήδη το risk-based ευρωπαϊκό μοντέλο. Το Συμβούλιο της Ευρώπης έχει υιοθετήσει AI Convention (Συμβούλιο της Ευρώπης, Σύμβαση-Πλαίσιο για την Τεχνητή Νοημοσύνη, τα Ανθρώπινα Δικαιώματα, τη Δημοκρατία και το Κράτος Δικαίου, CETS No. 225), η οποία μπορεί να μην ακολουθεί το risk-based μοντέλο, παραμένει όμως σε υψηλό (όχι ιδιαίτερα πρακτικό) επίπεδο, επιδιώκοντας την προστασία θεμελιωδών αξιών και δικαιωμάτων.

3.3.4. Η Ευρώπη επιχειρεί να ρυθμίσει τις τεχνολογίες τρίτων – Η ψηφιακή της κυριαρχία

Όλες οι εφαρμογές ΤΝ σήμερα, τουλάχιστον τα Μεγάλα Γλωσσικά τους Μοντέλα που έχουν απασχολήσει και την κοινή γνώμη, έχουν αναπτυχθεί εκτός Ευρώπης, κυρίως στην Αμερική και την Κίνα (με μια μοναδική ευρωπαϊκή εξαίρεση, η οποία, όμως, δεν έχει τον αριθμό χρηστών των άλλων). Ίδια, άλλωστε, όπως είναι γνωστό, ισχύουν και για κάθε άλλη ψηφιακή τεχνολογία σήμερα. Όλες οι online πλατφόρμες, (σχεδόν) όλα τα βασικά εργαλεία επεξεργασίας προσωπικών δεδομένων (αν αναλογιστεί κανείς σε ποιους ανήκουν τα διάφορα πακέτα λογισμικού τύπου «Office») και (σχεδόν) όλες οι βασικές τεχνολογίες δικτύου (και επομένως και κυβερνοασφάλειας) αναπτύσσονται εκτός Ευρώπης. Με άλλα λόγια, η Ευρώπη επιχειρεί να ρυθμίσει τις ψηφιακές τεχνολογίες τρίτων, τεχνολογίες που έχουν αναπτυχθεί και συνεχίζουν να αναπτύσσονται μακριά της, των οποίων δηλαδή οι εταιρείες-ιδιοκτήτες δεν έχουν την έδρα τους εντός της εδαφικής της επικράτειας.

Είναι άραγε κακό αυτό; Θεωρητικά ίσως όχι, αφού η ευρωπαϊκή νομοθεσία ισχύει στο ευρωπαϊκό έδαφος και προστατεύει τους Ευρωπαίους πολίτες. Δηλαδή, όλες οι παραπάνω μη ευρωπαϊκές εταιρείες, εφόσον επιθυμούν να διαθέτουν τα προϊόντα και τις υπηρεσίες τους στην Ευρώπη, οφείλουν να συμμορφώνονται με τους ευρωπαϊκούς κανόνες. Άλλωστε, τα 500 εκατομμύρια Ευρωπαίων με το μορφωτικό και οικονομικό επίπεδο που διαθέτουν δεν μπορούν εύκολα να αγνοηθούν από οποιαδήποτε επιχείρηση. Με άλλα λόγια, η Ευρώπη μπορεί να είναι μόνο χρήστης (user) ψηφιακών τεχνολογιών, των οποίων οι developers δεν είναι Ευρωπαίοι, τουλάχιστον, όμως, είναι ένας power(ful) user.

Επομένως, η τελική εμπειρία που απολαμβάνει ο Ευρωπαίος χρήστης είναι σύμφωνη με τον ευρωπαϊκό νόμο. Το πρόβλημα, όμως, είναι βαθύτερο, και είναι πολιτικό, όχι νομικό. Τελικά, αφορά τον έλεγχο και σε τελική ανάλυση την (ψηφιακή) κυριαρχία· όχι την (απλή) χρήση.

Πρακτικά, αν η κυβέρνηση της έδρας της επιχείρησης ενός μοντέλου ΤΝ (ή μιας άλλης ψηφιακής τεχνολογίας) απαιτήσει «ειδική μεταχείριση», η διοίκηση της επιχείρησης είναι πολύ αμφίβολο αν θα αντισταθεί, ακόμα και τις καλύτερες προθέσεις να έχει.⁹⁵ Τέτοια ειδική μεταχείριση μπορεί να είναι είτε τεχνική είτε εμπορική – για παράδειγμα, backdoors στον κώδικά της⁹⁶ ή απαγόρευση διάθεσης σε συγκεκριμένες χώρες (περιπτώσεις που και οι δύο έχουν συμβεί).

⁹⁵ Εντελώς ενδεικτικά βλ. Pete Hegseth wages war on Anthropic: Should AI labs unquestioningly obey the Pentagon's orders? (2026, February 24). *The Economist*. <https://www.economist.com/business/2026/02/24/pete-hegseth-wages-war-on-anthropic>

⁹⁶ Μια κερκόπορτα (backdoor) στον κώδικα είναι μια κρυφή, σκόπιμη μέθοδος παράκαμψης της ταυτοποίησης (έλεγχος ταυτότητας ή δικαιώματα χρήστη) ή κρυπτογράφησης σε ένα υπολογιστικό σύστημα, μια εφαρμογή λογισμικού ή μια ενσωματωμένη συσκευή, με σκοπό την απόκτηση μη εξουσιοδοτημένης πρόσβασης.

Επιπλέον, εξίσου πρακτικά είναι ξεκάθαρο ότι, εφόσον παρέχονται εθνικά κεφάλαια και πολιτική υποστήριξη σε τέτοιες εταιρείες, οι κυβερνήσεις τους θα έχουν πρώτη και προνομιακή πρόσβαση στις αντίστοιχες τεχνολογίες, διαιωνίζοντας (και βαθαίνοντας) το ψηφιακό χάσμα, δηλαδή την απόστασή τους από τις τρίτες χώρες (σε αυτήν την περίπτωση από την Ευρώπη).

Αλλά και σε νομικό επίπεδο νομίζω ότι άλλη ισχύ θα είχαν τυχόν απαγορεύσεις στη χώρα έδρας και άλλη σε μια χώρα πώλησης, ακόμα και τόσο σημαντική όσο η Ευρώπη. Δηλαδή, υποθέτω ότι μια απαγόρευση εθνικού αμερικανικού Δικαίου σε μια εφαρμογή ΤΝ στην Αμερική θα επηρέαζε ολόκληρο τον κώδικά της, ενώ μια αντίστοιχη απαγόρευση στην Ευρώπη απλά θα επηρέαζε το μοντέλο που διατίθεται στην Ευρώπη (ενδεχομένως, κάνοντάς το χειρότερο, μέσω ακρωτηριασμού, ενώ στην πρώτη περίπτωση το μοντέλο θα έπρεπε να παραμείνει λειτουργικό).

Τα παραπάνω επηρεάζουν και την επίκαιρη συζήτηση περί ψηφιακής κυριαρχίας (digital sovereignty). Παρότι η ψηφιακή αυτάρκεια και αυτονομία αποτελεί ευρωπαϊκό στόχο τουλάχιστον από το 2019 και ενώ το διεθνές, πλέον, κλίμα μετέτρεψε αυτόν τον στόχο σε ανάγκη, τα εμπόδια ώστε να επιτευχθεί φαίνονται αζεπέραστα. Σε συνθήκες ψηφιακής αποικιοκρατίας (digital colonialism), όπως περιγράφηκαν παραπάνω, πώς θα μπορούσε άραγε η Ευρώπη από χρήστη (έστω, power)⁹⁷ να γίνει προγραμματιστής (developer); Η απόσταση που θα πρέπει να καλυφθεί είναι τεράστια. Επιπλέον, αυτός είναι ένας στόχος στον βωμό του οποίου θα πρέπει να θυσιαστεί τεράστιο οικονομικό αλλά και πολιτικό κεφάλαιο – ασφαλώς δεν θα είναι εύκολο, για παράδειγμα, να πειστούμε ξαφνικά να μην αποθηκεύουμε τα καθημερινά μας αρχεία στις εφαρμογές cloud που ήδη χρησιμοποιούμε ή να αλλάξουμε άλλες καθημερινές μας συνήθειες, όπως εργαλεία «Office», online κοινωνικά δίκτυα, υπηρεσίες ανταλλαγής μηνυμάτων κοκ.

Είναι η προσπάθεια απέλπιδα; Όχι αναγκαστικά. Αφενός, το πρώτο βήμα έγινε, το πρόβλημα εντοπίστηκε και ο στόχος τέθηκε. Αυτό από μόνο του είναι εξαιρετικά σημαντικό, αφού πριν από λίγα χρόνια συζητήσεις περί εντοπισμού δεδομένων (data localisation) και ψηφιακής κυριαρχίας (digital sovereignty) αντιμετωπίζονταν ως σχεδόν σκοταδιστικές. Αφετέρου, η Ευρώπη μπορεί σήμερα να υστερεί στις «μεγάλες» εφαρμογές γενικού πληθυσμού, όμως αυτό είναι μια σχετικά πρόσφατη εξέλιξη: Πριν από τα smartphones η κινητή τηλεφωνία ήταν ευρωπαϊκή υπόθεση, ενώ άλλες τεχνολογίες, για παράδειγμα το Minitel στη Γαλλία ή το ISDN στη Γερμανία, άντεξαν όσο μπόρεσαν. Παρά την αυτο-λύπηση και την ομφαλοσκόπηση που χαρακτηρίζει την Ευρώπη αυτήν τη χρονική περίοδο, μέχρι και πριν

⁹⁷ Power user – ισχυρός χρήστης.

από λίγα χρόνια η αισιοδοξία και οι ελπίδες για το τεχνολογικό, και μάλιστα το ψηφιακό, μέλλον ήταν (και) ευρωπαϊκές.

3.3.5. Υπάρχει ειδική πολιτική μέσα στη γενική πολιτική;

Και η Ελλάδα; Μπορεί να έχει τη δική της πολιτική για την ΤΝ μέσα στην ευρύτερη ευρωπαϊκή πολιτική; Προφανώς στα μεγάλα ζητήματα όχι – οι αποφάσεις (η ρύθμιση, το μοντέλο διάκρισης κινδύνων, ακόμα και οι κατηγοριοποιήσεις) έχουν αποφασιστεί αλλού, κεντρικά. Ακόμα και στα μικρότερα, επιμέρους ζητήματα, και πάλι θα χωρήσει ευρωπαϊκός συντονισμός – δεν νοείται η θέση ενός κράτους-μέλους να είναι ριζικά διαφορετική από εκείνη κάποιου άλλου σε οποιοδήποτε θέμα ΤΝ, εξαιτίας του Κανονισμού και των πολιτικών, κοινών, αποφάσεων που τον στηρίζουν.

Ούτε, άλλωστε, ο γενικότερος πολιτικός στόχος περί υιοθέτησης της ΤΝ είναι ιδιαίτερα πρωτότυπος. Σήμερα, οποιαδήποτε κυβέρνηση ή επιχείρηση σέβεται τον εαυτό της και σκέφτεται το μέλλον αναζητά τρόπους χρήσης της ΤΝ ούτως ή άλλως (εννοείται, εντός του εκάστοτε ρυθμιστικού πλαισίου). Επομένως, η Ελλάδα καλά κάνει και ήδη προωθεί την κατά το δυνατόν εκτεταμένη, και ταυτόχρονα νόμιμη, χρήση της ΤΝ σε δημόσιο και ιδιωτικό τομέα – όπως κάνουν όλοι οι άλλοι, άλλωστε.

Τι απομένει; Η επιθετικότητα. Όταν το πλαίσιο υπάρχει και το θέτουν άλλοι (έστω και με τη συμμετοχή μας), το μόνο που μπορεί κανείς να κάνει είναι να τοποθετηθεί μέσα σε αυτό. Πού θα σταθεί; Στο ένα άκρο, στη μέση ή στο άλλο άκρο του;

Επιθετική υιοθέτηση της ΤΝ θα ήταν, για παράδειγμα, η επιδότηση μαθητών για τη συνδρομή τους σε μοντέλα ΤΝ. Η χρήση μοντέλων ΤΝ, συνδρομητικά, από σχολεία. Η στρατηγική συνεργασία πανεπιστημιακών τμημάτων και ακαδημαϊκών εργαστηρίων με εταιρείες ΤΝ. Η παροχή ανοιχτών δεδομένων του Δημοσίου σε εφαρμογές ΤΝ (σε μια προσπάθεια, άλλωστε, προστασίας της ίδιας της γλώσσας μας). Η απελευθέρωση όσο το δυνατόν περισσότερων έργων που προστατεύονται από δικαιώματα διανοητικής ιδιοκτησίας, καθώς και η διάθεσή τους σε εφαρμογές ΤΝ. Οι φορολογικές απαλλαγές για τα κέρδη εταιρειών ΤΝ. Η «εξάντληση» του regulatory sandbox – γιατί, για παράδειγμα, να μην είχαμε χρηματιστηριακές εταιρείες ή τράπεζες αμιγώς ΤΝ στην Ελλάδα;

Αμυντική υιοθέτηση της ΤΝ θα ήταν, για παράδειγμα, η εκπόνηση επιπλέον ρυθμιστικού πλαισίου πάνω στο κοινοτικό ρυθμιστικό πλαίσιο. Ο με νόμο ή εσωτερικό κανονισμό αποκλεισμός της ΤΝ από τα ελληνικά σχολεία και πανεπιστήμια. Ο περιορισμός πρόσβασης των μοντέλων της σε δεδομένα του Δημοσίου.

Η λίστα μπορεί να είναι μακρά και προς τη μία και προς την άλλη κατεύθυνση. Στη βάση της βρίσκεται η πολιτική βούληση και η συν-διαμόρφωση της κοινής γνώμης. Αν η ελληνική κοινή γνώμη αντιληφθεί την ΤΝ σαν πρόβλημα, ο νομοθέτης δεν μπορεί παρά να την περιορίσει. Αν τη δει σαν ευκαιρία ανάπτυξης, θα εξαντλήσει τα περιθώριά του για τη γενικευμένη χρήση της. Στο τέλος της διαδρομής, σε λίγα χρόνια από σήμερα, έτσι κι αλλιώς η ΤΝ θα ενσωματωθεί στη ζωή μας όπως ακριβώς το σφυρί (ή το άλογο ή το ηλεκτρικό ρεύμα ή οι υπολογιστές): αβίαστα και σχεδόν ασυνείδητα, χωρίς να μας πολυαπασχολεί. Ο νόμος, δηλαδή, για την ΤΝ νομίζω ότι στο τέλος της διαδρομής θα εξαφανιστεί, χαμένος στην πλήρη ενσωμάτωσή της στην ανθρώπινη, νέα, καθημερινότητα. Αν μέχρι τότε αμυνόμαστε ή επιτιθέμεθα, αν δηλαδή, στο αρχικό παράδειγμα απ' όπου ξεκινήσαμε, όλοι αγοράσουμε σφυριά για να δούμε τι μπορούμε να κάνουμε με αυτά ή αν αποφασίσουμε να ρυθμίσουμε τις χρήσεις τους ώστε να περιοριστεί ο αριθμός τους, είναι θέμα που θα αποφασίσουμε τώρα και τα αποτελέσματά του θα αντηχούν για δεκαετίες.

Βιβλιογραφία

- Almada, M., & Petit, N. (2025). The EU AI Act: Between the rock of product safety and the hard place of fundamental rights. *Common Market Law Review*, 62(1), 85-120. <https://doi.org/10.54648/COLA2025004>
- Nemitz, P. (2018, November 28). Constitutional democracy and technology in the age of artificial intelligence. *Philosophical Transactions of the Royal Society*, 376(2133). <https://doi.org/10.1098/rsta.2018.0089>
- Papakonstantinou, V., & De Hert, P. (2024). *The Regulation of Digital Technologies in the EU: Act-ification, GDPR Mimesis and EU Law Brutality at Play*. Routledge.
- Veale, M., & Borgesius, F. Z. (2021). Demystifying the Draft EU Artificial Intelligence Act Demystifying the Draft EU Artificial Intelligence Act – Analysing the good, the bad, and the unclear elements of the proposed approach. *Computer Law Review International* 22(4), 97-112. <https://doi.org/10.9785/cr-2021-220402>
- Μήτρου, Λ. (Επιμ.) (2023). *Μπορεί ο αλγόριθμος...* Πανεπιστημιακές Εκδόσεις Κρήτης.

3.4. Το όραμα της ψηφιακής κυριαρχίας της Ευρωπαϊκής Ένωσης

Σε αντίθεση με τις προσεγγίσεις που φαίνεται να κυριαρχούν, τουλάχιστον κανονιστικά, δυτικά και ανατολικά της Ευρώπης, η Ευρωπαϊκή Ένωση είχε επενδύσει στον συνδυασμό των τεχνολογικών και των βιομηχανικών πλεονεκτημάτων της με τις υψηλής ποιότητας ψηφιακές υποδομές και ένα κανονιστικό πλαίσιο που θα βασίζεται στις θεμελιώδεις αξίες της, προκειμένου να καταστεί παγκόσμια ηγέτιδα δύναμη στον τομέα της καινοτομίας που αφορά την οικονομία των δεδομένων και τις εφαρμογές της.⁹⁸ Όπως σημειώνεται, το 2020, σε μια ανακοίνωση του Ευρωπαϊκού Κοινοβουλίου, η έννοια της τεχνολογικής ή ψηφιακής κυριαρχίας (technological/digital sovereignty) αναδεικνύεται ως μέσο προώθησης της ευρωπαϊκής ηγεσίας και της στρατηγικής αυτονομίας στον ψηφιακό τομέα. Σε αυτό το πλαίσιο, η ψηφιακή κυριαρχία αναφέρεται στην ικανότητα της Ευρώπης να δρα ανεξάρτητα στον ψηφιακό κόσμο και πρέπει να γίνεται κατανοητή τόσο υπό το πρίσμα προστατευτικών μηχανισμών όσο και επιθετικών εργαλείων για την ενίσχυση της ψηφιακής καινοτομίας, συμπεριλαμβανομένης και της συνεργασίας με εταιρείες εκτός ΕΕ.⁹⁹

Η έννοια της ψηφιακής κυριαρχίας διατυπώνει την κυριαρχία επί ψηφιακών αγαθών, τα οποία μπορεί να είναι υλικά ή άυλα στοιχεία και, επομένως, να «εντοπίζονται» δυνητικά σε έναν χώρο που υπερβαίνει τα φυσικά σύνορα.¹⁰⁰ Η έννοια της ψηφιακής κυριαρχίας δηλώνει περαιτέρω την κυριαρχία στο πλαίσιο ενός ψηφιακού οικοσυστήματος και δεν εξαντλείται στην άσκηση κανονιστικής εξουσίας, αλλά συνδέεται με την ικανότητα μιας χώρας ή μιας ένωσης, όπως η ΕΕ, να θέτει τους κανόνες και τις προδιαγραφές της, καθώς και με την ικανότητα να επηρεάζει και τους άλλους. Συνδέεται με την ικανότητα να αποφασίζει και να δρα αυτόνομα σε σχέση με τις ψηφιακές διαστάσεις του μέλλοντος της οικονομίας, της κοινωνίας και της δημοκρατίας.

⁹⁸ European Commission (2020). *White paper on artificial intelligence: A European approach to excellence and trust* (COM(2020) 65 final), p. 2. https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

⁹⁹ European Parliamentary Research Service (2020). *Digital sovereignty for Europe* (EPRS Ideas Paper No. 651992). European Parliament. [https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI\(2020\)651992](https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2020)651992)

¹⁰⁰ Celeste, E. (2021). Digital sovereignty in the EU: challenges and future perspectives. In F. Fabbrini, E. Celeste & J. Quinn, *Data protection beyond borders: Transatlantic perspectives on extraterritoriality and sovereignty* (pp. 211, 217-218). Hart.

Στο πλαίσιο αυτό η ψηφιακή κυριαρχία θεωρείται ως μία δημόσια σφαίρα που είναι προσανατολισμένη σε επιλογές και αξίες, ρυθμισμένη και, συνεπώς, έλλογη και ασφαλής για τους πολίτες της και τα δικαιώματά τους.¹⁰¹

Αξίζει να σημειωθεί ότι την 3η Ιουνίου 2026 η Ευρωπαϊκή Επιτροπή υιοθέτησε μια φιλόδοξη δέσμη μέτρων με στόχο την ενίσχυση της ψηφιακής αυτονομίας της ΕΕ, καθώς, όπως σημείωσε η Ursula von der Laien, η τεχνολογική κυριαρχία της Ευρώπης αποτελεί βασική προϋπόθεση για την ενίσχυση της ανταγωνιστικότητας, της ανθεκτικότητας και της στρατηγικής της αυτονομίας σε ένα ταχέως μεταβαλλόμενο ψηφιακό περιβάλλον. Κατά την Ευρωπαϊκή Επιτροπή, η ανάπτυξη της πολιτικής της τεχνολογικής κυριαρχίας θα επιτρέψει στην ΕΕ να διασφαλίσει τη στρατηγική της αυτονομία σε βασικές ψηφιακές τεχνολογίες, να υποστηρίξει ανοιχτές και δίκαιες ψηφιακές αγορές, να προστατεύσει τα δικαιώματα των πολιτών και τις δημοκρατικές διαδικασίες στον διαδικτυακό χώρο, να ενισχύσει την καινοτομία και να διασφαλίσει τη μακροπρόθεσμη τεχνολογική πρωτοπορία.¹⁰²

¹⁰¹ Αναλυτικότερα βλ. Μήτρου, Λ. (2024). Το Σύνταγμα και οι Νέες Τεχνολογίες. Στο *Η Θεωρία της Συνταγματικής Πράξης – Μελέτες διαλόγου με το έργο του Ευάγγελου Βενιζέλου* (σσ. 107-130, σ. 129 επ.). Επίσης, τις ενδιαφέρουσες σκέψεις του Στάγκου, Π. (2004). Επιτεύγματα και όρια μιας ευρωπαϊκής ψηφιακής κυριαρχίας. Στο Ν. Τολιόπουλος (Επιμ.), *Η Ιδιωτικότητα στην ψηφιακή εποχή* (σσ. 236-276). Νομική Βιβλιοθήκη.

¹⁰² European Commission (n.d.). *Strengthening Europe's Tech Sovereignty*. <https://digital-strategy.ec.europa.eu/en/policies/eu-tech-sovereignty>

3.5. Digital Omnibus: Απλοποίηση ή απορρύθμιση;

Η προσπάθεια να εξισορροπηθούν η υποστήριξη της καινοτομίας και η ρυθμιστική οριοθέτηση των συστημάτων Τεχνητής Νοημοσύνης, όπως αποτυπώθηκε στην Πράξη για την Τεχνητή Νοημοσύνη, συναντά σημαντικές αντιδράσεις. Η ταχεία διεύρυνση και πύκνωση του ψηφιακού κανονιστικού πλέγματος, που περιλαμβάνει από την προστασία προσωπικών δεδομένων (GDPR) και τη ρύθμιση των ψηφιακών αγορών και πλατφορμών (DMA, DSA), έως τη συνδεσιμότητα, την κυβερνοασφάλεια, την επιβολή του νόμου και την ΤΝ, αντιμετωπίστηκε ακόμη πιο κριτικά ύστερα από τη δημοσίευση της έκθεσης του Mario Draghi (2024).¹⁰³ Η έκθεση για το «Μέλλον της Ευρωπαϊκής Ανταγωνιστικότητας» συνέδεε ευθέως το ζήτημα των κανόνων και της κανονιστικής συμμόρφωσης με τις απαιτήσεις για ενίσχυση της ανταγωνιστικότητας της ευρωπαϊκής οικονομίας. Υπό την πίεση πολιτικών και επιχειρηματικών παραγόντων (πρωτίστως εκτός ΕΕ), αλλά και προς ελάφρυνση του «βάρους της συμμόρφωσης» (ή και «κόπωσης της συμμόρφωσης»), η Ευρωπαϊκή Επιτροπή εισηγήθηκε πρόσφατα αλλαγές στην AI Act, στο νομοθέτημα που φιλοδοξούσε «να ρυθμίσει πώς θα λειτουργεί μια τεχνολογία».

Οι τροποποιήσεις αυτές εντάσσονται στο ψηφιακό πακέτο (digital package), το οποίο προβλήθηκε σε συνδυασμό με τους στόχους της απλούστευσης, της εξοικονόμησης (πόρων), της καινοτομίας. Ιδίως της ψηφιακής καινοτομίας, που, με τη σειρά της, αναμένεται να ενισχύσει την παραγωγικότητα και την ανταγωνιστικότητα της ευρωπαϊκής καινοτομίας. Η διευκόλυνση της συμμόρφωσης με ένα αναμορφούμενο κανονιστικό πλαίσιο σε καίρια πεδία συνιστά προϋπόθεση και ταυτόχρονα σκοπό της νέας κανονιστικής πρωτοβουλίας της Ευρωπαϊκής Επιτροπής.

Οι προτάσεις χαιρέτιστηκαν από αρκετούς ως άρση κανονιστικών εμποδίων και εξορθολογισμός. Οι «παρεμβάσεις» ως προς την έναρξη εφαρμογής της AI Act φαίνεται να «διορθώνουν» την αρχική αισιοδοξία του ενωσιακού

¹⁰³ Draghi, M. (2024). *The future of European competitiveness*. European Commission. https://commission.europa.eu/document/download/97e481fd-2dc3-412d-be4c-f152a8232961_en?filename=The%20future%20of%20European%20competitiveness%20%20A%20competitiveness%20strategy%20for%20Europe.pdf

νομοθέτη ως προς την ετοιμότητα εφαρμογής ενός τέτοιου νομοθετήματος. Από την άλλη πλευρά, οι προτάσεις αυτές χαρακτηρίστηκαν ως απόπειρα απορρύθμισης, απομάκρυνση από τις αξίες της Ένωσης και (αυτο)υπονόμευση της κανονιστικής ψηφιακής κυριαρχίας της. Την προβληματική αυτή εκθέτει αναλυτικότερα ο Μιχάλης Κρητικός στο κείμενό του αναφορικά με τη ρύθμιση ως άξονα μιας κοινωνικά βιώσιμης οικονομίας.

Είναι πράγματι αναγκαία αυτή η ευρεία νομοθετική πρωτοβουλία; Διασφαλίζει το εργαλείο του omnibus¹⁰⁴ αποφάσεις που εγγυώνται την ισόρροπη ανάπτυξη, αλλά και τον σεβασμό των αξιών της ΕΕ; Πρόκειται για μια διορθωτική κίνηση¹⁰⁵ που θα ενισχύσει την καινοτομία χωρίς να αγνοεί τον ανθρωποκεντρικό χαρακτήρα της προσέγγισης της ΕΕ; Ή αυτός ο εξορθολογισμός και η απλούστευση των νομικών απαιτήσεων αλλάζουν το πνεύμα θεμελιωδών κανονιστικών επιλογών ή και τον πυρήνα ατομικών δικαιωμάτων; Είναι το δίλημμα «καινοτομία ή ρύθμιση» ουσιαστικό και υπαρκτό ή ένα ψευδοδίλημμα;

Πέραν των αμφισβητήσεων του κατά πόσον η –κατανοητή από την άποψη της αποτελεσματικότητας– κανονιστική φιλοσοφία και μέθοδος omnibus είναι συμβατή με θεμελιώδεις αρχές του Ενωσιακού Δικαίου,¹⁰⁶ ένα μείζονος σημασίας ζήτημα αφορά την έλλειψη τεκμηρίωσης των κανονιστικών επεμβάσεων. Η οριζόντια φύση και η ταχύτητα με την οποία προωθούνται οι εν λόγω μεταρρυθμίσεις έχουν ως αποτέλεσμα τη γενικευμένη υποδήλωση ή τον στιγματισμό υφιστάμενων κανονιστικών υποχρεώσεων συμμόρφωσης ως δυσανάλογων εμποδίων για την καινοτομία, την τεχνολογική ανάπτυξη και την ανταγωνιστικότητα, χωρίς επαρκή τεκμηρίωση που να δικαιολογεί μια τόσο εκτεταμένη αναθεώρηση θεμελιωδών ρυθμιστικών εγγυήσεων.¹⁰⁷ Η διενέργεια εκτενών εκτιμήσεων των επιπτώσεων, στο πλαίσιο των οποίων αξιολογούνται οι δυνητικές οικονομικές, κοινωνικές και περιβαλλοντικές

¹⁰⁴ Ένας προτεινόμενος Κανονισμός ή μία προτεινόμενη Οδηγία τύπου omnibus συνιστά ενιαίο νομοθετικό εργαλείο, μέσω του οποίου επιχειρείται η ταυτόχρονη τροποποίηση πολλαπλών υφιστάμενων πράξεων του Ενωσιακού Δικαίου, οι οποίες εκτείνονται σε διαφορετικούς τομείς δημόσιας πολιτικής. Ήδη από το 2024, τα πακέτα omnibus εντάσσονταν σε έναν ευρύτερο λόγο περί ευρωπαϊκής ανταγωνιστικότητας, τόσο στην Έκθεση Letta όσο και στις πολιτικές Κατευθυντήριες Γραμμές της προέδρου της Επιτροπής, Ursula von der Leyen, για την περίοδο 2024-2029, καθώς και στην έκθεση Draghi.

¹⁰⁵ Αξίζει να αναφέρουμε ότι ένα σπονδυλωτό άρθρο δέκα συγγραφών που πρόκειται να δημοσιευτεί στην τελική μορφή του στο *Computer Law and Security Review* (2026) και πραγματεύεται τη συζήτηση για αλλαγές στον ΓΚΠΔ φέρει τον εύγλωττο τίτλο: If it ain't broke, don't fix it? Ten improvements for the upcoming tenth anniversary of the General Data Protection Regulation, <https://dial.uclouvain.be/pr/boreal/object/boreal:303765>

¹⁰⁶ Όπως η αρχή της αναλογικότητας κατά το άρθρο 5 παράγραφος 4 της Συνθήκης για την Ευρωπαϊκή Ένωση (ΣΕΕ), καθώς και οι αρχές της τεκμηριωμένης, διαφανούς και συμμετοχικής νομοθέτησης, όπως κατοχυρώνονται στο άρθρο 11 της ΣΕΕ και στο άρθρο 296 της Συνθήκης για τη Λειτουργία της Ευρωπαϊκής Ένωσης (ΣΛΕΕ). Βλ. Baldon Avocats (2025). *Potential legal grounds challenging the validity of the Omnibus Directive under EU law*. ClientEarth. https://www.clientearth.org/media/yh2dnt0m/20250623_legal_analysis_omnibus-challenges.pdf

¹⁰⁷ Alemanno, A. (2025). *The legality of omnibus legislation under EU law: A preliminary analysis of Omnibus I simplification directive of CSRD and CSDD and its legal consequences on the EU legal order* [HEC Paris Research Paper No. LAW-2025-1588]. <https://ssrn.com/abstract=5727822>

συνέπειες συναφών πρωτοβουλιών, καθώς και η αναγκαιότητα, η αναλογικότητα, οι επιπτώσεις των προτεινόμενων παρεμβάσεων, αλλά και τα αναμενόμενα οφέλη, δεν κατέστη –σύμφωνα με την Ευρωπαϊκή Επιτροπή– εφικτή, καθώς θα απαιτούσε υπερβολικά μεγάλο χρονικό διάστημα, ενώ η ανάγκη για απορρύθμιση ήταν επείγουσα.¹⁰⁸

Υπό τις συνθήκες του διεθνούς γεωπολιτικού, οικονομικού και τεχνολογικού ανταγωνισμού είναι αναμενόμενες και κατανοητές οι πρωτοβουλίες ενίσχυσης της οικονομίας με έμφαση στην ψηφιακή διάσταση.

Η αντιμετώπιση των ρυθμιστικών ορίων ως ανάχωμα στην καινοτομία παραβλέπει ορισμένες βασικές παραμέτρους:

α) ότι η άσκηση της συμμόρφωσης συνιστά μια ευκαιρία ορθολογικής και ποιοτικής διαχείρισης και διακυβέρνησης της Τεχνητής Νοημοσύνης, νοούμενης ως μιας πρακτικής που αποσκοπεί στη μεγιστοποίηση των θετικών επιπτώσεών της, αλλά και στην ελαχιστοποίηση των πιθανών κινδύνων ή βλαβών,

β) ότι, παρά τους φρενήρεις ρυθμούς διείσδυσης των εφαρμογών ΤΝ στην καθημερινότητα σε βαθμό και με τρόπο που οδηγεί σε εθισμό,¹⁰⁹ η εμπιστοσύνη ως προς τη σύννομη και θεμιτή χρήση της συνιστά όρο αποδοχής, αλλά και ανταγωνιστικό πλεονέκτημα για όσους επενδύουν στην ηθική ΤΝ,

γ) ότι η ανάπτυξη που δεν υπόκειται σε κανόνες και δεν είναι ισόρροπη, συνεκτική και βιώσιμη κινδυνεύει να αποδειχθεί, ήδη μεσοπρόθεσμα, βραχύβια και εύθραυστη.

Εν προκειμένω, είναι σημαντικό να προσεγγίσουμε ολιστικά και τον ρόλο του κράτους (και της Ένωσης). Μια βασική παράμετρος των πολιτικών για την ΤΝ που έχουν υιοθετηθεί συνδέεται άμεσα με τον ρόλο του κράτους στην αλγοριθμική εποχή και με την αλληλεπίδρασή του με μη κρατικούς φορείς και δρώντες. Αυτοί οι διαφορετικοί ρόλοι συνδέονται στενά με τις διαφορετικές ρυθμιστικές προσεγγίσεις και τους αντίστοιχους στόχους τους. Το κράτος συχνά παρουσιάζεται ως «διευκολυντής», που θέτει τα θεμέλια για την καινοτομία και παρέχει στα ενδιαφερόμενα μέρη τα μέσα και τους πόρους, ως «προωθητής», που λειτουργεί ως υποστηρικτής και εκφραστής της αλλαγής, ως «διαμεσολαβητής» ανάμεσα σε διαφορετικές απόψεις και στάσεις σχετικά με τα οφέλη και τους κινδύνους της ΤΝ, καθώς και ως «διαιτητής» ή «διαπραγματευτής» μεταξύ διαφορετικών κοινωνικών και

¹⁰⁸ Mariniello, M. (2025). *Efficiency and distribution in the European Union's digital deregulation push* [Bruegel Policy Brief, No. 31/2025]. Bruegel <https://www.bruegel.org/policy-brief/efficiency-and-distribution-european-unions-digital-deregulation-push>

¹⁰⁹ Δεν μπορεί να παραβλέψει κανείς την έντονη δημόσια συζήτηση για την ευθύνη των πλατφορμών αναφορικά με τον εθιστικό σχεδιασμό και τις ευρείες και σοβαρές επιπτώσεις του.

πολιτικών θέσεων των εμπλεκόμενων φορέων. Ωστόσο, το κράτος πρέπει ταυτόχρονα να είναι «εγγυητής», από τον οποίο αναμένεται να διασφαλίζει την προστασία απέναντι σε κινδύνους που αφορούν την ασφάλεια, την προστασία, τα θεμελιώδη δικαιώματα και το κράτος δικαίου.¹¹⁰

¹¹⁰ Mitrou, L. (2023). Regulation vs innovation? A comparative analysis of EU and UK policy responses on Artificial Intelligence. In E. Celeste, R. á Costello, E. Harbinja & N. Stathoulis (Eds.), *Data Protection and Digital Sovereignty Post-Brexit* (pp. 125-146). Hart.

3.6. Η ρύθμιση ως άξονας μιας κοινωνικά βιώσιμης καινοτομίας

Σε έναν κόσμο που διαμορφώνεται από τον τεχνολογικό ανταγωνισμό, η ρύθμιση –ιδιαίτερα κανονιστικά πλαίσια, όπως η Πράξη για την Τεχνητή Νοημοσύνη της Ευρωπαϊκής Ένωσης– θεωρείται εμπόδιο στην καινοτομία και ανασχετικός παράγοντας ακριβώς τη στιγμή που οι Ηνωμένες Πολιτείες και η Κίνα επιταχύνουν τις επενδύσεις και την ανάπτυξη. Το ενωσιακό μοντέλο ταυτίζεται με γραφειοκρατικούς φραγμούς και παρουσιάζεται ως ανασταλτικός παράγοντας για την ανταγωνιστικότητα. Αυτές οι ανησυχίες επικεντρώνονται στο κανονιστικό κόστος που συνεπάγεται η συμμόρφωση με νομοθετικές διατάξεις και στη «δυσκινησία» την οποία δημιουργεί η νέα διοικητική και οργανωτική αρχιτεκτονική, που διαμορφώνεται με τη σταδιακή εφαρμογή των σχετικών διατάξεων.

Ο εντεινόμενος τεχνολογικός ανταγωνισμός επαναφέρει διαρκώς στη δημόσια συζήτηση το εάν τα νέα κανονιστικά και ηθικά πρότυπα θα επηρεάσουν την ικανότητα της ευρωπαϊκής βιομηχανίας, περιλαμβανομένων των ευρωπαϊκών νεοφυών επιχειρήσεων, να παράγει «καινοτομία» στο επίπεδο ανάπτυξης βασικών μοντέλων, κατασκευής ημιαγωγών και ψηφιακών πλατφορμών. Το γεγονός ότι ο ανταγωνισμός αυτός έχει αποκτήσει γεωπολιτικά χαρακτηριστικά και φαίνεται να καθορίζει οικονομικές, ακόμα και στρατιωτικές, ισορροπίες έχει ενισχύσει τις φωνές εκείνες που ζητούν απλοποίηση και χαλάρωση των σχετικών νομοθετικών διατάξεων.

Είναι βάσιμες αυτές οι ανησυχίες; Το τεχνολογικό τοπίο που διαμορφώνεται αποκαλύπτει, επίσης, μια διαφορετική πραγματικότητα. Πρώτον, η ΤΝ εγκυμονεί συστημικούς κινδύνους, οι οποίοι, εάν παραμείνουν αρρύθμιστοι, θα μπορούσαν να υπονομεύσουν τη δημοκρατία, την κοινωνική συνοχή και τα θεμελιώδη δικαιώματα των πολιτών. Για παράδειγμα, οι εκστρατείες παραπληροφόρησης που τροφοδοτούνται από την Παραγωγική ΤΝ απειλούν τις ίδιες τις εκλογικές διαδικασίες, ενώ οι μεροληπτικοί αλγόριθμοι υπονομεύουν την κοινωνική συνοχή και μπορούν να οδηγήσουν σε τρομερές αδικίες και κοινωνικά δράματα. Οι τεχνολογίες επιτήρησης διαβρώνουν ήδη τις πολιτικές ελευθερίες σε πολλές χώρες, ενώ η διείσδυση αυτής της τεχνολογίας σε «ευαίσθητους» χώρους, όπως στον ιατρικό, στον εκπαιδευτικό, αλλά και στην αστυνόμευση, δημιουργεί πρωτόγνωρες προκλήσεις που δεν μπορούν να αντιμετωπιστούν με μεθόδους αυτορρύθμισης ή γενικόλογες συστάσεις σε επίπεδο soft law.

Επιπρόσθετα, θα ήταν άστοχο να αποδοθεί το χάσμα της Ευρώπης στην ψηφιακή καινοτομία με τις Ηνωμένες Πολιτείες και την Κίνα –το οποίο προηγείται των Κανονισμών όπως ο GDPR και η Πράξη για την Τεχνητή Νοημοσύνη– αποκλειστικά σε υποτιθέμενες υπερβολικές ρυθμίσεις της ΕΕ. Οι βαθύτερες αιτίες αυτού του κενού είναι πολυπαραγοντικές και σχετίζονται με τη γλωσσική ποικιλομορφία της ΕΕ, την έλλειψη ολοκληρωμένης ψηφιακής αγοράς, την απουσία συνεκτικής ευρωπαϊκής βιομηχανικής πολιτικής, την έλλειψη σημαντικών δημόσιων και ιδιωτικών επενδύσεων και ψηφιακών υποδομών, τις δυσκολίες προσέλκυσης και διατήρησης ταλέντων και την απουσία μιας ενιαίας κουλτούρας καινοτομίας.

Εάν σχεδιαστούν κατά τρόπο συμπεριληπτικό και με σεβασμό στην αρχή της αναλογικότητας, οι ρυθμιστικές παρεμβάσεις οδηγούν, άλλωστε, σε βάθος χρόνου σε στρατηγικά πλεονεκτήματα: Δημιουργούν εμπιστοσύνη στις αγορές, παρέχουν ασφάλεια δικαίου σε ένα διαρκώς μεταβαλλόμενο τεχνολογικό πεδίο, προσφέρουν μια πολύτιμη εργαλειοθήκη στους δημοκρατικούς θεσμούς και ενισχύουν τον διακριτό χαρακτήρα της Ευρώπης ως τεχνολογικό παράγοντα. Στην πραγματικότητα, η ρύθμιση μπορεί να δημιουργήσει προβλεψιμότητα και εμπιστοσύνη και ένα σαφές πλαίσιο κανόνων του τεχνολογικού παιχνιδιού.

Επιπλέον, τα συστήματα ΑΙ λειτουργούν σε ευαίσθητους τομείς, όπως η υγειονομική περίθαλψη, η εκπαίδευση, το πεδίο της απασχόλησης και η επιβολή του νόμου. Σε περίπτωση λάθους, τυχόν κοινωνικές αντιδράσεις μπορεί να ανατρέψουν την ίδια την ανάπτυξη της τεχνολογίας. Η αντίδραση κατά των τεχνολογιών αναγνώρισης προσώπου σε πολλές πόλεις παγκοσμίως δείχνει πόσο εύθραυστη μπορεί να είναι η αποδοχή. Από τους τοξικούς αλγόριθμους των μέσων κοινωνικής δικτύωσης έως τα προκατειλημμένα προγνωστικά συστήματα αστυνόμευσης, η κακώς ρυθμιζόμενη ΤΝ μπορεί να διαβρώσει την εμπιστοσύνη του κοινού και να προκαλέσει βαριές αντιδράσεις αργότερα.

Επομένως, η ρυθμιστική έμφαση που δίνει η Ευρώπη στη διαφάνεια, στη λογοδοσία και τα θεμελιώδη δικαιώματα ενισχύει την εμπιστοσύνη του κοινού και αποσκοπεί στο να διασφαλίσει ότι η καινοτομία θα ανθίζει χωρίς επιβλαβείς παρενέργειες – μια στάση που ενισχύει, δεν αποδυναμώνει, το ανταγωνιστικό της τοπίο. Είναι σημαντικό να σημειωθεί ότι όλες οι σχετικές έρευνες δείχνουν σταθερά ότι οι Ευρωπαίοι πολίτες είναι επιφυλακτικοί σχετικά με τις εφαρμογές ΤΝ που θεωρούνται παρεμβατικές ή μεροληπτικές. Μάλιστα, 80% των πολιτών της ΕΕ λέει ότι η ανάπτυξη της ΤΝ πρέπει να ρυθμίζεται προσεκτικά, ακόμη και αν επιβραδύνει την εξέλιξή της.¹¹¹

¹¹¹ Βλ. European Commission (2025). *Special Eurobarometer 554: Artificial Intelligence and the future of work (European Union)* [Dataset]. https://data.europa.eu/data/datasets/s3222_101_4_sp554_eng και Eurobarometer Survey 3222, & European Commission (2026). *Digital decade 2026 - Special Eurobarometer*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2759/4731876>

Η Πράξη για την Τεχνητή Νοημοσύνη ανταποκρίνεται άμεσα σε αυτές τις ανησυχίες. Για τους παροικούντες την τεχνολογική Ιερουσαλήμ, η εμπιστοσύνη δεν είναι απλώς ηθικό κεφάλαιο, είναι οικονομικό και πολιτικό κεφάλαιο στρατηγικού χαρακτήρα. Όταν οι πολίτες πιστεύουν ότι τα συστήματα είναι δίκαια και ασφαλή, η υιοθέτηση αυξάνεται. Η υψηλότερη υιοθέτηση παράγει περισσότερα δεδομένα, τα οποία βελτιώνουν τα μοντέλα και τροφοδοτούν την ανάπτυξη. Υπό αυτήν την έννοια, η ηθική δεν αντιμετωπίζεται πλέον ως εξωτερικός περιορισμός στην καινοτομία. Είναι απαραίτητη προϋπόθεση για την υπεύθυνη κλιμάκωσή της.

Σε έναν κόσμο όπου η εμπιστοσύνη του κοινού στην ΤΝ είναι χαμηλή –με έντονες ανησυχίες σχετικά με την παρακολούθηση, την προκατάληψη και την απώλεια ελέγχου– η έμφαση της Ευρώπης στην ηθική μπορεί να μετατραπεί σε ανταγωνιστικό πλεονέκτημα. Οι πολίτες είναι γενικά πιο πρόθυμοι να υιοθετήσουν τεχνολογίες που εμπιστεύονται. Οι επιχειρήσεις B2B προτιμούν συνεργάτες των οποίων τα συστήματα συμμορφώνονται με ισχυρά νομικά πρότυπα. Οι κυβερνήσεις και οι δημόσιοι οργανισμοί επιλέγουν με μεγαλύτερη ασφάλεια λύσεις ΤΝ που τηρούν τις ηθικές και νομικές αρχές. Η εμπιστοσύνη, μόλις χτιστεί, γίνεται πολλαπλασιαστής της αγοράς.¹¹²

Άλλωστε, η ρύθμιση δεν αποτελεί, από μόνη της, εμπόδιο στην καινοτομία. Αν σχεδιαστεί με βάση ενδεδειγμένες μελέτες προκαταρκτικών κοινωνικών και περιβαλλοντικών επιπτώσεων, μπορεί να αποτελέσει εφαλτήριο για μια κοινωνικά επιθυμητή καινοτομία, δηλαδή την καινοτομία που ωφελεί όχι μόνο οικονομικούς σκοπούς, αλλά πρωτίστως την ανθρωπινή και την πλανητική ευημερία. Η ρύθμιση, παρέχοντας τα πρότυπα και τα κριτήρια που καθοδηγούν την καινοτομία, είναι ένα μέσο για την επιχειρησιακή εφαρμογή της κοινωνικά επιθυμητής καινοτομίας, εάν πρόκειται για «υπεύθυνη καινοτομία», «πράσινη καινοτομία» ή «καινοτομία χωρίς αποκλεισμούς».

Η ρύθμιση της αγοράς, σε αυτό το πλαίσιο, μπορεί να διαδραματίσει βασικό ρόλο στην παροχή κινήτρων για την τήρηση ηθικών απαιτήσεων και στη διασφάλιση της ενσωμάτωσής τους σε ψηφιακά αγαθά και υπηρεσίες, παρέχοντας ποιοτικό ανταγωνιστικό πλεονέκτημα έναντι προϊόντων που συμμορφώνονται αποκλειστικά με οικονομικές απαιτήσεις. Υπό αυτήν την έννοια, τα ρυθμιστικά πρότυπα της ΕΕ έχουν την προοπτική να αναδειχθούν σε πλεονέκτημα της αγοράς και σε μέσο διασφάλισης ότι η καινοτομία λαμβάνει υπ' όψιν και υπηρετεί το δημόσιο συμφέρον.

Η προβληματική λογική που αντιπαράθετεί την καινοτομία έναντι της ρύθμισης αγνοεί επιπλέον τον *κεντρικό ρόλο που διαδραματίζει η νομοθεσία στην*

¹¹² Βλ. Choung, H., David, P., & Ross, A. (2023). Trust and ethics in AI. *AI & Society*, 38, 733-745. <https://doi.org/10.1007/s00146-022-01473-4> και Reinhardt, K. (2023). Trust and trustworthiness in AI ethics. *AI Ethics*, 3, 735-744. <https://doi.org/10.1007/s43681-022-00200-5>

προώθηση της δίκαιης κατανομής των πιθανών οφελών και κινδύνων των ψηφιακών τεχνολογιών και στη διασφάλιση ότι τα οφέλη της καινοτομίας διαχέονται στην κοινωνία. Η νομολογία του Δικαστηρίου της ΕΕ επιβεβαιώνει ότι η ρύθμιση δεν είναι αντίθετη προς την καινοτομία. Επιπλέον, έχει επανειλημμένα υποστηρίξει ότι η τεχνολογική ανάπτυξη πρέπει να παραμείνει υποταγμένη στην προστασία της ανθρώπινης αξιοπρέπειας, της αυτονομίας και της ισότητας. Η ρύθμιση θα πρέπει να νοείται ως ένα μέσο για την τήρηση των ευρωπαϊκών δεσμεύσεων σε αξίες και αρχές, όπως η ισότητα, η κοινωνική δικαιοσύνη, το κράτος δικαίου, η ανθρώπινη αξιοπρέπεια και τα θεμελιώδη δικαιώματα. Η καινοτομία και η ρύθμιση δεν είναι αντίθετα. Είναι αλληλεξαρτώμενες δυνάμεις που διαμορφώνουν την τροχιά της αναδυόμενης τεχνολογίας.

Το κράτος και οι υπερεθνικές οντότητες, όπως η Ευρωπαϊκή Ένωση, έχουν κρίσιμους ρόλους στη διαμόρφωση πλαισίων διακυβέρνησης που προστατεύουν τους πολίτες, ενισχύουν την εμπιστοσύνη και δημιουργούν περιβάλλοντα όπου η καινοτομία μπορεί να ευδοκιμήσει με υπευθυνότητα. Η ΤΝ αποτελεί παράδειγμα τόσο της υπόσχεσης όσο και της πολυπλοκότητας αυτής της αλληλεπίδρασης. Μέσω ρυθμίσεων με επίγνωση των κινδύνων, με βάση τα δικαιώματα –σε συνδυασμό με στρατηγική υποστήριξη για έρευνα, δεξιότητες και συνεργασία–, οι κυβερνήσεις μπορούν να κατευθύνουν την ΤΝ προς αποτελέσματα που ενισχύουν την ανθρώπινη άνθηση, ενώ παράλληλα προστατεύονται από βλάβες.

Η καινοτομία δεν θα πρέπει να μετράται αποκλειστικά με τον αριθμό των νεοφυών επιχειρήσεων ή την κλίμακα της χρηματοδότησης κεφαλαίων επιχειρηματικού κινδύνου. Πρέπει, επίσης, να αξιολογηθεί ως προς την ανθεκτικότητα και τον κοινωνικό αντίκτυπο. Τεχνολογικές καινοτομίες που υπονομεύουν το κοινωνικό κεφάλαιο, διαβρώνουν τη θεσμική εμπιστοσύνη και εντείνουν την κοινωνική αστάθεια δύνανται να λειτουργήσουν ανασταλτικά ως προς τη βιώσιμη οικονομική ανάπτυξη. Για παράδειγμα, η εξάπλωση του Bitcoin και άλλων crypto, που αμφισβητούν τα παραδοσιακά χρηματοπιστωτικά συστήματα, έχει δημιουργήσει ρυθμιστική αβεβαιότητα, σε πολλές περιπτώσεις οικονομικές φούσκες, αλλά και αστάθεια σε αγορές. Με την ενσωμάτωση της ηθικής στον σχεδιασμό, η Ευρώπη επιδιώκει να διασφαλίσει ότι η ανάπτυξη της ΤΝ παραμένει συμβατή με τις δημοκρατικές αξίες. Αυτός ο μακροπρόθεσμος προσανατολισμός αντανάκλα μια ευρύτερη ευρωπαϊκή φιλοσοφία: ο οικονομικός δυναμισμός και η κοινωνική προστασία δεν αλληλοαποκλείονται. Ενισχύουν το ένα το άλλο.

Παρά τις εντεινόμενες γεωπολιτικές πιέσεις που αφορούν την εξάπλωση και τη διάχυση της ΤΝ, οι οποίες οξύνονται από τις επιθετικές επενδύσεις από τις Ηνωμένες Πολιτείες και την Κίνα, η Ευρώπη δεν πρέπει να μείνει άπραγη. Υπό το πρίσμα ενός εντεινόμενου παγκόσμιου ανταγωνισμού για ταλέντο και κεφάλαιο, η Ευρώπη πρέπει να αντιτάξει ένα γενναίο χρηματοδοτικό

πλαίσιο έρευνας και τη δημιουργία πανευρωπαϊκών νησίδων καινοτομίας που θα στοχεύουν στην ανάπτυξη εφαρμογών που θα είναι και εφαρμόσιμες και θα ανταποκρίνονται σε συγκεκριμένες κοινωνικές ανάγκες.

Ο γεωπολιτικός ανταγωνισμός συχνά ενθαρρύνει την απορρύθμιση στο όνομα της ταχύτητας. Αλλά μια κούρσα προς τα κάτω στη διακυβέρνηση της ΤΝ κινδυνεύει να αποσταθεροποιήσει τις δημοκρατικές κοινωνίες. Τα συστήματα κοινωνικής βαθμολόγησης, η μαζική επιτήρηση και η ανεξέλεγκτη αλγοριθμική λήψη αποφάσεων μπορεί να προσφέρουν βραχυπρόθεσμη αποτελεσματικότητα, αλλά υπονομεύουν τη μακροπρόθεσμη νομιμότητα. Επομένως, *η εγκατάλειψη της ηθικής διακυβέρνησης, επιδιώκοντας μια καινοτομία χωρίς φρένα, θα ήταν στρατηγικό λάθος και βασισμένη σε κοντόφθαλμες συγκρίσεις.*

Η Ευρώπη έχει επιλέξει διαφορετικό δρόμο ακολουθώντας το κανονιστικό μονοπάτι. Η ρύθμιση, όταν είναι «έξυπνη» και «προσαρμοστική», έχει πολλαπλά οφέλη: Παρέχει την απαιτούμενη ασφάλεια δικαίου, εμπνέει εμπιστοσύνη στην αγορά και αυξάνει τη διεθνή επιρροή της Ένωσης, ενισχύοντας το περίφημο «Brussels Effect», όπως αυτό έχει διαμορφωθεί ιστορικά στο πεδίο του Δικαίου του ανταγωνισμού, της προστασίας των προσωπικών δεδομένων και της προστασίας του περιβάλλοντος. Η προσέγγιση της Ευρώπης, που βασίζεται στο συγκεκριμένο κανονιστικό μοντέλο, θα μπορούσε να καθορίσει παγκόσμια πρότυπα και να προσελκύσει επενδύσεις άλλου τύπου. Αυτή η στάση ενισχύει την ήπια δύναμη της Ευρώπης. Οι χώρες στην Αφρική, στη Λατινική Αμερική και τη Νοτιοανατολική Ασία που παρατηρούν τον παγκόσμιο αντίκτυπο της ΤΝ διαμορφώνουν ήδη μοντέλα διακυβέρνησης που προστατεύουν τα δικαιώματα αντί να τα διαβρώνουν. Η ρυθμιστική ηγεσία της Ευρώπης την καθιστά σημαντικό σημείο αναφοράς για όσες χώρες αναζητούν έναν ισορροπημένο και κοινωνικά δίκαιο ψηφιακό μετασχηματισμό.

Ένα τέτοιο ρυθμιστικό πλαίσιο θα μπορούσε ταυτόχρονα να αποτελέσει από μόνο του και ισχυρό παράγοντα καινοτομίας, καθώς δύναται να οδηγήσει στη δημιουργία εφαρμογών που είναι εξ αρχής εύληπτες, κατανοητές και προστατεύουν την ιδιωτικότητα «από κατασκευής». Η Ιστορία δείχνει ότι οι «περιορισμοί» ή τα καλά καθορισμένα όρια συχνά οδηγούν σε μια νέου τύπου δημιουργικότητα. Στην ΤΝ, οι ηθικές και νομικές απαιτήσεις έχουν ήδη αποτελέσει το εφαλτήριο για τη διαμόρφωση νέων τεχνικών λύσεων και ενός διαφορετικού μοντέλου καινοτομίας, που επιτρέπουν για παράδειγμα στα μοντέλα ΑΙ να εκπαιδεύονται σε αποκεντρωμένα δεδομένα χωρίς να εκθέτουν ευαίσθητες πληροφορίες.

Τα ευρωπαϊκά πανεπιστήμια και οι νεοφυείς επιχειρήσεις βρίσκονται στην πρώτη γραμμή αυτών των εξελίξεων, ακριβώς επειδή το απόρρητο αντιμετωπίζεται ως απαίτηση σχεδιασμού και όχι ως εκ των υστέρων σκέψη.

Επιπλέον, τα συστήματα «υψηλού κινδύνου», σύμφωνα με το Δίκαιο της ΕΕ, πρέπει να παρέχουν διαφάνεια και λογοδοσία. Αυτό έχει ωθήσει την πρόοδο στα εργαλεία επεξήγησης που ερμηνεύουν τις εξόδους των νευρωνικών δικτύων, ανιχνεύουν τη μεροληψία και την απόδοση του μοντέλου ελέγχου. Αυτό που ξεκίνησε ως αναγκαιότητα συμμόρφωσης γίνεται παγκόσμιο ερευνητικό σύνολο.

Οι αξιολογήσεις συμμόρφωσης της Πράξης για την Τεχνητή Νοημοσύνη δημιουργούν ζήτηση για υπηρεσίες ελέγχου τρίτων, πλατφόρμες συμμόρφωσης και λογισμικό διακυβέρνησης. Ένας νέος τομέας –η διασφάλιση της ΤΝ– αναδύεται. Οι ευρωπαϊκές εταιρείες μπορούν να ηγηθούν αυτής της αγοράς, αξιοποιώντας την πρώιμη ρυθμιστική ευθυγράμμιση.

Ήδη έχουν κάνει την εμφάνισή τους στην αγορά τεχνολογικές εφαρμογές και μοντέλα που συμμορφώνονται με τα πρότυπα του «privacy by design» αλλά και του «ethics by design», και τεχνικές, όπως η «διαφορική ιδιωτικότητα» και η ομοσπονδιακή μάθηση, επεξηγήσιμες μέθοδοι ΤΝ για καλύτερη λογοδοσία, πλατφόρμες διακυβέρνησης ΑΙ που παρακολουθούν τη συμμόρφωση. Αυτές οι καινοτομίες δεν θα υπήρχαν χωρίς ηθική και κανονιστική πίεση.

Ταυτόχρονα, και καθώς τα τσιπ ΤΝ, οι υποδομές cloud και τα Μεγάλα Γλωσσικά Μοντέλα συγκεντρώνονται όλο και περισσότερο σε λίγους παγκόσμιους παίκτες, η κανονιστική σαφήνεια της Ευρώπης μπορεί να προσελκύσει εναλλακτικούς παρόχους που δίνουν προτεραιότητα και στη συμμόρφωση με συγκεκριμένα ηθικά πρότυπα. Για παράδειγμα, οι νεοφυείς επιχειρήσεις που αναπτύσσουν μοντέλα ΤΝ ανοικτού κώδικα μπορούν να προσφέρουν αξιόπιστες εναλλακτικές λύσεις σε ένα αδιαφανές τεχνολογικό τοπίο. Αυτά τα παραδείγματα απεικονίζουν μια ευρύτερη αρχή: Όταν οι αξίες διαμορφώνουν την τεχνολογική ανάπτυξη, παράγουν διαφοροποιημένη καινοτομία.

Με άλλα λόγια, η ρυθμιστική προσέγγιση της Ευρώπης *πρέπει να εκληφθεί όχι ως αμυντική, αλλά ως στρατηγική προσέγγιση ενός διαφορετικού μοντέλου καινοτομίας σε μια ασταθή τεχνολογική κούρσα εξοπλισμών που βασίζεται σε αξίες*. Η Ευρώπη δεν βρίσκεται σε έναν αγώνα δρόμου για να είναι η πιο γρήγορη – είναι σε έναν αγώνα δρόμου για να είναι η πιο αξιόπιστη, η πιο υπεύνη και, τελικά, η πιο ανθεκτική. Στην εποχή της ΤΝ, αυτές οι ιδιότητες δεν είναι πολυτέλειες – είναι προϋποθέσεις για μακροπρόθεσμη ηγεσία. Το ερώτημα δεν είναι αν η Ευρώπη μπορεί να καινοτομήσει βάσει Κανονισμών. Το ερώτημα είναι εάν η καινοτομία χωρίς ηθικά προστατευτικά κιγκλιδώματα είναι βιώσιμη – οικονομικά, κοινωνικά ή γεωπολιτικά. Σε έναν κόσμο όπου η ΤΝ διαμορφώνει ολοένα και περισσότερο τη γεωπολιτική, *αυτή μπορεί να αποδειχθεί η πιο αποφασιστική και επιδραστική καινοτομία και συνεισφορά της Ευρώπης στο παγκόσμιο τεχνολογικό γίγνεσθαι και να ενισχύσει τη στρατηγική της αυτονομία*.

[ΠΩΣ] ΜΠΟΡΕΙ Η ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΝΑ ΕΙΝΑΙ
ΗΘΙΚΗ ΚΑΙ ΣΥΝΝΟΜΗ;
Ιούλιος 2026

Τα θεμελιώδη δικαιώματα,
οι δημοκρατικές διαδικασίες
και η ρύθμιση της Τεχνητής
Νοημοσύνης



Λίλιαν Μήτρου

Οι επιπτώσεις της Τεχνητής Νοημοσύνης στην απόλαυση και άσκηση θεμελιωδών δικαιωμάτων είχαν επισημανθεί ήδη από την Ομάδα Εμπειρογνομόνων Υψηλού Επιπέδου για την Τεχνητή Νοημοσύνη στις Κατευθυντήριες Γραμμές Δεοντολογίας για Αξίопιστη Τεχνητή Νοημοσύνη του 2019. Η χρήση ΤΝ μπορεί να συνεπιφέρει κινδύνους για πολλά θεμελιώδη δικαιώματα και ελευθερίες. Οι απειλές και οι επεμβάσεις μπορεί να αφορούν την ισότητα και την προστασία από διακρίσεις, την προσωπική ελευθερία, την προστασία της προσωπικότητας, την προστασία προσωπικών δεδομένων, το δικαίωμα αποτελεσματικής προστασίας, την ελευθερία της έκφρασης, το δικαίωμα του συνέρχεσθαι και του συνεταιρίζεσθαι. Έχει, επίσης, επισημανθεί ότι η χρήση ΤΝ μπορεί να έχει σοβαρό αντίκτυπο στις δημοκρατικές διαδικασίες και κατ' επέκταση στη λειτουργία του δημοκρατικού πολιτεύματος. Εάν η χρήση ΤΝ συνιστά επέμβαση ή επίταση της επέμβασης στα ατομικά, στα πολιτικά και τα κοινωνικά δικαιώματα, από τα δικαιώματα αυτά απορρέουν και οι αρχές που πρέπει να διέπουν τις χρήσεις αυτές.

4.1. Η Τεχνητή Νοημοσύνη και τα θεμελιώδη δικαιώματα

Η αξιοποίηση συστημάτων Τεχνητής Νοημοσύνης έχει, χωρίς αμφιβολία, επιπτώσεις σε δικαιώματα που είναι κομβικής σημασίας στην εθνική και την ευρωπαϊκή δικαιοταξία. Σε πολλές περιπτώσεις μια χρήση μπορεί να επεμβαίνει σε περισσότερα δικαιώματα: Π.χ., μια απόφαση που ενέχει μεροληψία μπορεί να αφορά το δικαίωμα στην ισότητα και τη μη διάκριση, την προστασία των προσωπικών δεδομένων και το δικαίωμα σε αποτελεσματική προσφυγή. Αντίστοιχα, μια απαίτηση σύννομης και θεμιτής χρήσης, όπως το δικαίωμα στην επεξήγηση ή/και στην ανθρώπινη παρέμβαση, μπορεί να ικανοποιεί περισσότερα δικαιώματα, όπως η προστασία προσωπικών δεδομένων, το δικαίωμα σε αποτελεσματική προσφυγή-έννομη προστασία ή το δικαίωμα στη χρηστή διοίκηση.

Δεν είναι παράδοξο που η πρόσφατη έκθεση του Συμβουλίου της Ευρώπης¹¹³ έχει ως αφετηρία της ανάλυσης των επιπτώσεων στα δικαιώματα την ανθρώπινη αξιοπρέπεια (human dignity), που συνιστά τον πυρήνα του συστήματος προστασίας της Ευρωπαϊκής Σύμβασης Δικαιωμάτων του Ανθρώπου (ΕΣΔΑ), επισημαίνοντας ότι οι χρήσεις της ΤΝ μπορεί να αποανθρωπούν τα πρόσωπα (dehumanise), να τα υποβιβάζουν σε σημεία δεδομένων ή να υπονομεύουν την αυτονομία τους. Μια αξιοσημείωτη απαίτηση που απορρέει από τη Σύμβαση-Πλαίσιο για την ΤΝ (βλ. παρακάτω) είναι η αποφυγή του ανθρωπομορφισμού της ΤΝ κατά τρόπο που παραβιάζει την ανθρώπινη αξιοπρέπεια.¹¹⁴ Ζητήματα προσβολής της ανθρώπινης αξιοπρέπειας εγείρουν χρήσεις, όπως η «πρόβλεψη τέλεσης μιας αξιόποινης πράξης», με βάση αποκλειστικά την κατάρτιση προφίλ ενός φυσικού προσώπου ή την αξιολόγηση των γνωρισμάτων και των χαρακτηριστικών της προσωπικότητάς του, ή η κοινωνική βαθμολόγηση (social scoring), δηλαδή η αξιολόγηση φυσικών προσώπων ή ομάδων προσώπων με βάση την κοινωνική τους συμπεριφορά ή γνωστά, συναγόμενα ή προβλεπόμενα προσωπικά χαρακτηριστικά τους ή

¹¹³ Steering Committee for Human Rights (CDDH) (2026). *Human Rights and Artificial Intelligence: Handbook of the Steering Committee for Human Rights*. Council of Europe. <https://rm.coe.int/artificial-intelligence-handbook/48802b5bf2>

¹¹⁴ Council of Europe. (2024). *Framework convention on artificial intelligence, human rights, democracy and the rule of law: Explanatory report*. <https://rm.coe.int/framework-convention-on-artificial-intelligence-explanatory-report/1680afae3f>

χαρακτηριστικά της προσωπικότητάς τους.¹¹⁵ Οι σοβαρότεροι κίνδυνοι του αλγοριθμικού profiling εντοπίζονται ακριβώς στη δυνατότητα χρήσης του ως προβλεπτικού και προληπτικού εργαλείου. Η κατηγοριοποίηση, η ένταξη των προσώπων σε ομάδες με προ-προσδιορισμένα χαρακτηριστικά ή σε μορφότυπους συμπεριφοράς θέτουν ζητήματα ως προς τον –επιβαλλόμενο από το Σύνταγμα– σεβασμό της «αξίας του ανθρώπου», καθώς ένα πρόσωπο κινδυνεύει να αντιμετωπίζεται, στις διάφορες εκφάνσεις της ζωής του, όχι ως αυτόνομη οντότητα, αλλά ως φορέας ενός «μορφότυπου».

Η σύνδεση μεταξύ αξιοπρέπειας και προσωπικής αυτονομίας –του δικαιώματος των ατόμων να ορίζουν αυτόνομα τη ζωή τους, όπως αναγνωρίζεται στο άρθρο 8 της ΕΣΔΑ, στη νομολογία του Ευρωπαϊκού Δικαστηρίου Δικαιωμάτων του Ανθρώπου (ΕΔΔΑ), αλλά και στα άρθρα 2 παρ. 1, 5 και 9Α του Συντάγματος– είναι ιδιαίτερα προφανής στις περιπτώσεις κατά τις οποίες συστάσεις, αποφάσεις που υποστηρίζονται ή λαμβάνονται από συστήματα ΤΝ μπορεί είτε εκ σχεδιασμού είτε εκ του αποτελέσματος να επενεργούν στη συμπεριφορά ή/και στην ελευθερία λήψης αποφάσεων ενός ατόμου χωρίς επαρκείς δυνατότητες συμμετοχής, εξήγησης ή αμφισβήτησης και κατά τρόπο που περιορίζεται η αυτονομία του.¹¹⁶

Η χρήση ΤΝ μπορεί να έχει σοβαρές επιπτώσεις στην ιδιωτική ζωή και στο δικαίωμα προσωπικών δεδομένων. Ένα σύστημα ΤΝ μπορεί να επεξεργαστεί τεράστια ποσότητα δεδομένων, ενώ η επεξεργασία χαρακτηρίζεται από ακρίβεια, αλλά και ταχύτητα παραγωγής και επικαιροποίησης αποτελεσμάτων. Από ακατέργαστα, συχνά ασύνδετα μεταξύ τους, δεδομένα, τα συστήματα ΤΝ επιτρέπουν να αναδειχθούν πληροφορίες ή να παραχθούν νέες, να αναλυθούν μοτίβα, να εντοπιστούν συσχετίσεις και μοντέλα ή να συναχθούν συμπεράσματα και προγνώσεις. Από δεδομένα καθημερινά ή «αβλαβή» μπορούν να συναχθούν ή να παραχθούν ευαίσθητες πληροφορίες που μπορεί να αφορούν άγνωστα σε τρίτους (ή ακόμη και στο ίδιο το άτομο) προσωπικά χαρακτηριστικά ή καταστάσεις (σεξουαλικός προσανατολισμός, κατάσταση υγείας ή προδιάθεση ασθένειας, οικονομικές δυσχέρειες κ.ά.). Τα «δεδομένα εξόδου», τα συναγόμενα στοιχεία (inferences), μπορεί να χρησιμοποιηθούν για τη λήψη αποφάσεων, συχνά κρίσιμων για τη ζωή ενός προσώπου (όπως η απόλυσή του, η άρνηση ενός δανείου, η αντιμετώπισή του από δικαστικές Αρχές) ή για τη στοχευμένη πληροφόρηση ή/και διαφήμιση. Οι μοναδικές ιδιότητες της ΤΝ μπορούν να οδηγήσουν σε νέες μορφές επέμβασης στην πληροφοριακή ιδιωτικότητα, ενώ η απαίτηση για μαζικά δεδομένα που χρησιμοποιούνται και για την εκπαίδευση των συστημάτων και των Μεγάλων Γλωσσικών Μοντέλων καθιστά οξύτερη

¹¹⁵ Για τον λόγο αυτόν, άλλωστε, μια τέτοια χρήση συνιστά μη αποδεκτό κίνδυνο και απαγορεύεται από το άρθρο 5 της AI Act.

¹¹⁶ Είναι χαρακτηριστικό ότι στις μη αποδεκτές χρήσεις, σύμφωνα με την AI Act, συγκαταλέγεται η απαγόρευση τεχνικών που απευθύνονται στο υποσυνείδητο ενός προσώπου, υπερκεράζοντας το συνειδητό του, ή σκόπιμα χειριστικών ή παραπλανητικών τεχνικών, με στόχο ή με αποτέλεσμα την ουσιώδη στρέβλωση της συμπεριφοράς κατά τρόπο που προκαλεί ή εύλογα ενδέχεται να προκαλέσει στο εν λόγω πρόσωπο, σε άλλο πρόσωπο ή σε ομάδα προσώπων σημαντική βλάβη (φυσική, ψυχολογική, οικονομική, κοινωνική).

τη διείσδυση στην ιδιωτική ζωή. Η προσβολή της ιδιωτικότητας συνδέεται άρρηκτα με την αυτονομία. Έχοντας απολέσει τον έλεγχο επί των ιδίων πληροφοριών, τα πρόσωπα καθίστανται ευρέως και ευχερώς ικνηλάσιμα, ελέγξιμα και χειραγωγήσιμα, με επιπτώσεις στη συμπεριφορά, στις επιλογές τους και τα δικαιώματά τους, όπως η ανάπτυξη της προσωπικότητας ή η ελευθερία της έκφρασης.

Προφανής είναι ο κίνδυνος για το δικαίωμα της ισότητας και της μη διάκρισης, ένα δικαίωμα που προστατεύεται από το Σύνταγμά μας (άρθρο 4), αλλά και την ΕΣΔΑ (άρθρο 14), το άρθρο 2 της Συνθήκης της Ευρωπαϊκής Ένωσης και τον Χάρτη των Θεμελιωδών Δικαιωμάτων της ΕΕ (άρθρο 21).¹¹⁷ Ο κίνδυνος των διακρίσεων και της μεροληψίας μοιάζει να είναι εγγενής σε πολλές περιπτώσεις, καθώς σκοπός των αλγορίθμων μηχανικής μάθησης είναι να κατηγοριοποιούν, να ταξινομούν και να διαχωρίζουν. Η άμεση ή έμμεση διάκριση μέσω της χρήσης αλγορίθμων που περιλαμβάνουν μεγάλα δεδομένα θεωρείται μία από τις πιο πιεστικές προκλήσεις στη χρήση τεχνολογιών που βασίζονται στην ΤΝ.¹¹⁸ Η μεροληψία και οι διακρίσεις, συμπεριλαμβανομένων των διακρίσεων λόγω φύλου, στη λήψη αλγοριθμικών αποφάσεων, που υποστηρίζεται από δεδομένα, μπορούν να προκύψουν για διάφορους λόγους και σε πολλά επίπεδα των συστημάτων ΤΝ.¹¹⁹ Συχνά η διακριτική μεταχείριση μπορεί να οφείλεται στην ποιότητα, αλλά και στη μεροληψία που ενυπάρχει στα δεδομένα (μεροληψία δεδομένων). Οι κίνδυνοι αυτοί εντοπίζονται στον τομέα των εργασιακών σχέσεων, αλλά και στον τομέα της προγνωστικής αστυνόμευσης ή της επιβολής του νόμου, όταν η ανάλυση και αξιολόγηση βασίζεται σε μεροληπτικά «ιστορικά δεδομένα».

Η χρήση συστημάτων ΤΝ μπορεί να έχει επιπτώσεις και στο δικαίωμα πρόσβασης στη Δικαιοσύνη, στην έννομη προστασία και τη δίκαιη δίκη. Οι κίνδυνοι αυτοί συνδέονται με την έλλειψη διαφάνειας και επεξηγησιμότητας και τα ελλείμματα κατανόησης και ενημέρωσης δυσχεραίνουν τη δυνατότητα των προσώπων να εντοπίσουν ευθύνες και υπευθύνους και να διεκδικήσουν τα δικαιώματά τους, προσβάλλοντας την απόφαση που τα θίγει ή τα επηρεάζει αρνητικά. Υπό την έννοια αυτήν, επηρεάζεται και το δικαίωμα των προσώπων σε δίκαιη δίκη (άρθρο 6 της ΕΣΔΑ), καθώς τίθεται εν προκειμένω ζήτημα και αναφορικά με την αρχή της ισότητας των

¹¹⁷ Το άρθρο 21 του Χάρτη Θεμελιωδών Δικαιωμάτων του Ανθρώπου απαγορεύει κάθε διάκριση η οποία βασίζεται σε οποιοδήποτε λόγο, όπως το φύλο, η φυλή, το χρώμα, η εθνική ή κοινωνική καταγωγή, τα γενετικά χαρακτηριστικά, η γλώσσα, η θρησκεία ή οι πεποιθήσεις, οι πολιτικές ή άλλες γνώμες, η συμμετοχή σε εθνική μειονότητα, η περιουσία, η γέννηση, η αναπηρία, η ηλικία ή ο γενετήσιος προσανατολισμός.

¹¹⁸ Fundamental Rights Agency (2018). *#BigData: Discrimination in data-supported decision making* (p. 3). Publications Office of the European Union. https://fra.europa.eu/sites/default/files/fra_uploads/fra-2018-focus-big-data_en.pdf

¹¹⁹ Fundamental Rights Agency (2020). *Getting the future right: Artificial intelligence and fundamental rights*. Publications Office of the European Union. <https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-rights>

όπλων.¹²⁰ Αντίστοιχα, είναι τα ζητήματα που εγείρονται από τη χρήση ΤΝ στο πλαίσιο της παροχής διοικητικών υπηρεσιών.

Τα εγγενή χαρακτηριστικά και, ταυτόχρονα, οι προκλήσεις της ΤΝ, όπως η αδιαφάνεια και η περιορισμένη έως ανύπαρκτη ικνηλασιμότητα και επεξηγησιμότητα των διεργασιών, των συστάσεων, των αποφάσεων και, εν γένει, των «στοιχείων εξόδου», καθώς και η μεροληψία που μπορεί να απορρέει είτε από την έλλειψη αντιπροσωπευτικότητας είτε από στρεβλώσεις των δεδομένων ή ακόμα να εμφιλοχωρεί στον ίδιο τον σχεδιασμό ενός συστήματος, ενδέχεται να οδηγήσουν σε περιορισμούς των δικαιωμάτων που συνδέονται με την πρόσβαση και την απόλαυση κοινωνικών και προνοιακών υπηρεσιών,¹²¹ την εκπαίδευση, την εργασία και την απασχόληση.¹²²

¹²⁰ Βλ. το Κεφάλαιο 7 της παρούσας μελέτης.

¹²¹ Χαρακτηριστικό, αλλά όχι μοναδικό παράδειγμα είναι η περίπτωση του *Systeem Risico Indicatie* (SyRI) στην Ολλανδία, όπου το Περιφερειακό Δικαστήριο της Χάγης στην υπόθεση *Federatie Nederlandse Vakbeweging v. Netherlands* έκρινε το 2020 ότι ένα αλγοριθμικό σύστημα ανίχνευσης απάτης που προέβαινε σε κατάρτιση προφίλ πολιτών βάσει κοινωνικο-οικονομικών χαρακτηριστικών παραβίαζε το δικαίωμα στην ιδιωτική ζωή του άρθρου 8 της ΕΣΔΑ, θεμελιώνοντας ορόσημο για την εφαρμογή του Δικαίου των ανθρωπίνων δικαιωμάτων στη χρήση της ΤΝ στον τομέα κοινωνικής πρόνοιας.

¹²² Βλ. το Κεφάλαιο 8 της παρούσας μελέτης.

4.2. Η Σύμβαση-Πλαίσιο του Συμβουλίου της Ευρώπης για την Τεχνητή Νοημοσύνη

Ορόσημο για τη ρύθμιση της Τεχνητής Νοημοσύνης ήταν η Σύμβαση-Πλαίσιο για την ΤΝ (AI Framework Convention) που υιοθέτησε το Συμβούλιο της Ευρώπης. Η Σύμβαση-Πλαίσιο υιοθετεί –σε γενικές γραμμές, αν και με διαφορετικές διατυπώσεις– τη στοχοθεσία της AI Act: Σύμφωνα με το Προοίμιο, βασικός σκοπός της είναι να διασφαλίσει ότι οι δυνατότητες των τεχνολογιών ΤΝ να προάγουν την ανθρώπινη ευημερία, το ατομικό και κοινωνικό ευ ζην, και να καθιστούν τον κόσμο πιο παραγωγικό, καινοτόμο και ασφαλή αξιοποιούνται με υπεύθυνο τρόπο που σέβεται, προστατεύει και υλοποιεί τις κοινές αξίες των μερών (που θα υπογράψουν και θα κυρώσουν τη σύμβαση) και είναι σύμφωνος με τα ανθρώπινα δικαιώματα, τη δημοκρατία και το κράτος δικαίου.

Σε αντίθεση με την AI Act, η Σύμβαση δεν εισάγει ρυθμίσεις που αφορούν την ασφάλεια προϊόντων. Εστιάζει ακριβώς στην προστασία δικαιωμάτων, του κράτους δικαίου και της δημοκρατίας και δεν ρυθμίζει ρητά τις οικονομικές διαστάσεις και την αγορά των συστημάτων ΤΝ. Η Σύμβαση-Πλαίσιο δεν χαρακτηρίζεται από την προσέγγιση με βάση τον κίνδυνο και τη ρυθμιστική πυκνότητα της AI Act, ούτε εισάγει ουσιαστικούς μηχανισμούς εφαρμογής, αλλά εδράζεται περισσότερο σε μια προσέγγιση επί τη βάση αρχών με αποδέκτες τα συμβαλλόμενα μέρη που πρέπει να υιοθετήσουν κανονιστικά, διοικητικά και άλλα μέτρα, προκειμένου να διασφαλίσουν ότι η χρήση της ΤΝ εναρμονίζεται με τις αρχές αυτές.

Παρά την αφηρημένη διατύπωση, άξια αναφοράς είναι η επιταγή προς κάθε συμβαλλόμενο μέρος να διασφαλίζει ότι τα συστήματα ΤΝ δεν χρησιμοποιούνται κατά τρόπο που να υπονομεύει την ακεραιότητα, την ανεξαρτησία και την αποτελεσματικότητα των δημοκρατικών θεσμών και διαδικασιών, συμπεριλαμβανομένης της αρχής της διάκρισης των εξουσιών, του σεβασμού της ανεξαρτησίας της Δικαιοσύνης και της πρόσβασης στη Δικαιοσύνη. Η δεύτερη, δε, παράγραφος του ίδιου άρθρου διευκρινίζει ότι τα μέτρα που πρέπει να λάβει ένα συμβαλλόμενο μέρος περιλαμβάνουν τη δίκαιη πρόσβαση των ατόμων στον δημόσιο διάλογο, τη συμμετοχή τους σε αυτόν και τη δυνατότητά τους να διαμορφώνουν ελεύθερα τις απόψεις τους.

Η Σύμβαση-Πλαίσιο αναδεικνύει σε αφετηριακή αρχή τον σεβασμό της ανθρώπινης αξιοπρέπειας και της ατομικής αυτονομίας, ενώ, όπως και η AI Act, υιοθετεί την αρχή της διαφάνειας και εποπτείας, της λογοδοσίας και της ευθύνης, την αρχή της ισότητας και της απαγόρευσης διακρίσεων, την προστασία της ιδιωτικότητας και των προσωπικών δεδομένων και την αξιοπιστία των συστημάτων ΤΝ. Στη φάση των διαπραγματεύσεων είχε δοθεί ιδιαίτερη έμφαση στην εκτίμηση των επιπτώσεων, η οποία απέληξε στην υιοθέτηση ενός πλαισίου διαχείρισης κινδύνων με το οποίο θα πρέπει να εντοπίζονται, να αξιολογούνται, να προλαμβάνονται και να μετριάζονται οι επιπτώσεις της χρήσης ΤΝ στα ανθρώπινα δικαιώματα, στη δημοκρατία και το κράτος δικαίου.

Η ρυθμιστική εμβέλειά της μπορεί να θεωρηθεί περιορισμένη ή έστω λιγότερο στιβαρή σε σχέση με την AI Act, καθώς η ευελιξία που αναγνωρίζεται στα συμβαλλόμενα μέρη ως προς τα μέτρα που θα λάβουν,¹²³ όπως, π.χ., η ευχέρεια εφαρμογής της στον ιδιωτικό τομέα, είναι πιθανό να έχει ως αποτέλεσμα έλλειψη συνοχής και διαφοροποιήσεις ως προς το επίπεδο προστασίας. Ωστόσο, η επεξεργασία και υιοθέτηση της Σύμβασης αυτής και μάλιστα την ίδια περίοδο με την AI Act (2024) έχει ιδιαίτερη σημασία, όχι μόνο λόγω της έμφασης στα δικαιώματα και τη δημοκρατία, αλλά και γιατί η διαπραγμάτευση της Σύμβασης δεν πραγματοποιήθηκε μόνο από κράτη-μέλη του Συμβουλίου της Ευρώπης, κύκλος ούτως ή άλλως ευρύτερος της ΕΕ, αλλά και με τη συμμετοχή κρατών όπως οι ΗΠΑ, ο Καναδάς, η Ιαπωνία, η Ουρουγουάη και το Ισραήλ. Η επιλογή του εργαλείου της Σύμβασης-Πλαίσιο απηχεί τη φιλοδοξία να λειτουργήσει ως ένα διεθνές σημείο αναφοράς για την προστασία των ανθρωπίνων δικαιωμάτων σε σχέση με τη χρήση συστημάτων ΤΝ, καθιερώνοντας κοινές αρχές για ένα ευρύτερο φάσμα κρατών. Ακόμη και αν πρόκειται για ρυθμίσεις στο επίπεδο του ελάχιστου κοινού παρονομαστή, η συναίνεση στην καθιέρωση κάποιων κοινών κανόνων έχει βαρύτητα όταν η τεχνολογία, οι χρήσεις, οι επιπτώσεις και οι κίνδυνοί της έχουν de facto «διεθνοποιημένο» χαρακτήρα και υπερβαίνουν προφανώς τα εθνικά ή τα ενωσιακά σύνορα.

¹²³ Σύμφωνα με τη γενική προσέγγιση, κάθε συμβαλλόμενο μέρος οφείλει να εφαρμόζει τις γενικές κοινές αρχές κατά τρόπο πρόσφορο προς το εσωτερικό του νομικό σύστημα.

4.3. Η Τεχνητή Νοημοσύνη και οι δημοκρατικές διαδικασίες

Ένα πεδίο στο οποίο εκφράζονται επιφυλάξεις για τις πιθανές επιπτώσεις της χρήσης συστημάτων Τεχνητής Νοημοσύνης είναι οι λεγόμενες «δημοκρατικές διαδικασίες». Όπως σημειώνεται στην προαναφερόμενη έκθεση του Συμβουλίου της Ευρώπης, ως δημοκρατικές διαδικασίες αντιλαμβανόμαστε δραστηριότητες που σχετίζονται με την οργάνωση και τη διοίκηση εκλογικών διαδικασιών και δημοψηφισμάτων, τη διαχείριση των εκλογικών καταλόγων, την καταχώριση υποψηφίων, τη διεξαγωγή της ψηφοφορίας, την καταμέτρηση των ψήφων, τη συγκέντρωση και ανακοίνωση των αποτελεσμάτων, καθώς και τους μηχανισμούς εποπτείας και παρακολούθησης. Η έννοια των δημοκρατικών διαδικασιών καλύπτει, επίσης, τον σχεδιασμό και την υλοποίηση εκστρατειών δημόσιας ενημέρωσης, την εκπαίδευση των ψηφοφόρων, τη διάδοση εκλογικών πληροφοριών και μέτρα για τη διασφάλιση της προσβασιμότητας και της ακεραιότητας της εκλογικής διαδικασίας.

Οι χρήσεις αυτές μπορεί να αφορούν από εργαλεία υποστήριξης των εκλογέων (χρήση chatbots και virtual assistants, αλλά και εργαλεία υποστήριξης προσβασιμότητας για άτομα με αναπηρίες), εργαλεία που διασφαλίζουν την ακεραιότητα των εκλογών (εργαλεία ΤΝ για τη σάρωση, καταμέτρηση και επαλήθευση ψηφοδελτίων, καθώς και συστήματα ανίχνευσης κυβερνοαπειλών), έως εργαλεία και χρήσεις που μπορεί να επηρεάσουν την κοινή γνώμη, το είδος και την κατεύθυνση της πληροφόρησης των εκλογέων και την ίδια την εκλογική συμπεριφορά τους (ανάλυση συναισθήματος, π.χ., μέσω της επεξεργασίας αναρτήσεων και δεδομένων από μέσα κοινωνικής δικτύωσης, επεξεργασία ιστορικών εκλογικών δεδομένων, δημογραφικών στοιχείων και κοινωνικών τάσεων, με στόχο την πρόβλεψη της εκλογικής συμμετοχής, τον εντοπισμό περιοχών που χρήζουν ενίσχυσης της συμμετοχής και τη βέλτιστη κατανομή πόρων από τους υποψηφίους). Πέραν των συστημάτων που εστιάζουν στη (μικρο)στοχευμένη συμπεριφορική πολιτική διαφήμιση με εξατομικευμένα πολιτικά μηνύματα σε συγκεκριμένους ψηφοφόρους και ψηφιακά περιβάλλοντα, κατά τα τελευταία χρόνια βαίνει αύξουσα η αξιοποίηση (ή κατάχρηση) της Παραγωγικής ΤΝ σε πολιτικές εκστρατείες με χρήση συνθετικού κειμένου, ήχου ή οπτικοακουστικού υλικού (π.χ., τεχνολογίες deepfake) για σκοπούς πολιτικής επικοινωνίας και άσκησης πολιτικής επιρροής.

Οι χρήσεις ΤΝ ενδέχεται να θέτουν ζητήματα παραβίασης σημαντικών δικαιωμάτων, όπως το δικαίωμα στις ελεύθερες εκλογές,¹²⁴ την ελευθερία της έκφρασης, την προστασία των προσωπικών δεδομένων, το δικαίωμα του συνέρχεσθαι και του συνεταιρίζεσθαι, και το δικαίωμα συμμετοχής σε πολιτικά κόμματα. Οι επιπτώσεις αυτές δεν αφορούν μόνο τα πολιτικά και τα ατομικά δικαιώματα των προσώπων, αλλά έχουν αντίκτυπο στην ίδια τη λειτουργία και την ποιότητα της δημοκρατίας. Η παραπληροφόρηση που καθοδηγείται από την ΤΝ, η αλγοριθμική ενίσχυση πολιτικού περιεχομένου, η ψυχογραφική μικροστόχευση ψηφοφόρων και τα συνθετικά μέσα που παράγονται από την ΤΝ (deepfakes) δημιουργούν συλλογικά ένα περιβάλλον στο οποίο οι προϋποθέσεις για γνήσια και ενημερωμένη δημοκρατική συμμετοχή υπονομεύονται συστηματικά.

¹²⁴ Το δικαίωμα σε ελεύθερες εκλογές κατά το άρθρο 3 του Πρώτου Πρόσθετου Πρωτοκόλλου της ΕΣΔΑ απαιτεί να διεξάγονται οι εκλογές υπό συνθήκες που επιτρέπουν την «ελεύθερη έκφραση της γνώμης του λαού». Το δικαίωμα αυτό δεν προστατεύεται, εάν το πληροφοριακό περιβάλλον εντός του οποίου αυτές διεξάγονται διαστρεβλώνεται συστηματικά από χειραγώγηση μέσω ΤΝ. Το ΕΔΔΑ έχει ερμηνεύσει τη διάταξη αυτή ως απαιτούσα ουσιαστικές συνθήκες δημοκρατικής συμμετοχής, όχι απλώς απουσία τυπικών εμποδίων στην ψήφο – ερμηνεία που πρέπει να επεκταθεί ώστε να περιλάβει και το πληροφοριακό περιβάλλον ως συστατικό στοιχείο της δημοκρατικής ελευθερίας. Βλ. Steering Committee for Human Rights (CDDH) (2026). *Human Rights and Artificial Intelligence: Handbook of the Steering Committee for Human Rights* (p. 62 ff.). Council of Europe. <https://rm.coe.int/artificial-intelligence-handbook/48802b5bf2>

4.4. Τεχνητή Νοημοσύνη, ενημέρωση και παραπληροφόρηση

Η πρόσβαση στην πληροφόρηση, η απόλαυση, η άσκηση και η προστασία ατομικών δικαιωμάτων αποτελούν, εν μέρει, αντικείμενο «ρύθμισης» από ιδιώτες ψηφιακούς φορείς μέσω της αξιοποίησης αυτοματοποιημένων εργαλείων, με χαρακτηριστικό παράδειγμα τη χρήση συστημάτων αυτόματου ελέγχου (moderation) και επιλογής (curation) περιεχομένου. Λόγω της αλγοριθμικής αρχιτεκτονικής, οι ψηφιακές πλατφόρμες μπορούν να εφαρμόσουν οιονεί (ημι)δημόσιες λειτουργίες σε κλάσματα του δευτερολέπτου, χωρίς να υπόκεινται σε καταρχήν έλεγχο δημόσιας εξουσίας, αν και κατά τα τελευταία χρόνια η Πράξη για τις Ψηφιακές Υπηρεσίες (Digital Services Act – DSA) έχει επιβάλει υποχρεώσεις και έχει εισαγάγει εργαλεία ελέγχου.

Όλο και περισσότεροι πάροχοι ψηφιακών υπηρεσιών/online intermediaries, όπως πλατφόρμες ψηφιακών δικτύων και μηχανές αναζήτησης, χρησιμοποιούν συστήματα Τεχνητής Νοημοσύνης 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. Οι αλγόριθμοι ενισχύουν τη μικροστόχευση, καθώς συνδυάζουν μεγάλη ποσότητα πληροφορίας και στη συνέχεια πρόσωπα με συγκεκριμένο προφίλ γίνονται αποδέκτες συγκεκριμένου περιεχομένου. Τα συστήματα συστάσεων¹²⁵ που βασίζονται σε αλγορίθμους (algorithm-based recommendation systems) ασκούν μια μάλλον αμφιλεγόμενη επιρροή σχετικά με τον εντοπισμό και τη διάδοση της πληροφορίας. Η προσωποποίηση του περιεχομένου επί τη βάση του profiling μπορεί να οδηγήσει στα λεγόμενα «echo chambers» και «filter bubbles».¹²⁶ Η –συχνά μέσω αλγορίθμων– στοχευμένη ενημέρωση μπορεί να οδηγήσει στη «στένωση» του ορίζοντα και κατ' αποτέλεσμα των απόψεων και επιλογών μας.¹²⁷

¹²⁵ Χρησιμοποιούν αλγορίθμους για να προτείνουν στον χρήστη περιεχόμενο, προϊόντα, υπηρεσίες ή επιλογές που εκτιμάται ότι τον ενδιαφέρουν.

¹²⁶ Τα «echo chambers» («θάλαμοι ηχούς») είναι περιβάλλοντα όπου κάποιος εκτίθεται κυρίως σε απόψεις ίδιες με τις δικές του. Οι ίδιες ιδέες «αντηχούν» συνεχώς, χωρίς ουσιαστική επαφή με διαφορετικές γνώμες. Τα «filter bubbles» («φούσκες φίλτρων») αφορούν κυρίως τους αλγορίθμους πλατφορμών όπως YouTube, TikTok, Instagram ή Google. Οι αλγόριθμοι παρακολουθούν τι βλέπει κάποιος, τι πατά, τι του αρέσει και του προβάλλουν παρόμοιο περιεχόμενο.

¹²⁷ Εκτενέστερη ανάλυση σε Μήτρου, Λ. (2024). Ψηφιακή δημοκρατία, ψηφιακή συμμετοχή και ψηφιακές απειλές. Στο Γ. Καραβοκύρης (Επιμ.), *Κράτος Δικαίου και Δημοκρατία στην ψηφιακή εποχή* (σ. 53 επ.). Ίδρυμα της Βουλής των Ελλήνων. Επίσης, Ασκητής, Α. (2020). *Ψηφιακή Ελεύθερη Έκφραση*. Νομική Βιβλιοθήκη.

Η ΤΝ αλλάζει όλη την αλυσίδα της είδησης, από τη συλλογή της πληροφορίας έως την παραγωγή, τη διάδοση και την κατανάλωση της πληροφορίας. Μπορεί να επηρεάσει τη δημιουργία και τη διάδοση περιεχομένου λόγω/μέσω της αλλαγής των οικονομικών συνθηκών παραγωγής ειδήσεων. Η πίεση του φθηνού ανταγωνισμού οδηγεί συχνά σε αυτοματοποιημένη παραγωγή περιεχομένου. Περαιτέρω απολήγει και σε σημαντικές αλλαγές στον τρόπο αναζήτησης και ανάδειξης της είδησης: μέσω του scraping και του tagging αλλάζει το περιεχόμενο των ειδήσεων, αλλά και την ανάδειξή τους. Δεν πρέπει, επίσης, να διαφεύγει την προσοχή μας ότι πολλές εταιρείες μέσωσ ενημέρωσης εξαρτώνται από τις υπηρεσίες των παρόχων περιεχομένου και κοινωνικών δικτύων για να αξιοποιήσουν λύσεις ΤΝ: από την ύπαρξη εργαλείων για την ανάλυση δεδομένων έως το fact-checking και τον έλεγχο της προέλευσης περιεχομένου και την αυτόματη παραγωγή και διάδοση περιεχομένου.

Η Πράξη για τις Ψηφιακές Υπηρεσίες δεν εισάγει μόνο την υποχρέωση λογοδοσίας των παρόχων υπηρεσιών διαμεσολάβησης, αλλά ένα νέο νομικό πλαίσιο για τη ρύθμιση περιεχομένου. Η DSA επιβάλλει υποχρεώσεις διαφάνειας και λογοδοσίας ως προς τον έλεγχο του περιεχομένου, τη διαφήμιση και τις αλγοριθμικές διαδικασίες που εφαρμόζουν οι πλατφόρμες, αλλά και την υποχρέωση αιτιολογίας, καθώς και τη δυνατότητα αμφισβήτησης των αποφάσεων ελέγχου του περιεχομένου των πλατφορμών ώστε να αναπτύξουν κατάλληλα εργαλεία διαχείρισης κινδύνων για την προστασία της ακεραιότητας των υπηρεσιών τους έναντι της χρήσης τεχνικών χειραγώγησης. Συγκεκριμένα, οι πολύ μεγάλες επιγραμμικές πλατφόρμες και μηχανές αναζήτησης οφείλουν να αποτρέπουν την κατάχρηση των συστημάτων τους λαμβάνοντας μέτρα βάσει κινδύνου, συμπεριλαμβανομένης της εποπτείας μέσω ανεξάρτητων ελέγχων των μέτρων διαχείρισης κινδύνου.

Σύμφωνα με τη Σύσταση του Συμβουλίου της Ευρώπης,¹²⁸ οι ειδησεογραφικοί οργανισμοί θα πρέπει να αποκαλύπτουν εάν και πώς χρησιμοποιούν συστήματα ΤΝ. Η αποκάλυψη θα πρέπει να πραγματοποιείται σε περιπτώσεις όπου η χρήση ΤΝ μπορεί εύλογα να επηρεάσει τα δικαιώματα ενός ατόμου ή εν γένει του ακροατηρίου ή να επηρεάσει την ερμηνεία των αποτελεσμάτων. Αντίστοιχη ενημέρωση θα πρέπει να υπάρχει και εντός των δημοσιογραφικών οργανισμών αναφορικά με τη χρήση τέτοιων συστημάτων, τους σκοπούς για τους οποίους σχεδιάστηκαν, τις «αξίες» που τα διέπουν, την εκπαίδευση και την εποπτεία. Το human-in the-loop¹²⁹ είναι ίσως αναγκαία, αλλά όχι επαρκής συνθήκη για τη διασφάλιση της απόδοσης προτεραιότητας σε «ανθρώπινες» αξίες. Στο ίδιο πλαίσιο είναι αναγκαία η συντακτική εποπτεία, ιδίως στη χρήση

¹²⁸ Council of Europe, Steering Committee on Media and Information Society. (2023, November 30). *Guidelines on the responsible implementation of artificial intelligence systems in journalism*. <https://rm.coe.int/cdmsi-2023-014-guidelines-on-the-responsible-implementation-of-artific/1680adb4c6>

¹²⁹ Το human-in the-loop (HITL) είναι μοντέλο λειτουργίας της Τεχνητής Νοημοσύνης, όπου ο άνθρωπος παραμένει μέσα στη διαδικασία λήψης αποφάσεων, ελέγχου ή εκπαίδευσης του συστήματος.

αυτοματισμών και Παραγωγικής ΤΝ, προκειμένου να αποφευχθούν σφάλματα ή προκατειλημμένες διαδικασίες και αποτελέσματα. Η συντακτική εποπτεία δεν θα πρέπει να περιορίζεται στα αποτελέσματα, αλλά να εκτείνεται στις διαδικασίες που παράγουν αυτά τα αποτελέσματα. Απαιτείται, δε, κατά μείζονα λόγο όταν πρόκειται για «ευαίσθητα αποτελέσματα» (λόγω συνεπειών σε πρόσωπα ή σε σημαντικές για την κοινωνία διαδικασίες, όπως οι εκλογές). Σύμφωνα με τη Media Freedom Act,¹³⁰ οι πάροχοι πολύ μεγάλων ψηφιακών πλατφορμών (VLOPs) πρέπει να δηλώνουν ότι δεν παρέχουν περιεχόμενο που παράγεται από συστήματα ΤΝ χωρίς να το υποβάλουν σε επανεξέταση από άνθρωπο ή συντακτικό έλεγχο (άρθρο 18 παρ. 1ε).

Η διεκδίκηση της ποιότητας και της αυτονομίας καθίσταται ακόμη πιο αναγκαία, καθώς η δυναμική των αγορών, η εξάρτηση από τις μεγάλες πλατφόρμες και οι διαχειριστικές πρακτικές αυξάνουν την πίεση για τη βελτιστοποίηση της λογικής της αγοράς και της απομείωσης της επαγγελματικής αυτονομίας. Στο πλαίσιο αυτό, έχουν ιδιαίτερη σημασία πρωτοβουλίες αυτορρύθμισης και αυτοελέγχου, όπως ο Κώδικας Δεοντολογίας για τη Χρήση της Παραγωγικής Τεχνητής Νοημοσύνης από Δημοσιογράφους, ο οποίος εκπονήθηκε πρόσφατα από την Πανελλήνια Ομοσπονδία Ενώσεων Συντακτών (ΠΟΕΣΥ), που αποτέλεσε προϊόν συνεργασίας δημοσιογράφων, νομικών και πληροφορικών. Όπως σημειώνει ο Σπύρος Βλαχόπουλος, που υπήρξε και πρόεδρος της επιτροπής που συνέταξε τον Κώδικα, αυτός δεν μπορεί να επιλύσει όλα τα προβλήματα, αλλά η προέχουσα συνεισφορά του έγκειται στην παιδευτική του αξία, αφού υπενθυμίζει στον δημοσιογράφο τις βασικές δεοντολογικές αρχές που θα πρέπει να τηρεί όταν χρησιμοποιεί εφαρμογές ΤΝ. Ταυτόχρονα, ο ίδιος Κώδικας αναπτύσσει και μια προστατευτική για τον δημοσιογράφο λειτουργία απέναντι στον εργοδότη του, δίνοντάς του το δικαίωμα να αρνηθεί νομίμως εντολές που έρχονται σε αντίθεση με τις αρχές του Κώδικα.¹³¹

Ο κίνδυνος διάδοσης ψευδών ειδήσεων και σκόπιμης παραπληροφόρησης αυξάνεται εκθετικά μέσω της ευρείας χρήσης της Παραγωγικής ΤΝ (Generative AI). Η παραπληροφόρηση γίνεται φθηνότερη και πιο αποτελεσματική, και το ποσοτικό ζήτημα εξελίσσεται σε ποιοτικό πρόβλημα: Η παραπληροφόρηση δεν γεννιέται με τη Δημιουργική ΤΝ, αλλά η διαδικτυακή διάδοση παραπληροφόρησης ενισχύεται σε όγκο και εμβέλεια. Οι δυνατότητες

¹³⁰ Πρόκειται για τον Κανονισμό 2024/1083 της 11ης Απριλίου 2024 σχετικά με τη θέσπιση κοινού πλαισίου για τις υπηρεσίες μέσων ενημέρωσης στην εσωτερική αγορά και την τροποποίηση της Οδηγίας 2010/13/ΕΕ (Ευρωπαϊκός Κανονισμός για την ελευθερία των μέσων ενημέρωσης) που τέθηκε σε ισχύ στις 7 Μαΐου 2024, ενώ οι περισσότερες από τις διατάξεις του άρχισαν να εφαρμόζονται στις 8 Αυγούστου 2025. Εισάγει ένα σύνολο κανόνων για την προστασία της πολυφωνίας και της ανεξαρτησίας των μέσων ενημέρωσης στην ΕΕ, προστατεύοντας μεταξύ άλλων τη συντακτική ανεξαρτησία και τις δημοσιογραφικές πηγές. Ενισχύει τη διαφάνεια της ιδιοκτησίας των μέσων ενημέρωσης, της κρατικής διαφήμισης και της «μέτρησης κοινού». Προς διασφάλιση της παρακολούθησης της εφαρμογής συστάθηκε ένα νέο ανεξάρτητο Ευρωπαϊκό Συμβούλιο Υπηρεσιών Μέσων Ενημέρωσης, αποτελούμενο από εκπροσώπους των εθνικών αρχών και φορέων εποπτείας των μέσων ενημέρωσης.

¹³¹ Βλαχόπουλος, Σ. (2025). Ο κώδικας δεοντολογίας για τη χρήση τεχνητής νοημοσύνης από δημοσιογράφους: Πώς θα πρέπει να ρυθμίζουμε την τεχνητή νοημοσύνη; ΔΙΤΕ, 2, 53.

των γλωσσικών παραγωγικών μοντέλων μπορούν να αξιοποιηθούν για να οδηγήσουν τη σκόπιμη παραπληροφόρηση και τη χειραγώγηση σε εντελώς νέα επίπεδα.¹³² Οι κίνδυνοι της παραπληροφόρησης είναι δυνητικά υψηλότεροι όταν αυτά τα γλωσσικά μοντέλα εκπαιδεύονται με επικαιροποιημένη πληροφορία, καθώς οι εκστρατείες σκόπιμης παραπληροφόρησης συχνά εκκινούν από και βασίζονται σε τρέχοντα γεγονότα και καθημερινά ζητήματα.

Αξίζει να σημειωθεί ότι οι άνθρωποι τείνουν να πιστέψουν περισσότερο μια πληροφορία όταν τη βλέπουν να αναρτάται και να διαδίδεται από περισσότερους, αλλά και να επιδεικνύουν μεγαλύτερο βαθμό εμπιστοσύνης σε αποτελέσματα που παράγει η ΤΝ, ιδίως εάν έχει έναν χαρακτήρα επίσημο, αυθεντίας, ή εάν οι άνθρωποι βρίσκονται σε ευάλωτη κατάσταση. Η δύναμη της πειθούς ενισχύεται ακόμη περισσότερο από τον ανθρωπομορφισμό που χαρακτηρίζει συστήματα όπως το ChatGPT.

Η ταχύτητα και η βελτιούμενη συντακτική ακρίβεια αυτών των μοντέλων τα καθιστούν τέλειο εργαλείο για τη μαζική δημιουργία άψογων, φαινομενικά θεμελιωμένων σε δεδομένα, ψευδών ειδήσεων και τη χειραγώγηση της κοινής γνώμης, καθώς και τη δραστική επίπτωση σε δημοκρατικές διαδικασίες, όπως οι εκλογές. Η Ομάδα Εργασίας της Ευρωπαϊκής Επιτροπής για τις ψευδείς ειδήσεις και τη διαδικτυακή σκόπιμη παραπληροφόρηση εντόπισε το πρόβλημα στο ότι τα ψηφιακά μέσα καθοδηγούνται από παράγοντες (κρατικούς ή μη κρατικούς πολιτικούς παράγοντες, κερδοσκοπικούς παράγοντες, πολίτες μεμονωμένους ή σε ομάδες) και από χειραγωγικές χρήσεις των επικοινωνιακών υποδομών που έχουν σχεδιαστεί για να παράγουν, να κυκλοφορούν και να διαδίδουν παραπληροφόρηση σε κλίμακα μεγαλύτερη από ό,τι προηγουμένως.

Η παραπληροφόρηση δημιουργεί άμεσες ή έμμεσες επιπλοκές σε δικαιώματα, όπως η ελευθερία της σκέψης και το δικαίωμα της δημοκρατικής συμμετοχής. Ενέχει τον κίνδυνο της διάβρωσης της δημόσιας εμπιστοσύνης στους θεσμούς, αυξάνοντας την πόλωση και τις ποικίλες εξτρεμιστικές θέσεις και απόψεις. Η διάβρωση της δημοκρατίας επιταχύνεται από τα βασικά χαρακτηριστικά της «οικονομίας της προσοχής και των πλατφορμών» (attention and platform economy), με τις τεχνικές της προσωποποίησης και του κατακερματισμού του ακροατηρίου, αλλά και της χειραγωγικής στόχευσης. Μια στόχευση που διευκολύνεται ιδιαίτερα από τη χρήση της ΤΝ.

Ένα ειδικότερο ζήτημα είναι η επιμόλυνση του πληροφοριακού οικοσυστήματος με όλο και πιο ρεαλιστικά bots και συνθετικά μέσα, όπως τα deepfakes. Όπως σημειώνει ο Βασίλης Καρκατζούνης, ειδικός γραμματέας Τεχνητής

¹³² Ήδη το 2024, μερικούς μήνες ύστερα από την εκρηκτική διάδοση του ChatGPT, η Washington Post ανέφερε σε ρεπορτάζ ότι μέσα σε 7-8 μήνες οι ιστοσελίδες που φιλοξενούσαν fake news, τα οποία δημιουργήθηκαν μέσω ΤΝ, αυξήθηκαν πάνω από 1.000%.

Νοημοσύνης και Διακυβέρνησης Δεδομένων, σύμφωνα με έρευνες, τα deepfake βίντεο αυξήθηκαν κατά 550% την περίοδο 2019-2023, ενώ μέχρι το τέλος του 2025 αναμενόταν να διαμοιραστούν στο Διαδίκτυο έως και 8 εκατομμύρια deepfakes, από περίπου 500.000 το 2023, γεγονός που αντιστοιχεί σε σχεδόν διπλασιασμό κάθε έξι μήνες. Όπως επισημαίνει, υπολογίζεται ότι πάνω από το μισό περιεχόμενο στο Διαδίκτυο έχει παραχθεί μέσω ΤΝ, με τους χρήστες –ακόμη και ειδικούς στον χώρο– να δυσκολεύονται ολοένα και περισσότερο να διακρίνουν το συνθετικό από το πραγματικό.¹³³ Η Πράξη για την Τεχνητή Νοημοσύνη περιλαμβάνει ρυθμίσεις που επιβάλλουν την υποχρέωση σήμανσης τέτοιων προϊόντων βαθυπαραποίησης (deepfakes).

Η απάντηση, πράγματι, δεν μπορεί να είναι μόνο τεχνολογική, αλλά απαιτεί συνδυασμό νομικής ρύθμισης, διαφάνειας και θεσμικής εποπτείας: Στο επίκεντρο της ελληνικής παρέμβασης είναι, αφενός, η «πιστοποίηση της γνησιότητας του περιεχομένου», σε συνδυασμό με ένα «υβριδικό μοντέλο εμπιστοσύνης» με μηχανισμό πιστοποίησης της ανθρώπινης δημιουργίας. Πέραν των νομικών υποχρεώσεων και των τεχνικών ανίχνευσης των deepfakes, επιβάλλεται η ενίσχυση του ψηφιακού γραμματισμού, ώστε να «εκπαιδεύεται» ο θεατής/αναγνώστης. Ειδικά, η αδυναμία διάκρισης του αληθινού και του ψεύτικου, η αίσθηση ότι «δεν ξέρεις πια τι να πιστέψεις», θα απολήξει στη δημιουργία κοινωνιών μηδενικής εμπιστοσύνης, με περαιτέρω συνέπεια τη διάβρωση της εμπιστοσύνης στους θεσμούς.

Οι νέες τεχνολογίες έχουν επηρεάσει ριζικά τον τρόπο με τον οποίο οι άνθρωποι απολαμβάνουν, βιώνουν εν είδει «user experience» την ελευθερία της έκφρασης ως θεμελιώδες δικαίωμα. (Θετική) υποχρέωση του κράτους είναι να διασφαλίζει ότι η ιδιωτική διακυβέρνηση του ψηφιακού περιβάλλοντος δεν βρίσκεται στον αντίποδα της προστασίας της ελευθερίας του λόγου, υπό την παθητική και ενεργητική εκδοχή της. Οι «παραδοσιακές αξίες», όπως η θεμιτή πληροφόρηση, η ελευθερία, η αυτονομία, η έλλειψη προκαταλήψεων, η αναζήτηση της αλήθειας, η αντικειμενικότητα, παραμένουν βασικές αρχές, αλλά ίσως θα πρέπει να επανελέγονται διαρκώς και εν ανάγκη να εμπλουτίζονται και να αναδιατυπώνονται υπό το πρίσμα και τις απαιτήσεις που απορρέουν από τη χρήση συστημάτων ΤΝ.

¹³³ Καρκατζούνης, Β. Μ. (2026, 18 Φεβρουαρίου). Η συμβολή της Ελλάδας στη στρατηγική της ΕΕ για τα deepfakes. Το Βήμα. <https://www.tovima.gr/2026/02/18/the-key-issue/i-symvoli-tis-elladas-sti-stratigiki-tis-ee-gia-ta-deepfakes/>

[ΠΩΣ] ΜΠΟΡΕΙ Η ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΝΑ ΕΙΝΑΙ
ΗΘΙΚΗ ΚΑΙ ΣΥΝΝΟΜΗ;
Ιούλιος 2026

Οι προϋποθέσεις εφαρμογής και αποτελεσματικότητας των νομικών κανόνων



Ιρένε Καμάρα, Λίλιαν Μήτρου, Χαράλαμπος Τσέκερης,
Κατερίνα Χρυσανθοπούλου

Συγγραφική συνεισφορά ανά υποκεφάλαιο:

5.1., 5.2., 5.3., 5.5., 5.7. – Λίλιαν Μήτρου

5.4. – Ιρένε Καμάρα

5.6. – Χαράλαμπος Τσέκερης, Κατερίνα Χρυσανθοπούλου

5.1. Ο μηχανισμός εφαρμογής της Πράξης για την Τεχνητή Νοημοσύνη

Παρά τη φύση της Πράξης για την Τεχνητή Νοημοσύνη ως Κανονισμού, που σημαίνει ότι πρόκειται για κείμενο αμέσου εφαρμογής και υπέρτερης ισχύος, ο εθνικός νομοθέτης θα πρέπει να υιοθετήσει τις απαραίτητες εθνικές διατάξεις για την εφαρμογή της AI Act. Οι διατάξεις αυτές περιλαμβάνουν:

α) τον ορισμό των αρμόδιων Αρχών εποπτείας της αγοράς, σύμφωνα με τις διατάξεις της AI Act, τον εγγύτερο καθορισμό των αρμοδιοτήτων τους και της συνεργασίας/του συντονισμού με άλλες Αρχές. Όλες οι Αρχές εποπτείας της αγοράς που ορίζονται βάσει του AI Act πρέπει να πληρούν τις απαιτήσεις ανεξαρτησίας και αμεροληψίας,

β) τον ορισμό των εθνικών Αρχών προστασίας θεμελιωδών δικαιωμάτων,

γ) τον ορισμό της κοινοποιούσας Αρχής (Notifying Authority) και τον καθορισμό του τρόπου άσκησης των αρμοδιοτήτων της,

δ) τον προσδιορισμό των κυρώσεων στο πλαίσιο και στο περιθώριο που θέτει η AI Act. Οι κυρώσεις που προδιαγράφει η AI Act είναι ιδιαίτερα υψηλές. Ωστόσο, το ακριβές επίπεδο των κυρώσεων εναπόκειται εν τέλει στα κράτη-μέλη.

Το βασικό πρόβλημα που καλείται να αντιμετωπίσει ο νομοθέτης, αλλά και η διοίκηση, όταν πρέπει να προσδιορίσουν και να οργανώσουν ένα σύστημα εποπτείας, είναι η σύμμετρη επίτευξη και εναρμόνιση περισσότερων προϋποθέσεων και στόχων. Συγκεκριμένα, ο εθνικός νομοθέτης θα πρέπει να προβεί σε επιλογές που διασφαλίζουν ταυτόχρονα:

α) την ανεξαρτησία και την αποτελεσματικότητα των μηχανισμών εφαρμογής και εποπτείας της τήρησης των απαιτήσεων του κανονιστικού πλαισίου,

β) την εισαγωγή κυρώσεων που είναι αποτελεσματικές και αποτρεπτικές, λαμβάνοντας υπ' όψιν τις δεδομένες δυνατότητες όσων υποχρεούνται σε συμμόρφωση,

γ) την εναρμόνιση των κανόνων με τα υφιστάμενα κανονιστικά πλαίσια, όπως, π.χ., η προστασία προσωπικών δεδομένων και η ασφάλεια προϊόντων, εφόσον και στο μέτρο που αυτό δεν προβλέπεται ήδη από την AI Act.

Ιδίως ως προς την επίτευξη του πρώτου στόχου είναι σημαντικό να αποτυπωθούν με σαφήνεια και να διασφαλιστούν:

α) ο σαφής καθορισμός και διαχωρισμός των αρμοδιοτήτων των Αρχών, που είτε προϋφίστανται είτε πρόκειται να ιδρυθούν, ώστε να διασφαλιστούν ο συντονισμός και η συνεργασία, και να αποφευχθεί η αλληλεπικάλυψη. Αυτό αφορά ιδίως την Αρχή εποπτείας της αγοράς σε σχέση με τις τομεακές Αρχές εποπτείας της αγοράς,

β) η δημιουργία μηχανισμών συνεργασίας μεταξύ των Αρχών εποπτείας, ιδίως μεταξύ των Αρχών που είναι αρμόδιες για τη διασφάλιση θεμελιωδών δικαιωμάτων και των Αρχών για τον έλεγχο της αγοράς, ώστε να υπάρχει αποτελεσματική και έγκαιρη διαχείριση παραβάσεων, καταγγελιών και παραπόνων,

γ) η δημιουργία μηχανισμών και φορέων υποστήριξης της συμμόρφωσης με δεδομένη την ανάγκη διασφάλισης της αναγκαίας τεχνογνωσίας και εμπειρογνωμοσύνης,

δ) η διασφάλιση τεχνικών, οικονομικών και ανθρώπινων πόρων και υποδομών που θα επιτρέπουν στις Αρχές να εκπληρώνουν τα καθήκοντα και να ασκούν αποτελεσματικά τις αρμοδιότητες που (θα) τους έχουν ανατεθεί,

ε) η εκπόνηση μεθοδολογιών εκτίμησης επιπτώσεων της χρήσης συστημάτων ΤΝ, που θα μπορέσουν να λειτουργήσουν υποστηρικτικά στην εφαρμογή των συναφών ρυθμίσεων, λαμβάνοντας ιδίως υπ' όψιν τα ζητήματα ποιότητας και αντιπροσωπευτικότητας των δεδομένων, την τεχνική σκοπιμότητα και επεκτασιμότητα, τον σχεδιασμό και τις λειτουργίες των σχετικών αλγορίθμων, την προστασία δικαιωμάτων και την ανάγκη η σχετική αξιολόγηση να συμπεριλαμβάνει την ανθρωποκεντρική προσέγγιση, τις ηθικές διαστάσεις και τις επιπτώσεις στην κοινωνία.

5.2. Η διακυβέρνηση της Τεχνητής Νοημοσύνης

Η εφαρμογή των κανονιστικών ρυθμίσεων πρέπει να πλαισιώνεται και να υποστηρίζεται από ένα συνολικότερο πρόγραμμα διακυβέρνησης της Τεχνητής Νοημοσύνης (AI Governance),¹³⁴ που συνιστά υποπερίπτωση ή μετεξέλιξη της ψηφιακής διακυβέρνησης. Ακόμη και εάν δεν υπάρχει καθολικά αποδεκτός ορισμός της διακυβέρνησης της ΤΝ,¹³⁵ όλοι συγκλίνουν ότι πρόκειται για πρακτική που αποσκοπεί στη μεγιστοποίηση των θετικών επιπτώσεων και στην ελαχιστοποίηση των αρνητικών συνεπειών κατά την ανάπτυξη και χρήση συστημάτων ΑΙ.¹³⁶

Η διακυβέρνηση που αποσκοπεί στην οργάνωση της αποδοτικής και αποτελεσματικής υιοθέτησης συστημάτων ΑΙ για τις ανάγκες της ψηφιακής διακυβέρνησης σε δημόσιο και ιδιωτικό τομέα και στην εφαρμογή των απαιτήσεων που απορρέουν από την ΑΙ Act πρέπει να συμπεριλαμβάνει μέτρα για την ασφάλη, σύννομη και ηθική χρήση της ΤΝ, ιδίως στο πλαίσιο λειτουργίας των κυβερνητικών και διοικητικών Αρχών και φορέων.

Όπως αναφέρεται στη σχετική έκθεση του ΟΟΣΑ,¹³⁷ θα πρέπει να διασφαλίζονται οι:

α) Αποδοτικότητα (efficiency), που αναφέρεται στη βελτίωση των εσωτερικών λειτουργιών της δημόσιας διοίκησης.

¹³⁴ Η έννοια της ψηφιακής διακυβέρνησης εξελίχθηκε από τον Luciano Floridi και αφορά την πρακτική της θέσπισης και εφαρμογής πολιτικών, διαδικασιών και προτύπων για την ορθή ανάπτυξη, χρήση και διαχείριση της πληροφορίας (infosphere). Βλ. Floridi, L. (2014). *The Fourth Revolution: How the infosphere is reshaping human reality*. Oxford.

¹³⁵ Οι James Butcher και Irakli Beridze χαρακτηρίζουν τη διακυβέρνηση της ΤΝ ως «μια ποικιλία εργαλείων, λύσεων και μοχλών που επηρεάζουν την ανάπτυξη και τις εφαρμογές της ΤΝ». Butcher, J., & Beridze, I. (2019). What is the state of artificial intelligence governance globally? *THE RUSI Journal*, 164(5-6), 88-96. <https://doi.org/10.1080/03071847.2019.1694260>

¹³⁶ Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications policy*, 44(6), <https://doi.org/10.1016/j.telpol.2020.101976>.

¹³⁷ Organisation for Economic Co-operation and Development (OECD) (2024). *Governing with artificial intelligence: Are governments ready?* [OECD Artificial Intelligence Papers No. 20]. <https://doi.org/10.1787/26324bc2-en>

β) Αποτελεσματικότητα (effectiveness), που αναφέρεται στη βελτίωση της διαδικασίας χάραξης πολιτικής, δηλαδή της διαδικασίας λήψης αποφάσεων για τον σχεδιασμό, την εφαρμογή και –όπου απαιτείται– την τροποποίηση δημόσιων πολιτικών και προγραμμάτων.

γ) Ανταπόκριση (responsiveness), που αναφέρεται στη βελτίωση της παροχής δημόσιων υπηρεσιών, τόσο ως προς την παροχή τους όσο και ως προς τον σχεδιασμό τους.

δ) Λογοδοσία (accountability), που αναφέρεται στην ενίσχυση των μηχανισμών εποπτείας, τόσο εσωτερικών όσο και εξωτερικών.

Η διακυβέρνηση της ΤΝ είναι μια σύνθετη διαδικασία, καθώς περιλαμβάνει τη διακυβέρνηση συστημάτων και υποδομών, τη διακυβέρνηση δεδομένων και τη διαχείριση του ανθρώπινου δυναμικού. Εγγενές στοιχείο της διακυβέρνησης είναι η συνολική προσέγγιση και όχι η αντιμετώπιση της ΤΝ ως πανάκεια. Η ολιστική προσέγγιση περιλαμβάνει καταρχάς τις αποφάσεις αναφορικά με:

α) τη σκοπιμότητα και αναγκαιότητα υιοθέτησης και ανάπτυξης συστημάτων ΤΝ,

β) τον προσδιορισμό των μερών ή/και Αρχών που θα εμπλακούν,

γ) τη διασφάλιση της σύννομης και ηθικής ανάπτυξης και υλοποίησης,

δ) τη διασφάλιση της παρακολούθησης των σχετικών υλοποιήσεων και των πολιτικών που τις υποστηρίζουν.

5.3. Πρότυπα και προδιαγραφές

Κρίσιμη παράμετρο για τη λυσιτελή εφαρμογή των κανονιστικών επιλογών αποτελεί η ανάπτυξη προτύπων που θα «μεταφράζουν» τις νομικές απαιτήσεις σε τεχνικά πρότυπα. Η ουσιαστική εξειδίκευση των απαιτήσεων που εισάγονται για τα συστήματα ΤΝ «υψηλού κινδύνου» (όπως η διαχείριση κινδύνων, η ποιότητα των δεδομένων, η τεχνική τεκμηρίωση, η ανθρώπινη εποπτεία, η ακρίβεια, η ανθεκτικότητα και η κυβερνοασφάλεια) επιτυγχάνεται μέσω εναρμονισμένων ευρωπαϊκών προτύπων.¹³⁸

Η AI Act υιοθετεί το εργαλείο των ευρωπαϊκών εναρμονισμένων προτύπων με σκοπό:

- α) την ενίσχυση της εμπιστοσύνης και της αποδοχής από την αγορά και κατ' επέκταση την προώθηση της καινοτομίας και της ανταγωνιστικότητας,
- β) την ενίσχυση της ασφάλειας δικαίου,
- γ) τη διασφάλιση της ευχερέστερης και αποτελεσματικής εφαρμογής των νομικών απαιτήσεων με μείωση του κόστους της συμμόρφωσης,
- δ) την επίτευξη συνεκτικής και ενιαίας εφαρμογής των κανονιστικών απαιτήσεων,
- ε) τη διευκόλυνση της αξιολόγησης και του ελέγχου συμμόρφωσης των συστημάτων ΤΝ.

Η εφαρμογή των προτύπων παραμένει εθελοντική. Ωστόσο, η συμμόρφωση προς εναρμονισμένα πρότυπα μπορεί να θεμελιώσει τεκμήριο συμμόρφωσης με τις αντίστοιχες νομικές απαιτήσεις της AI Act.

¹³⁸ Η σύνδεση της ύπαρξης εναρμονισμένων προτύπων με την αποτελεσματική συμμόρφωση είναι τέτοια που η Ευρωπαϊκή Επιτροπή, στο πλαίσιο του Digital Omnibus, πρότεινε να συνδεθεί η έναρξη εφαρμογής των κανόνων που διέπουν τα συστήματα ΤΝ «υψηλού κινδύνου» με τη διαθεσιμότητα εργαλείων υποστήριξης, συμπεριλαμβανομένων, μεταξύ άλλων, και των προτύπων.

Η Ευρωπαϊκή Επιτροπή έχει αναθέσει στους ευρωπαϊκούς οργανισμούς τυποποίησης, ιδίως στους CEN και CENELEC,¹³⁹ την εκπόνηση εναρμονισμένων προτύπων για βασικούς τομείς της Πράξης, όπως τα συστήματα διαχείρισης κινδύνου, η διακυβέρνηση δεδομένων, η καταγραφή συμβάντων, η διαφάνεια, η ανθρώπινη εποπτεία, η ακρίβεια, καθώς και η κυβερνοασφάλεια. Οι οργανισμοί αυτοί, υιοθετώντας την προσέγγιση «international first», συμμετέχουν ενεργά στους διεθνείς οργανισμούς τυποποίησης και συμβάλλουν, συνεπώς, στη διαμόρφωση ενός ευρύτερου παγκόσμιου πλαισίου προτύπων για την ΤΝ.¹⁴⁰

Εκτενέστερη ανάλυση των ζητημάτων που συνδέονται με τα εναρμονισμένα πρότυπα και την πιστοποίηση γίνεται στη συμβολή της Ιρένε Καμάρα στην παρούσα μελέτη.

¹³⁹ Οι European Committee for Standardization (CEN) και European Committee for Electrotechnical Standardization (CENELEC) είναι οι δύο βασικοί ευρωπαϊκοί οργανισμοί τυποποίησης που αναπτύσσουν κοινά πρότυπα (standards) για προϊόντα, υπηρεσίες και διαδικασίες στην Ευρώπη.

¹⁴⁰ Εκτενέστερη ανάλυση των ρυθμίσεων στην ιστοσελίδα της Ευρωπαϊκής Επιτροπής: European Commission (2026). *Understanding the standardisation of the AI Act*. Shaping Europe's Digital Future. <https://digital-strategy.ec.europa.eu/en/faqs/understanding-standardisation-ai-act>

5.4. Η τεχνική τυποποίηση και πιστοποίηση ως πυλώνες για αξιόπιστα συστήματα Τεχνητής Νοημοσύνης

5.4.1. Η Πράξη για την Τεχνητή Νοημοσύνη, τεχνικά πρότυπα και πιστοποιήσεις

Η Τεχνητή Νοημοσύνη και οι διάφορες –συχνά καταχρηστικές και χειριστικές– χρήσεις της, από την κατάχρηση συνθετικών προσωπικών εικόνων έως τις εξελιγμένες επιθέσεις ηλεκτρονικού «ψαρέματος» (phishing), έχουν καταστήσει αναγκαία την κανονιστική καινοτομία. Το παράδειγμα της Πράξης για την Τεχνητή Νοημοσύνη, ενός νομοθετήματος που εστιάζει αμιγώς στην τεχνολογία, σηματοδοτεί τη μετάβαση από τη ρύθμιση της συμπεριφοράς και των δεδομένων στη ρύθμιση των (προβλεπόμενων) χρήσεων ενός συστήματος ΤΝ.

Η Πράξη ορίζει διάφορες νομικές απαιτήσεις για συστήματα ΤΝ «υψηλού κινδύνου», όπως η διαχείριση κινδύνου, η τεχνική τεκμηρίωση, η διακυβέρνηση δεδομένων, η ανθρώπινη εποπτεία, η κυβερνοασφάλεια και η ανθεκτικότητα (resilience). Στο άρθρο 40, γίνεται αναφορά σε ευρωπαϊκά (εναρμονισμένα) τεχνικά πρότυπα, ενώ η αιτιολογική σκέψη 121 ορίζει:

«Η τυποποίηση θα πρέπει να διαδραματίσει βασικό ρόλο στην παροχή τεχνικών λύσεων στους παρόχους για τη διασφάλιση της συμμόρφωσης με τον παρόντα Κανονισμό, σύμφωνα με την εξέλιξη της τεχνολογίας, με σκοπό την προώθηση της καινοτομίας, καθώς και της ανταγωνιστικότητας και της ανάπτυξης στην ενιαία αγορά. Η συμμόρφωση με τα εναρμονισμένα πρότυπα, όπως ορίζονται στο άρθρο 2 σημείο 1 στοιχείο γ) του Κανονισμού (ΕΕ) αριθμ. 1025/2012 του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου, τα οποία, υπό κανονικές συνθήκες, αναμένεται να ανταποκρίνονται στην εξέλιξη της τεχνολογίας, θα πρέπει να αποτελεί μέσο απόδειξης της συμμόρφωσης των παρόχων με τις απαιτήσεις του παρόντος Κανονισμού».¹⁴¹

Ο Ευρωπαίος νομοθέτης έχει ήδη ζητήσει από τους ευρωπαϊκούς οργανισμούς τυποποίησης να αναπτύξουν τεχνικά πρότυπα για την υποστήριξη

¹⁴¹ European Union (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) (Recital 121). Official Journal of the European Union.

της Πράξης για την Τεχνητή Νοημοσύνη. Η CEN και η CENELEC έχουν συμφωνήσει να εκτελέσουν το έργο της ανάπτυξης αυτών των εναρμονισμένων προτύπων.¹⁴²

Αναδύεται, επομένως, το ερώτημα: Γιατί ο Ευρωπαίος νομοθέτης επέλεξε αυτούς τους μηχανισμούς για να προσφέρουν καθοδήγηση για τόσο καίρια ζητήματα; Και, μάλιστα, μηχανισμοί που προτείνονται ειδικά όχι από νομοθετικό ή εκτελεστικό όργανο, αλλά από ιδιωτικούς οργανισμούς; Με άλλα λόγια, γιατί τα τεχνικά πρότυπα και οι πιστοποιήσεις είναι κατάλληλα για να διαδραματίσουν έναν τέτοιο ρόλο; Πριν απαντηθεί αυτό το ερώτημα, θα πρέπει να διερευνηθεί εν συντομία η εξέλιξη των τεχνικών προτύπων ως κανονιστικής τεχνικής στην Ευρωπαϊκή Ένωση, παράλληλα με την έναρξη της αλληλεπίδρασης μεταξύ τεχνικών προτύπων και ευρωπαϊκής νομοθεσίας.

5.4.2. Τεχνικά πρότυπα: Από την παροχή προδιαγραφών για βίδες στον σχεδιασμό αξιόπιστων συστημάτων Τεχνητής Νοημοσύνης

Τα τεχνικά πρότυπα είναι έγγραφα που παρέχουν τεχνογνωσία (state-of-the-art), απαιτήσεις, ελέγχους και ορθές πρακτικές για συστήματα, προϊόντα και διαδικασίες σε διαφορετικούς τομείς. Όπως εξηγούν οι JoAnne Yates και Craig N. Murphy:

«Η εθελοντική θέσπιση προτύπων σε οργανωμένη μορφή χρονολογείται από τη δεκαετία του 1880 στην Ευρώπη και τις Ηνωμένες Πολιτείες. Επιτροπές μηχανικών, οι οποίες αρχικά συνήθως χρηματοδοτούνταν από μηχανικούς συλλόγους και αργότερα από οργανισμούς θέσπισης προτύπων, άρχισαν να δημιουργούν πρότυπα που θα υιοθετούνταν από τους κατασκευαστές για να ικανοποιήσουν τις ανάγκες των εταιρικών πελατών».¹⁴³

Τα τεχνικά πρότυπα αναπτύσσονται είτε από *ad hoc* ομάδες εταιρειών ή από καθιερωμένα φόρα και οργανισμούς, το κύριο καθήκον των οποίων είναι η παραγωγή τέτοιων προδιαγραφών. Ένα γνωστό πρότυπο είναι, για παράδειγμα, το πρότυπο ποιότητας ISO 9001, το οποίο είναι ένα διεθνές πρότυπο που παρέχει κριτήρια για υψηλής ποιότητας προϊόντα και υπηρεσίες. Στην Ευρωπαϊκή Ένωση, τρεις μη κερδοσκοπικοί οργανισμοί έχουν το δικαίωμα να συντάσσουν, να αναπτύσσουν και να δημοσιεύουν ευρωπαϊκά τεχνικά πρότυπα: οι οργανισμοί CEN, CENELEC και ETSI. Κάθε κράτος-μέλος της ΕΕ διαθέτει έναν εθνικό φορέα τυποποίησης, όπως ο ΕΛΟΤ στην Ελλάδα.¹⁴⁴

¹⁴² Μέσω της εκτελεστικής απόφασης της Επιτροπής C(2023)3215 - Αίτημα τυποποίησης M/593 της 22ας.5.2023, η οποία αντικαταστάθηκε από την εκτελεστική απόφαση της Επιτροπής C(2025)3871 - Αίτημα τυποποίησης M/613 της 23ης.6.2025.

¹⁴³ Ιδία μετάφραση από το πρωτότυπο στην αγγλική. Yates, J., & Murphy, C. N. (2019). *Engineering rules: Global standard setting since 1880*. JHU Press.

¹⁴⁴ Βλ. ιστοσελίδα του Ελληνικού Οργανισμού Τυποποίησης <https://elot.gr/>

Ενώ κατά τις προηγούμενες δεκαετίες τα τεχνικά πρότυπα εστίαζαν κυρίως στη διαλειτουργικότητα, σήμερα διαδραματίζουν έναν εξίσου σημαντικό ρόλο: Διαμορφώνουν αθόρυβα και σχεδόν αόρατα τις πρακτικές και τις διαδικασίες σε εμπορικές εταιρείες, δημόσιους οργανισμούς και, κατ' επέκταση, την καθημερινότητα των πολιτών. Η επίδρασή τους είναι καθολική: από το ύψος ενός γραφείου που καθορίζεται από εργονομικά τεχνικά πρότυπα και την ασφάλεια των παιχνιδιών που εξαρτάται από πρότυπα χημικών ουσιών, μέχρι τις μεθοδολογίες αξιολόγησης για τα συστήματα πέδησης των αυτοκινήτων. Αντίστοιχα, η λειτουργία των ναυτιλιακών ραντάρ βασίζεται σε πρότυπα ηλεκτρομαγνητικής συμβατότητας και διαχείρισης του ραδιοφάσματος.

Η πιστοποίηση αποτελεί τη βεβαίωση από έναν ανεξάρτητο φορέα ότι ένα σύστημα, προϊόν ή διαδικασία συμμορφώνεται με ένα συγκεκριμένο τεχνικό πρότυπο. Στο πλαίσιο αυτό, ο ανεξάρτητος φορέας πιστοποίησης αξιολογεί τη συμμόρφωση ενός συστήματος και σε περίπτωση θετικού αποτελέσματος της αξιολόγησης εκδίδει τη σχετική πιστοποίηση. Παρόλο που τα πρότυπα τυποποίησης και οι πιστοποιήσεις συχνά συνδέονται, η υιοθέτηση ενός προτύπου από μια εταιρεία δεν συνεπάγεται αυτόματα και την υποχρέωση υποβολής σε διαδικασία αξιολόγησης.

Οι δύο έννοιες έχουν διαφορετικό σκοπό: Τα τεχνικά πρότυπα, αφενός, έχουν *εσωτερικό* προσανατολισμό – στοχεύουν στην καθοδήγηση ενός οργανισμού-αφετέρου, οι πιστοποιήσεις υπηρετούν έναν *εξωτερικό* σκοπό, καθώς λειτουργούν ως τεκμήριο προς τρίτους (π.χ., ρυθμιστικές Αρχές, πελάτες, ακόμη και ανταγωνιστές) ότι οι απαιτήσεις όντως πληρούνται στην πράξη.

Λόγω της ευελιξίας των τεχνικών προτύπων ως ρυθμιστικό μέσο και του μη υποχρεωτικού χαρακτήρα τους –με την έννοια ότι ένας οργανισμός είναι καταρχάς ελεύθερος να επιλέξει εάν θα συμμορφωθεί με ένα συγκεκριμένο τεχνικό πρότυπο–, τα τεχνικά πρότυπα άρχισαν να χρησιμοποιούνται στο Δίκαιο της ΕΕ για την ασφάλεια των προϊόντων τη δεκαετία του '80. Σύμφωνα με αυτό το ρυθμιστικό μοντέλο, το Δίκαιο της ΕΕ για την ασφάλεια των προϊόντων προέβλεπε βασικές νομικές προϋποθέσεις (*essential requirements*), ενώ οι συγκεκριμένες τεχνικές πτυχές δεν αποτελούσαν μέρος του ίδιου του Δικαίου, αλλά ορίζονταν σε ευρωπαϊκά τεχνικά πρότυπα. Αυτά τα τεχνικά πρότυπα αναπτύχθηκαν, και εξακολουθούν να αναπτύσσονται, με σκοπό την υποστήριξη της νομοθεσίας εναρμόνισης της ΕΕ («εναρμονισμένα ευρωπαϊκά πρότυπα»). Όσον αφορά την αξιολόγηση της συμμόρφωσης και τις πιστοποιήσεις, πριν από την τοποθέτηση ενός προϊόντος στην εσωτερική αγορά της ΕΕ, το προϊόν πρέπει να πληροί όλες τις νομικές απαιτήσεις που ορίζονται στη νομοθεσία για την ασφάλεια των προϊόντων. Ενώ για προϊόντα χαμηλότερου κινδύνου μπορεί να επιτραπεί αυτοαξιολόγηση από τον κατασκευαστή, για προϊόντα με υψηλότερους κινδύνους για την

υγεία και την ασφάλεια τρίτοι φορείς διεξάγουν τη διαδικασία (π.χ. φορείς πιστοποίησης).¹⁴⁵

Τόσο ο μηχανισμός των τεχνικών προτύπων όσο και των πιστοποιήσεων αποσκοπούν στην εναρμόνιση του πεδίου και την αύξηση του επιπέδου εμπιστοσύνης στην αγορά. Η Πράξη για την Τεχνητή Νοημοσύνη ακολουθεί το μοντέλο της νομοθεσίας της ΕΕ για την ασφάλεια των προϊόντων, όσον αφορά την τοποθέτηση συστημάτων ΤΝ «υψηλού κινδύνου» στην αγορά. Κατά συνέπεια, τα εναρμονισμένα πρότυπα και οι διαδικασίες αξιολόγησης της συμμόρφωσης αποτελούν σημαντικό μέρος αυτού του νέου νομικού πλαισίου.

5.4.3. Οφέλη από την τυποποίηση και την πιστοποίηση: Προς ένα αξιόπιστο μέλλον για τα συστήματα Τεχνητής Νοημοσύνης;

α) Καθοδήγηση σε έναν λαβύρινθο νομικών απαιτήσεων και επίδειξη νομικής συμμόρφωσης

Είναι ίσως αναπόφευκτο, όταν ο νομοθέτης θεσπίζει απαιτήσεις σχετικά με ένα σύστημα ΤΝ, να προκύπτει ανάγκη για διευκρινίσεις και περαιτέρω καθοδήγηση ως προς την τεχνική τους εφαρμογή. Χαρακτηριστικό παράδειγμα αποτελεί το άρθρο 10 παρ. 2 στοιχείο γ) της Πράξης για την Τεχνητή Νοημοσύνη, επί παραδείγματι, το οποίο ορίζει ότι τα συστήματα ΤΝ «υψηλού κινδύνου» πρέπει να συμμορφώνονται με πρακτικές που λαμβάνουν υπ' όψιν σχετικές λειτουργίες επεξεργασίας και προετοιμασίας δεδομένων και να υιοθετούν ορθές πρακτικές προετοιμασίας δεδομένων, όπως ο σχολιασμός (annotation), η επισήμανση (labelling), ο καθαρισμός (cleansing) και η συγκέντρωση (aggregation). Πώς μπορεί ένας κατασκευαστής ενός τέτοιου συστήματος ΤΝ, και ειδικότερα μια μικρομεσαία (ΜμΕ) ή μια νεοφυής επιχείρηση, να γνωρίζει τις βέλτιστες πρακτικές στην αγορά και να λάβει μια τεκμηριωμένη απόφαση σχετικά με το ποιες από αυτές τις πρακτικές είναι κατάλληλες για το εν λόγω σύστημα; Υπό αυτήν την έννοια, τα εναρμονισμένα πρότυπα στην περίπτωση της Πράξης για την Τεχνητή Νοημοσύνη αποτελούν ένα απαραίτητο στοιχείο για τη νομική συμμόρφωση. Αυτό αποδεικνύεται, επίσης, από τις σχετικές εκκλήσεις επιχειρήσεων προς την Ευρωπαϊκή Επιτροπή για τη μετάθεση της εφαρμογής της Πράξης έως ότου δημοσιευθούν τα σχετικά ευρωπαϊκά πρότυπα.¹⁴⁶ Όπως είχε τονίσει η Ομάδα Εμπειρογνομόνων Υψηλού Επιπέδου για την Τεχνητή Νοημοσύνη στις Κατευθυντήριες Γραμμές Δεοντολογίας για Αξιόπιστη Τεχνητή Νοημοσύνη, εφόσον η πλήρης κατανόηση της λειτουργίας ενός συστήματος δεν είναι εφικτή για όλους τους χρήστες, η πιστοποίηση λειτουργεί ως εγγύηση.

¹⁴⁵ Αυτοί ονομάζονται «κοινοποιημένοι οργανισμοί» (notified bodies).

¹⁴⁶ Grady, P. (2026). *On the AI Act, the EU Need Not Move Fast and Break Things*. Chamber of Progress. <https://progresschamber.org/insights/ai-act-eu-need-not-move-fast-break-things/>

Βεβαιώνει το ευρύτερο κοινό ότι το σύστημα ΤΝ συμμορφώνεται με τις αρχές της διαφάνειας, της αντικειμενικότητας και της νομιμότητας.

β) Ενίσχυση της ανταγωνιστικότητας και της καινοτομίας

Τα τεχνικά πρότυπα και οι πιστοποιήσεις είναι μηχανισμοί που έχουν περάσει με επιτυχία τη δοκιμασία του χρόνου και είναι πλέον αποδεκτό ότι δημιουργούν συνθήκες αξιοπιστίας και εμπιστοσύνης στην αγορά. Μελέτες καταδεικνύουν ότι σε αγορές με υψηλή αβεβαιότητα, όπως στον τομέα της ΤΝ, ο οποίος αλλάζει ραγδαία και χαρακτηρίζεται από ένα ετερογενές τεχνικό τοπίο, τα τεχνικά πρότυπα ενισχύουν την ικανότητα καινοτομίας των εταιρειών.¹⁴⁷ Η έκθεση του Mario Draghi για το μέλλον της ευρωπαϊκής ανταγωνιστικότητας¹⁴⁸ αναδεικνύει τα τεχνικά πρότυπα ως παράγοντες οι οποίοι διευκολύνουν την ενοποίηση σε διαφορετικούς τομείς, που επηρεάζονται από την ψηφιοποίηση και τις αναδυόμενες τεχνολογίες. Πράγματι, η τυποποίηση των βασικών διαδικασιών και απαιτήσεων επιτρέπει στους παράγοντες της αγοράς να αποδεσμεύσουν πόρους προς όφελος της καινοτομίας. Ειδικότερα στον τομέα της ΤΝ, η εφαρμογή κοινών προτύπων για την κυβερνοασφάλεια, την αξιολόγηση και τη διαχείριση κινδύνων επιτρέπει στον κλάδο να εστιάσει τις προσπάθειες σε καινοτόμους τομείς Έρευνας και Ανάπτυξης (R&D).

γ) Πρόληψη της κανονιστικής αποσύνδεσης

Ένα από τα βασικά οφέλη της τεχνικής τυποποίησης είναι η δυνατότητά της να αποτρέπει την κανονιστική αποσύνδεση. Οι νόμοι και οι Κανονισμοί σε τομείς όπως η ΤΝ και οι νέες τεχνολογίες ενδέχεται να καταστούν γρήγορα παρωχημένοι, ιδίως όταν ορίζουν απαιτήσεις που αφορούν συγκεκριμένες τεχνολογίες. Διατηρώντας βασικές νομικές προϋποθέσεις στο κείμενο του νόμου και τις τεχνικές προδιαγραφές σε εναρμονισμένα ευρωπαϊκά πρότυπα, μπορεί να αποφευχθεί ο κίνδυνος κανονιστικής αποσύνδεσης. Η τυποποίηση είναι ομολογουμένως ταχύτερη διαδικασία σε σχέση με τη νομοθετική διαδικασία, παρόλο που ομολογουμένως μπορεί να διαρκέσει πολλά χρόνια.

5.4.4. Στρατηγικές προτεραιότητες και ανοικτά ερωτήματα

Τα τεχνικά πρότυπα και οι πιστοποιήσεις είναι πρωτίστως ένα εργαλείο στα χέρια του νομοθέτη. Η δυνατότητά τους να λειτουργούν ως μηχανισμός εμπιστοσύνης εξαρτάται από μια σειρά παραγόντων, συμπεριλαμβανομένης της

¹⁴⁷ Blind, K., Petersen, S. S., & Riillo, C. A. (2017). The impact of standards and regulation on innovation in uncertain markets. *Research policy*, 46(1), 249-264. <https://doi.org/10.1016/j.respol.2016.11.003>

¹⁴⁸ Στις 9 Σεπτεμβρίου 2024 δημοσιεύθηκε η έκθεση «The future of European competitiveness» του τέως πρωθυπουργού της Ιταλίας και τέως προέδρου της Ευρωπαϊκής Κεντρικής Τράπεζας, για την ανταγωνιστικότητα της ΕΕ.

αγαστής συνεργασίας μεταξύ της Ευρωπαϊκής Επιτροπής και των τεχνικών επιτροπών των ευρωπαϊκών οργανισμών τυποποίησης, και της διασφάλισης ότι τα μέλη αυτών των τεχνικών επιτροπών διαθέτουν την κατάλληλη γνώση και τις δεξιότητες για να αναπτύξουν τα απαραίτητα εναρμονισμένα ευρωπαϊκά πρότυπα.

Η συμπερίληψη και η εκπροσώπηση ποικίλων συμφερόντων είναι υψίστης σημασίας, δεδομένου ότι τα εναρμονισμένα ευρωπαϊκά πρότυπα επηρεάζουν τον σχεδιασμό και τη λειτουργία των συστημάτων ΤΝ «υψηλού κινδύνου». Η τυποποίηση έχει αντίκτυπο σε κρίσιμα ζητήματα, όπως η αλγοριθμική μεροληψία και η ανάπτυξη κατάλληλων μεθοδολογιών αξιολόγησης των μέτρων εξάλειψής της από την πλευρά του κατασκευαστή. Στην περίπτωση που οι τεχνικές επιτροπές απαρτίζονται μόνο ή κυρίως από κατασκευαστές συστημάτων ΤΝ «υψηλού κινδύνου», το αποτέλεσμα θα ήταν πιθανότατα μονομερές: Η επιδίωξη της αποτελεσματικότητας θα υπερίσχυε της προσαρμογής θεμελιωδών δικαιωμάτων. Αντιθέτως, η συμμετοχή οργανώσεων προάσπισης πολιτικών δικαιωμάτων και κρατικών φορέων παρέχει εγγύα για ένα πιο δίκαιο και ισορροπημένο αποτέλεσμα.

Ομοίως, οι ΜμΕ και οι νεοφυείς επιχειρήσεις θα πρέπει, επίσης, να συμμετέχουν στη διαδικασία ανάπτυξης προτύπων τυποποίησης, είτε ως μέλη τεχνικών επιτροπών είτε μέσω διαβουλεύσεων. Οι ΜμΕ και οι startups στην Ελλάδα αποτελούν βασικό σπόνδυλο της οικονομίας, ωστόσο υστερούν σε τεχνογνωσία σε πολλά μέτωπα.¹⁴⁹ Η συμμετοχή σε μια διαδικασία τυποποίησης ωφελεί τόσο την ποιότητα των ίδιων των προτύπων όσο και τις συμμετέχουσες εταιρείες, μέσω της απόκτησης γνώσεων σε εξαιρετικά σύνθετα πεδία, όπως η τήρηση αρχείων από συστήματα ΤΝ ή η διασφάλιση ουσιαστικής ανθρώπινης εποπτείας.

Επιπλέον, τα εναρμονισμένα ευρωπαϊκά πρότυπα είναι απαραίτητα για την ευρωπαϊκή ψηφιακή κυριαρχία.¹⁵⁰ Όπως υπογράμμισε η πρόεδρος της Ευρωπαϊκής Επιτροπής στην ομιλία της για την Κατάσταση της Ένωσης το 2025, η τεχνολογική (και ψηφιακή) κυριαρχία είναι ζωτικής σημασίας.¹⁵¹ Υπό το πρίσμα των υφιστάμενων γεωπολιτικών εντάσεων, η επένδυση σε εγχώριες τεχνολογίες εντός της ΕΕ, δεξιότητες και δυνατότητες διασφαλίζει ότι σε στιγμές κρίσης, η εξάρτηση από ξένους παρόχους δεν θα διαταράξει τις εγχώριες εφοδιαστικές αλυσίδες και την οικονομική ανθεκτικότητα της Ένωσης. Το «Σχέδιο για τη μετάβαση της Ελλάδας στην εποχή της Τεχνητής

¹⁴⁹ Chrysolora, E. (2025, August 3). The paradox of Greece's SMEs. *eKathimerini*. <https://www.ekathimerini.com/economy/1276817/the-paradox-of-greeces-smes/>

¹⁵⁰ Περί ψηφιακής κυριαρχίας: Czerniawski, M. (2025). *EU's Digital Sovereignty and the Rights-Based Imperative Linking Enforcement, Competences and Fundamental Rights*. *Verfassungsblog*. <https://verfassungsblog.de/digital-sovereignty-and-the-rights/>

¹⁵¹ European Commission (2025). *State of the Union 2025 address by President Ursula von der Leyen* (Speech No. SPEECH/25/2053). https://ec.europa.eu/commission/presscorner/detail/en/speech_25_2053.

Νοημοσύνης»,¹⁵² το οποίο αναπτύχθηκε από τη Συμβουλευτική Επιτροπή Υψηλού Επιπέδου για την Τεχνητή Νοημοσύνη, υπογραμμίζει, επίσης, τη σημασία της ψηφιακής κυριαρχίας και τον ρόλο της Ελλάδας στην επίδειξη ηγετικής θέσης και αυτονομίας στον ψηφιακό τομέα. Προκειμένου να διαδραματίσει η τυποποίηση ρόλο στην ενίσχυση της ευρωπαϊκής ψηφιακής κυριαρχίας, τα ευρωπαϊκά πρότυπα πρέπει να ανταποκρίνονται στις ανάγκες της ενιαίας αγοράς της ΕΕ, αντανakλώντας τις αξίες και τις προτεραιότητες του συστήματος αξιών της ΕΕ.

Τέλος, ενώ η τυποποίηση και η πιστοποίηση έχουν όλες αυτές τις δυνατότητες, θα πρέπει να εδραιωθούν η διαφάνεια των διαδικασιών και η λογοδοσία των φορέων πιστοποίησης. Η διαφάνεια και η λογοδοσία αποτελούν εγγυήσεις ότι αυτοί οι δύο μηχανισμοί θα λειτουργήσουν ως πραγματικοί πυλώνες για αξιόπιστα συστήματα ΤΝ.

5.4.5. Καταληκτικές παρατηρήσεις

Η μετάβαση από τη ρύθμιση των δεδομένων στη διαχείριση των χρήσεων των συστημάτων ΤΝ σηματοδοτεί μια κομβική μετατόπιση στην ευρωπαϊκή ψηφιακή στρατηγική. Καθώς η Πράξη για την Τεχνητή Νοημοσύνη οδεύει προς την πλήρη εφαρμογή της, τα τεχνικά πρότυπα και οι πιστοποιήσεις αναδεικνύονται όχι απλώς ως προαιρετικά εργαλεία, αλλά ως απαραίτητες υποδομές για τη νομοθετική συμμόρφωση και την ενίσχυση εμπιστοσύνης της αγοράς.

Τα πρότυπα αποτελούν ακρογωνιαίο λίθο της ευρωπαϊκής ψηφιακής κυριαρχίας. Με τη θέσπιση εγχώριων τεχνικών απαιτήσεων, η ΕΕ μειώνει την εξάρτηση από ξένα τεχνολογικά οικοσυστήματα και προωθεί τη βασισμένη σε αξίες προσέγγισή της στη διεθνή σκηνή. Για να διασφαλιστεί η επιτυχία αυτού του μοντέλου συρρύθμισης, οι πολιτικές πρωτοβουλίες πρέπει να θέσουν ως προτεραιότητα τη συμπεριληπτικότητα και την ενσωμάτωση κατάλληλης εμπειρογνομοσύνης στη διαδικασία τυποποίησης. Η διαδικασία αυτή απαιτεί την ενεργό συμμετοχή των ΜμΕ, των νεοφυών επιχειρήσεων και της Κοινωνίας των Πολιτών, ώστε να διασφαλιστεί ότι τα τεχνικά κριτήρια αναφοράς αντικατοπτρίζουν τα θεμελιώδη δικαιώματα και τα ποικίλα οικονομικά συμφέροντα της Ένωσης.

¹⁵² Συμβουλευτική Επιτροπή Υψηλού Επιπέδου για την Τεχνητή Νοημοσύνη (2024). Σχέδιο για τη μετάβαση της Ελλάδας στην εποχή της Τεχνητής Νοημοσύνης. Ελληνική Δημοκρατία, Προεδρία της Κυβέρνησης & Ειδική Γραμματεία Μακροπρόθεσμου Σχεδιασμού. https://foresight.gov.gr/wp-content/uploads/2024/11/Sxedio_gia_tin_metavasi_TN_Gr.pdf

5.5. Ο ψηφιακός γραμματισμός

Η δημόσια συζήτηση για την Τεχνητή Νοημοσύνη είναι πλέον ευρεία, η χρήση της (κυρίως Παραγωγικής) ΤΝ σε ατομικό επίπεδο έχει γνωρίσει διάδοση σε τόσο βραχύ χρόνο όσο καμία άλλη αντίστοιχη τεχνολογία στο παρελθόν και η υιοθέτησή της στον ιδιωτικό και τον δημόσιο τομέα βαίνει αυξανόμενη. Κρίσιμη παράμετρος της σύννομης, θεμιτής και υπεύθυνης χρήσης είναι η γνώση και η επίγνωση τόσο αναφορικά με τις βασικές τεχνολογικές διαστάσεις όσο και σε σχέση με τους κινδύνους που η χρήση της ΤΝ μπορεί να επιφέρει ως προς τα θεμελιώδη δικαιώματα, αλλά και τη βλάβη που μπορεί να προκαλέσει.

Η Πράξη για την Τεχνητή Νοημοσύνη αναφέρεται ρητά στον γραμματισμό στην ΤΝ στο άρθρο 4, το οποίο φέρει τον τίτλο «AI Literacy» και αποτελεί μία από τις σημαντικότερες καινοτομίες του Κανονισμού σε σχέση με την ανθρωποκεντρική διάσταση της ΤΝ. Στους παρόχους και τους φορείς ανάπτυξης συστημάτων ΤΝ επάγεται η υποχρέωση να «λαμβάνουν μέτρα για να διασφαλίσουν, στο μέτρο του δυνατού, επαρκές επίπεδο γραμματισμού ΤΝ στο προσωπικό τους και σε κάθε άλλο πρόσωπο που ενεργεί για λογαριασμό τους και το οποίο ασχολείται με την ανάπτυξη και τη λειτουργία συστημάτων ΤΝ». Πρόκειται μεν για νομικά δεσμευτική απαίτηση, εντασσόμενη στο ευρύτερο πλαίσιο υπεύθυνης χρήσης και διακυβέρνησης της ΤΝ, αν και η μη τήρησή της δεν συνδέεται ρητά με την πρόβλεψη κάποιας κύρωσης.

Σε τι συνίσταται αυτή η υποχρέωση ακριβώς επιχειρεί να εξηγήσει η Ευρωπαϊκή Επιτροπή,¹⁵³ ενώ τα σχετικά ζητήματα αναπτύσσουν οι Χαράλαμπος Τσέκερης και Κατερίνα Χρυσανθοπούλου στη συμβολή τους στην παρούσα μελέτη. Κατά βάση πρόκειται για μία υποχρέωση εκπαίδευσης των τυχόν εμπλεκομένων με το σύστημα ΤΝ. Η εκπαίδευση αποσκοπεί:

α) στη δυνατότητα λήψης τεκμηριωμένων/ενήμερων αποφάσεων επί των συστημάτων ΤΝ από τους αρμόδιους εργαζομένους του παρόχου ΤΝ,

¹⁵³ European Commission (n.d.). *Shaping Europe's Digital Future: AI literacy – Questions & answers*. <https://digital-strategy.ec.europa.eu/en/faqs/ai-literacy-questions-answers>

β) στον εφοδιασμό όλων των εμπλεκόμενων στελεχών με την αντιληπτική ικανότητα που απαιτείται για να εξασφαλιστεί ότι ο πάροχος ΤΝ θα είναι σε θέση να συμμορφωθεί με τις υποχρεώσεις του βάσει της AI Act και

γ) στην κατά συνέπεια καλύτερη προστασία των θεμελιωδών δικαιωμάτων των τελικών χρηστών του εκάστοτε συστήματος ΤΝ.¹⁵⁴

Η εκπαίδευση αυτή μπορεί αναμφίβολα να συνεισφέρει στην κατανόηση της λειτουργίας, στην κριτική αξιολόγηση του περιεχομένου και των αποτελεσμάτων (των «στοιχείων εξόδου»), στην υπεύθυνη χρήση της ΤΝ και στην ανάπτυξη όχι μόνο της γνώσης, αλλά και της αντίληψης και της «ηθικής» της συμμόρφωσης τόσο κατά τον σχεδιασμό όσο και σε όλον τον κύκλο ζωής των συστημάτων ΤΝ. Για τον λόγο αυτόν προκαλεί εντύπωση ότι μια πρώτη επέμβαση που επιχειρεί η πρόταση της Ευρωπαϊκής Επιτροπής για το Digital Omnibus αφορά το εύρος μιας υποχρέωσης του άρθρου 4 της Πράξης για την Τεχνητή Νοημοσύνη, που, αν και ενείχε αρκετά στοιχεία αοριστίας, εισήγαγε ένα ουσιαστικό συστατικό υπευθυνότητας και λογοδοσίας για όσους παράγουν, θέτουν σε κυκλοφορία ή χρησιμοποιούν συστήματα ΤΝ, τον γραμματισμό στην Τεχνητή Νοημοσύνη. Ο ενωσιακός νομοθέτης, όπως εκφράστηκε με την AI Act, συνδύασε τον γραμματισμό στην ΤΝ με την άντληση των μέγιστων δυνατών οφελών από την ΤΝ, με παράλληλη προστασία των θεμελιωδών δικαιωμάτων, της Υγείας και της Ασφάλειας, και διασφάλιση του δημοκρατικού ελέγχου.

Όπως προαναφέρθηκε, το άρθρο 4 της AI Act, που ήδη εφαρμόζεται,¹⁵⁵ προβλέπει ότι οι πάροχοι και οι φορείς εφαρμογής συστημάτων ΤΝ λαμβάνουν μέτρα για να εξασφαλίσουν στον μέγιστο δυνατό βαθμό επαρκές επίπεδο γραμματισμού στον τομέα της ΤΝ για το προσωπικό τους και άλλα πρόσωπα που ασχολούνται με τη λειτουργία και τη χρήση συστημάτων ΤΝ για λογαριασμό τους. Η πρόταση Digital Omnibus, ωστόσο, μετατρέπει αυτήν την ήδη ισχύουσα υποχρέωση σε υποχρέωση πλέον της Ευρωπαϊκής Επιτροπής και των κρατών-μελών να «ενθαρρύνουν» τους παρόχους και τους φορείς ανάπτυξης και εφαρμογής να εξασφαλίσουν το εν λόγω επαρκές επίπεδο γραμματισμού.

Ο Ευρωπαίος Επόπτης Προστασίας Δεδομένων (ΕΕΠΔ-EDPS) και το Ευρωπαϊκό Συμβούλιο Προστασίας Δεδομένων (ΕΣΠΔ-EDPB), σε κοινή Γνωμοδότηση που δημοσίευσαν τον Ιανουάριο του 2026, καταλήγουν πως

¹⁵⁴ European Union (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) (Recital 20)*. Official Journal of the European Union. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>

¹⁵⁵ Από τον Φεβρουάριο του 2025. Βλ. European Union (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) (Article 113)*. Official Journal of the European Union. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>

η εν λόγω προτεινόμενη ρύθμιση δημιουργεί στην ουσία μια νέα υποχρέωση για τα κράτη-μέλη, μετατοπίζοντας το κανονιστικό βάρος από τους παρόχους και τους φορείς ανάπτυξης και εφαρμογής συστημάτων ΤΝ.¹⁵⁶ Με τον τρόπο αυτόν, αναιρείται η άμεση και δεσμευτική υποχρέωση των τελευταίων, γεγονός που αποδυναμώνει την αποτελεσματικότητα και τον επιδιωκόμενο σκοπό της ρύθμισης. Όπως καταλήγουν οι ΕΕΠΔ και ΕΣΠΔ στην κοινή τους Γνωμοδότηση, η οποιαδήποτε τυχόν νέα υποχρέωση «ενθάρρυνσης» των παρόχων και των φορέων εφαρμογής συστημάτων ΤΝ από την Επιτροπή και τα κράτη-μέλη θα έπρεπε να λειτουργεί μόνο συμπληρωματικά και όχι κατά τρόπο που υποκαθιστά την υφιστάμενη υποχρέωση του άρθρου 4 της AI Act.¹⁵⁷

¹⁵⁶ European Data Protection Board & European Data Protection Supervisor. (2026). *EDPB–EDPS Joint Opinion 1/2026 on the proposal for a regulation as regards the simplification of the implementation of harmonised rules on artificial intelligence (Digital Omnibus on AI)* (para. 44). https://www.edps.europa.eu/system/files/2026-01/2025-1124-edpb-edps-joint-opinion_simplification-implementation-harmonised-rules-ai_en.pdf

¹⁵⁷ Χελιουδάκης, Λ., & Μήτρου, Λ. (2025). Quo vadis Digital Omnibus: Καινοτομία (και απλοποίηση) ή (απορ)ρύθμιση; *ΔΙΤΕ*, 4, 550.

5.6. Ο γραμματισμός στην Τεχνητή Νοημοσύνη στην Ελλάδα

5.6.1. Γραμματισμός στην Τεχνητή Νοημοσύνη: Η αναγκαιότητα ενός νέου γνωστικού παραδείγματος

Η Ελλάδα βρίσκεται ήδη σε μια κρίσιμη φάση ψηφιακού μετασχηματισμού, καθώς η ραγδαία ενσωμάτωση των λεγόμενων «διαταρακτικών» τεχνολογιών δυνητικά διαμορφώνει νέες κοινωνικο-οικονομικές, εκπαιδευτικές και δημοκρατικές ισορροπίες. Ο γραμματισμός στην *Τεχνητή Νοημοσύνη* δεν αποτελεί πλέον εξειδικευμένη δεξιότητα, αλλά βασική προϋπόθεση για τη δημοκρατική συμμετοχή του πολίτη στη σύγχρονη Κοινωνία της Πληροφορίας, περιλαμβάνοντας γνώσεις και ικανότητες που επιτρέπουν την κριτική, ηθική και υπεύθυνη χρήση συστημάτων ΤΝ, με στόχο την ασφάλή, λειτουργική και αποτελεσματική ένταξη σε ένα αυξητικά ψηφιακό περιβάλλον. Παράλληλα, η ανάπτυξη γραμματισμού στην ηθική της ΤΝ, ως επέκταση της Παιδείας στα Μέσα και την Πληροφορία (Media & Information Literacy), αποτελεί στρατηγική προτεραιότητα για την εδραίωση μιας ακλόνητης αξιακής πυξίδας, ώστε η τεχνολογική πρόοδος να υπηρετεί τον άνθρωπο, τη δημοκρατία και την κοινωνική συνοχή.

Ο γραμματισμός στην Τεχνητή Νοημοσύνη προάγει στρατηγικά την εθνική τεχνο-ανθεκτικότητα και ψηφιακή κυριαρχία, υπερβαίνοντας την περιοριστική έννοια της τεχνικής κατάρτισης και λειτουργώντας ως θεμέλιο της κοινωνίας της γνώσης. Αποτελεί ένα δυναμικό πλαίσιο που συνδυάζει εφαρμόσιμες πρακτικές με θεμελιώδεις αξίες, όπως η δικαιοσύνη και η υπεροχή της ανθρώπινης κρίσης έναντι της αλγοριθμικής υποκατάστασης, σε μια εποχή καθολικής διάδοσης του Διαδικτύου (Τσέκερης κ.ά., 2023). Η υιοθέτηση ισχυρών προγραμμάτων γραμματισμού στην Τεχνητή Νοημοσύνη υποστηρίζει τη μετάβαση σε ένα μοντέλο ψηφιακής ανάπτυξης κατά το οποίο οι πολίτες δεν είναι παθητικοί καταναλωτές αλγοριθμικών υπηρεσιών, αλλά ενεργοί συμμετοχοί στη διαμόρφωση της πληροφοριακής σφαίρας και στη λήψη τεκμηριωμένων αποφάσεων, διασφαλίζοντας ότι η χρήση της ΤΝ υπηρετεί ηθικά τεκμηριωμένους σκοπούς και ενισχύει την ανθρώπινη βούληση.

Σε αυτό το συγκείμενο, το «Σχέδιο για τη μετάβαση της Ελλάδας στην εποχή της Τεχνητής Νοημοσύνης», που εκπονήθηκε από τη Συμβουλευτική Επιτροπή Υψηλού Επιπέδου για την ΤΝ, αναφέρεται στην προώθηση ενός

«πολυδιάστατου γραμματισμού» και στη δημιουργία μιας ολιστικής ανθρωποκεντρικής στρατηγικής για την ασφαλή και παραγωγική αξιοποίηση της ΤΝ στη χώρα, με ενίσχυση της επιχειρηματικότητας, της εκπαίδευσης, του ρυθμιστικού πλαισίου και της λειτουργίας του κράτους. Επίσης, προσδίδει ιδιαίτερη έμφαση στη συστηματική αναβάθμιση της έρευνας και της καινοτομίας, καθώς και σε υλοποιήσιμες «εμβληματικές πρωτοβουλίες» (flagship initiatives), που στοχεύουν στην ταυτόχρονη οικοδόμηση ψηφιακού ανθρώπινου κεφαλαίου, ψηφιακού κοινωνικού κεφαλαίου και ψηφιακού θεσμικού κεφαλαίου, παράλληλα με την ανάπτυξη των τεχνολογικών δυνατοτήτων (υπολογιστική ισχύς, γλωσσικά μοντέλα, αλγόριθμοι κοκ.).

5.6.2. Ενσωμάτωση του γραμματισμού στην Τεχνητή Νοημοσύνη στο πλαίσιο της εθνικής εκπαιδευτικής πολιτικής

Αναγνωρίζοντας την εκπαίδευση ως υψηλή εθνική προτεραιότητα στην εποχή της 4ης Βιομηχανικής Επανάστασης, ο γραμματισμός στην ΤΝ προσεγγίζεται ως μια θεμελιώδης ικανότητα για το σύνολο του πληθυσμού, απαραίτητη για την κοινωνική συνοχή και την οικονομική ανάπτυξη (Tsekeris, 2019). Ειδικά όσον αφορά σε εκπαιδευτικούς και μαθητές, κύριος στόχος είναι η καλλιέργεια ικανοτήτων που θα τους επιτρέπουν να κατανοούν και να αξιολογούν το πώς λειτουργούν τα συστήματα ΤΝ (άρα και να αμφισβητούν τα αποτελέσματά τους), όπως επίσης και να χρησιμοποιούν κριτικά τις νέες τεχνολογίες, έχοντας ενισχυμένη αλγοριθμική επίγνωση (algorithmic awareness) και υπολογιστική σκέψη και μειώνοντας την εξάρτηση από τις μηχανές. Με άλλα λόγια, να μάθουν να αποφεύγουν τη λεγόμενη «Παγίδα του Τούρινγκ» (The Turing Trap).¹⁵⁸

Η συστηματική αλληλεπίδραση με τεχνολογίες ΤΝ, ιδίως με «συνομιλιακά» LLMs, επιβάλλει ως θεμελιώδη αναγκαιότητα μια νέα «γνωστική εγρήγορση», υποστηριζόμενη από στοχευμένα προγράμματα Παιδείας στα Μέσα και την Πληροφορία, που θα διασφαλίζει την ηθική και σκεπτόμενη χρήση του ψηφιακού περιεχομένου. Η «ανθρωποκεντρική» ή «τεχνοανθρωπιστική» προσέγγιση του γραμματισμού στην Τεχνητή Νοημοσύνη φιλοδοξεί ότι η τεχνολογία θα συνεχίσει να λειτουργεί ως ενισχυτής της ανθρώπινης δημιουργικότητας και όχι ως υποκατάστατό της.

Σύμφωνα με πρόσφατη Γνώμη της Εθνικής Επιτροπής Βιοηθικής και Τεχνολογικής (2025),¹⁵⁹ η οποία επισημαίνει την προαναφερθείσα αναγκαιότητα,

¹⁵⁸ Η «Παγίδα του Τούρινγκ» είναι ένας όρος που επινοήθηκε από τον Erik Brynjolfsson (2022), καθηγητή του Πανεπιστημίου του Στάνφορντ, για να περιγράψει τον κίνδυνο τεχνολογικής εξάρτησης που ελλοχεύει όταν η ανάπτυξη της ΤΝ επικεντρώνεται αποκλειστικά σε συστήματα που μιμούνται την ανθρώπινη συμπεριφορά, αντί να «ενισχύουν» ή να «επαυξάνουν» (augment) τις ανθρώπινες ικανότητες.

¹⁵⁹ Εθνική Επιτροπή Βιοηθικής και Τεχνολογικής (ΕΕΒΤ) (2025). Γνώμη για τις εφαρμογές τεχνητής νοημοσύνης στο ελληνικό σχολείο. <https://bioethics.gr/announcements/nea-gnwmh-gia-tis-efarmoges-technhs-nohmosynhs-sto-ellhniko-sxoleio-18-martioy-2025-3219>

η Ελλάδα οφείλει να υιοθετήσει συγκεκριμένες στρατηγικές στο τρίπτυχο «κατανόηση-αξιολόγηση-ηθική χρήση» για την ομαλή ενσωμάτωση της ΤΝ και του γραμματισμού της στην εκπαίδευση. Η ουσιαστική και αποτελεσματική εισαγωγή του γραμματισμού στην Τεχνητή Νοημοσύνη πρέπει να βασιστεί στη θεμελίωση ενός «ισχυρού πλαισίου Παιδείας στα Μέσα και την Πληροφορία» σε όλες τις βαθμίδες της εκπαίδευσης, από την Πρωτοβάθμια έως την Τριτοβάθμια, σε τεχνικά και μη τεχνικά γνωστικά πεδία, αφού οποιαδήποτε τεχνική ή εργαλειακή ενσωμάτωση της ΤΝ προϋποθέτει οριζόντια καλλιέργεια δεξιοτήτων κριτικής ανάλυσης, αξιολόγησης πηγών, κατανόησης αλγοριθμικών συστημάτων και επίγνωσης των ηθικών και κοινωνικών συνεπειών της τεχνολογίας. Σε αυτά τα συμφραζόμενα, καθίστανται αναγκαίες η αναμόρφωση των προγραμμάτων σπουδών, η συστηματική επιμόρφωση των νέων και των εν ενεργεία εκπαιδευτικών, η εξειδίκευση στις παιδαγωγικές σχολές, η ανάπτυξη διδακτικού υλικού που ενισχύει ταυτόχρονα τις δεξιότητες γραμματισμού στην Τεχνητή Νοημοσύνη και την ικανότητα πνευματικού μόχθου, καθώς και η θέσπιση σαφών πρωτοκόλλων ασφάλειας και δεοντολογίας (Chryssanthopoulou, 2025-Χρυσανθοπούλου, 2025).

Παράλληλα, ζητούμενο παραμένει η συμμόρφωση και ανάληψη ευθύνης από τις μεγάλες τεχνολογικές εταιρείες, ώστε τα συστήματα ΤΝ να εκπαιδεύονται σε ηθικές αρχές και να λειτουργούν με διαφάνεια, λογοδοσία, αξιοπιστία και ασφάλεια, με υψηλής ποιότητας σύνολα δεδομένων (datasets), που ενσωματώνουν όλες τις κοινωνικές παραμέτρους. Για να προληφθεί ο κίνδυνος παραπλάνησης ή χειραγώγησης της κοινής γνώμης, καθίσταται επιτακτική η ενσωμάτωση του γραμματισμού στην Τεχνητή Νοημοσύνη και στον τομέα της επικοινωνίας και της διάχυσης πληροφοριών, προκειμένου να ενισχυθούν η κριτική πρόσληψη του ψηφιακού περιεχομένου και η ενημέρωση του κοινού τόσο σε θέματα επικαιρότητας όσο και σε στρατηγικά ζητήματα, αλλά και η υπεύθυνη χρήση των μέσων κοινωνικής δικτύωσης. Η διαφάνεια, η επεξηγησιμότητα των αλγορίθμων (Explainable AI), η συμμόρφωση με διεθνή πρότυπα, οι σαφείς μηχανισμοί εποπτείας και ρύθμισης, και η ευαισθητοποίηση της κοινωνίας εν γένει αποτελούν απαραίτητες προϋποθέσεις για να διατηρείται η ανθρώπινη παρέμβαση στον έλεγχο και τη λήψη αποφάσεων.

5.6.3. Το νομοθετικό πλαίσιο και η ευρωπαϊκή προοπτική

Πρόσφατα έχουν αναληφθεί σημαντικές κανονιστικές πρωτοβουλίες σε ευρωπαϊκό επίπεδο,¹⁶⁰ όπως ο Κανονισμός 2024/1689, ο οποίος αποτελεί το πρώτο ολοκληρωμένο, νομικά δεσμευτικό κείμενο για την ΤΝ σε

¹⁶⁰ European Union (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>

παγκόσμιο επίπεδο και δημιουργεί ένα ισχυρό πλαίσιο ρύθμισης και ασφάλειας¹⁶¹ για την ανάπτυξη και χρήση συστημάτων ΤΝ. Σε συνδυασμό με τις κατευθυντήριες αρχές του Οργανισμού Οικονομικής Συνεργασίας και Ανάπτυξης (OECD, 2022) και τη Σύμβαση-Πλαίσιο του Συμβουλίου της Ευρώπης (Council of Europe, 2024), οι ρυθμίσεις αυτές αποσκοπούν στη διασφάλιση του σεβασμού των θεμελιωδών δικαιωμάτων, της ασφάλειας και των ηθικών και δημοκρατικών αρχών απέναντι στους κινδύνους που ενδέχεται να προκύψουν από ιδιαίτερα ισχυρά και σύνθετα μοντέλα ΤΝ. Αναγκαία κρίνεται η περαιτέρω θεσμοθέτηση ενός ευρωπαϊκού «καθεστώτος ποιότητας» για τα εκπαιδευτικά δεδομένα, ώστε να ενισχυθούν η προστασία της ιδιωτικότητας και η αποφυγή αλγοριθμικών διακρίσεων.

Ο γραμματισμός στην Τεχνητή Νοημοσύνη εστιάζει στην αξιόπιστη ΤΝ, στην εξατομικευμένη μάθηση και τη γεφύρωση του ψηφιακού χάσματος και των ψηφιακών ανισοτήτων εν γένει. Έχοντας δικαιωματικό χαρακτήρα, απαιτεί γνωστική ετοιμότητα για την προστασία της δημοκρατίας και των θεμελιωδών ελευθεριών από την αυτοματοποιημένη λήψη αποφάσεων. Για τους ανηλίκους, η «παιδοκεντρική» προσέγγιση αποσκοπεί σε ασφαλή πλοήγηση με σεβασμό στην ιδιωτικότητα και τη δίκαιη εκπροσώπηση (UNICEF, 2021). Για τους ενήλικους, σε συνάρτηση με τον έξυπνο επαγγελματικό αλφαριθμητισμό, ο γραμματισμός στην Τεχνητή Νοημοσύνη ενισχύει την επανειδίκευση (reskilling) και την υπεύθυνη διαχείριση για την ένταξη του εργατικού δυναμικού στην αυτοματοποιημένη οικονομία (World Economic Forum, 2020).

Η εφαρμογή του γραμματισμού στην Τεχνητή Νοημοσύνη στην Ελλάδα προϋποθέτει την αντιμετώπιση σημαντικών δομικών προκλήσεων, όπως η γλωσσική και πολιτισμική μεροληψία των κυρίαρχων δυτικοκεντρικών μοντέλων και η ανάγκη ενίσχυσης της παραγωγής ποιοτικών ελληνό-γλωσσων μοντέλων^{162,163} και συνόλων δεδομένων, ώστε να περιοριστούν οι τεχνολογικές εξαρτήσεις. Παράλληλα, η θωράκιση του εκπαιδευτικού συστήματος απέναντι στη «γνωστική φτώχεια» (EEBT, 2025) συνδέεται με τη μετάβαση του σχολείου από χώρο παθητικής κατανάλωσης πληροφοριών σε περιβάλλον κριτικής σκέψης, δημιουργικής διερεύνησης και ενεργού μάθησης. Στο πλαίσιο αυτό, η εισαγωγή ρυθμιστικών πλαισίων δοκιμών, ή «ρυθμιστικών δοκιμαστηρίων» (sandboxes), και η ενίσχυση μηχανισμών λογοδοσίας, σύμφωνα και με τις κατευθύνσεις της Βίβλου

¹⁶¹ European Commission (n.d.). *Shaping Europe's Digital Future: AI Act*. <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>

¹⁶² Ερευνητικό Κέντρο «Αθηνά» (2025). *Meltemi: Το πρώτο ανοικτό Μεγάλο Γλωσσικό Μοντέλο για τα Ελληνικά*. <https://www.athenarc.gr/el/news/meltemi-proto-anoihto-megalo-glossiko-montelo-gia-ta-ellinika>

¹⁶³ Ερευνητικό Κέντρο «Αθηνά» (2025). *Παγκόσμια Ημέρα Ελληνικής Γλώσσας: Το Ερευνητικό Κέντρο Αθηνά παρουσιάζει το νέο Μεγάλο Γλωσσικό Μοντέλο Llama-Krikri*. <https://www.athenarc.gr/el/news/pagkosmia-imera-ellinikis-glossas-2025-neo-megalo-glossiko>

Ψηφιακού Μετασχηματισμού 2020-2025,¹⁶⁴ θεωρούνται κρίσιμες για τη μείωση της κανονιστικής αβεβαιότητας και την αποτροπή της αλγοριθμικής υποκατάστασης της ανθρώπινης απόφασης.

5.6.4. Στρατηγικές προτάσεις για τον γραμματισμό στην Τεχνητή Νοημοσύνη

Με βάση τις πρόσφατες κατευθύνσεις της Εθνικής Επιτροπής Βιοηθικής και Τεχνοηθικής (2025) και διεθνείς βέλτιστες πρακτικές, προτείνεται η υιοθέτηση ενός ανθρωποκεντρικού παιδαγωγικού μοντέλου, κατά το οποίο η τεχνολογία λειτουργεί ως εργαλείο ενίσχυσης και όχι υποκατάστασης της μάθησης. Συγκεκριμένα:

α) Η δημιουργία Εθνικού Παρατηρητηρίου Αλγοριθμικής Διαφάνειας θα συμβάλει στην ενημέρωση πολιτών και μαθητών για τη λειτουργία των συστημάτων ΤΝ, ενισχύοντας τον κριτικό γραμματισμό στην ΤΝ και την κατανόηση της αλγοριθμικής λήψης αποφάσεων.

β) Το Εθνικό Αποθετήριο Ελληνικών Δεδομένων θα υποστηρίξει εκπαιδευτικές εφαρμογές ΤΝ, συμβάλλοντας στην ενίσχυση της πολιτισμικής αντιπροσωπευτικότητας και της γλωσσικής επάρκειας.

γ) Τα ρυθμιστικά δοκιμαστήρια θα επιτρέπουν ασφαλή πειραματική αξιοποίηση της ΤΝ σε εκπαιδευτικά σενάρια, προάγοντας τη γνώση μαθητών και εκπαιδευτικών σχετικά με τη λειτουργία των αλγορίθμων με σεβασμό στον ΓΚΠΔ.

δ) Η Παιδεία στα Μέσα και την Πληροφορία αναγνωρίζεται ως βασικός γνωστικός μηχανισμός, ώστε οι πολίτες να χρησιμοποιούν αποτελεσματικά τα ψηφιακά μέσα, να αξιολογούν κριτικά το περιεχόμενο και να συμμετέχουν υπεύθυνα στον δημόσιο λόγο.

ε) Η παιδαγωγική έμφαση στη διαδικασία της σκέψης και τον «πνευματικό μόχθο» στοχεύει στην καλλιέργεια μεταγνωστικών δεξιοτήτων και τη διατήρηση της νοητικής αυτονομίας των μαθητών κατά την αλληλεπίδραση με συστήματα ΤΝ.

στ) Η συστηματική επιμόρφωση εκπαιδευτικών θα τους καταστήσει κριτικούς διαμεσολαβητές μάθησης, ικανούς να καθοδηγούν ηθικά την αξιοποίηση της Παραγωγικής ΤΝ και να καλλιεργούν δεξιότητες ενσυναίσθησης και συνεργασίας.

¹⁶⁴ Ελληνική Δημοκρατία, Υπουργείο Ψηφιακής Διακυβέρνησης (2021). Βίβλος Ψηφιακού Μετασχηματισμού 2020-2025. <https://digitalstrategy.gov.gr/>

ζ) Η διά βίου μάθηση ανανεώνεται με επανειδίκευση (reskilling) του εργατικού δυναμικού και των στελεχών της δημόσιας διοίκησης και των επιχειρήσεων.

η) Η δημιουργία Προγράμματος Ψηφιακής Ενδυνάμωσης Οικογένειας θα υποστηρίξει γονείς και κηδεμόνες στην κατανόηση της λειτουργίας της ΤΝ, ενισχύοντας τον κριτικό γραμματισμό στην Τεχνητή Νοημοσύνη και την υπεύθυνη οικογενειακή διαμεσολάβηση της τεχνολογίας. Ανοικτοί εκπαιδευτικοί πόροι και πιστοποιημένα σεμινάρια θα ενισχύσουν τη δυνατότητά τους να καθοδηγούν τα παιδιά σε ασφαλή και δημιουργική αξιοποίηση εφαρμογών ΤΝ.

θ) Η θέσπιση Εθνικού Προγράμματος Ψηφιακής Συμπερίληψης 65+ θα συμβάλει στη μείωση του ηλικιακού ψηφιακού χάσματος, ενισχύοντας την αυτονομία και την εμπιστοσύνη των πολιτών 65+ στη χρήση ψηφιακών υπηρεσιών με ΤΝ, διασφαλίζοντας προσβασιμότητα, προστασία προσωπικών δεδομένων, ενεργό κοινωνική συμμετοχή και συμπερίληψη, αλλά και διαγενεακή αλληλεγγύη.

Ο γραμματισμός στην Τεχνητή Νοημοσύνη επεκτείνεται πέραν του εκπαιδευτικού τομέα και σε άλλους κρίσιμους τομείς της κρατικής και κοινωνικής λειτουργίας. Στη δημόσια διοίκηση υποστηρίζει την εμπέδωση της ανθρωπίνης εποπτείας,¹⁶⁵ της αλγοριθμικής λογοδοσίας και της διαφανούς λήψης αποφάσεων· στην Υγεία συμβάλλει στην κριτική αξιολόγηση διαγνώσεων και τη διαχείριση μεγάλου όγκου ευαίσθητων ιατρικών δεδομένων· ενώ στον ενεργειακό τομέα ενισχύει τη χρήση ΤΝ για τη λειτουργία έξυπνων δικτύων και τη μοντελοποίηση κλιματικών δεδομένων. Στην αγορά εργασίας και τη δημιουργική οικονομία προωθεί τη συμπληρωματική συνεργασία ανθρώπου-ΤΝ (Human-AI Collaboration)· ενώ στον τομέα της Δικαιοσύνης υποστηρίζει την προστασία των δικαιωμάτων πνευματικής ιδιοκτησίας και τη δίκαιη αξιοποίηση αλγοριθμικών εργαλείων. Όσον αφορά το περιβάλλον, συνεισφέρει στην ανάλυση του ενεργειακού αποτυπώματος και του υλικού κόστους των υποδομών δεδομένων (OECD, 2022· International Energy Agency, 2024). Σε γεωπολιτικό επίπεδο, ο γραμματισμός στην Τεχνητή Νοημοσύνη συνδέεται με την κατανόηση της ψηφιακής κυριαρχίας και των κινδύνων εξάρτησης από εξωτερικές τεχνολογικές υποδομές (UNCTAD, 2021· EPRS, 2020)· ενώ ο «επιστημονικός γραμματισμός» (science literacy) συμβάλλει στην αντιμετώπιση της κρίσης αναπαραγωγικότητας μέσω προώθησης ανοικτού κώδικα και διαφανούς διάθεσης δεδομένων (Royal Society, 2012· UNESCO, 2021). Από την άλλη, η ψυχολογική και ηθική διάσταση του γραμματισμού στην Τεχνητή Νοημοσύνη επιδιώκει θωράκιση των χρηστών προς αποφυγή των επιπτώσεων του ανθρωπομορφισμού, της συναισθηματικής εξάρτησης και προσωποποίησης των συστημάτων ΤΝ

¹⁶⁵ Τσέκερης, Χ., & Ευθυμίου, Γ. (2024, 19 Νοεμβρίου). Τα ηθικά όρια της αλγοριθμικής διακυβέρνησης. *Η Καθημερινή*. <https://www.kathimerini.gr/politics/563329816/ta-ithika-oria-tis-algorithmikis-diakyvernisis/>

(European Commission, 2019), καθώς η παραπλανητική χρήση μεταφορών που αποδίδουν «κατανόηση» ή «σκέψη» ή «ενσυναίσθηση» σε αλγοριθμικά συστήματα μπορεί να δημιουργήσει μη ρεαλιστικές προσδοκίες (έως και ψυχικές διαταραχές). Επιπρόσθετα, η ενσωμάτωση της ΤΝ ως υποστηρικτικής τεχνολογίας για άτομα με αναπηρία ή νευροδιαφορετικότητα δύναται να ενισχύει τη συμπεριληπτική ψηφιακή πρόσβαση (UNICEF, 2021).

5.6.5. Εγγράμματοι στον αβέβαιο ψηφιακό κόσμο

Η ΤΝ αποτελεί ένα σύνολο «διαταρακτικών» τεχνολογιών που εκτείνονται από τα γλωσσικά μοντέλα και τη «δρώσα ΤΝ» (Agentic AI) έως τα συστήματα αναγνώρισης προτύπων και τις έξυπνες εφαρμογές καθημερινής χρήσης. Η Πολιτεία οφείλει να επιδιώξει μια ενεργητική γνωστική θωράκιση της κοινωνίας επενδύοντας στον γραμματισμό στην ΤΝ όχι ως απλή τεχνική δεξιότητα, αλλά ως βασική δημοκρατική αρετή. Η ανάπτυξη του γραμματισμού στην Τεχνητή Νοημοσύνη πρέπει να υπηρετεί το συλλογικό συμφέρον, ενισχύοντας την κριτική σκέψη, τη συνειδητή και αναστοχαστική χρήση της τεχνολογίας και την προστασία της νομικής αυτονομίας των πολιτών.

Ο γραμματισμός στην Τεχνητή Νοημοσύνη στέκεται απέναντι στους συνήθεις και επαναλαμβανόμενους κύκλους τεκνοφοβίας και «τεχνοθητικού πανικού».¹⁶⁶ Είναι ενδεικτικό ότι στην πιο πρόσφατη εμπειρική έρευνα του πέμπτου κύματος του «World Internet Project Greece» (WIP-GR) για την αναδυόμενη ελληνική κοινωνία της ΤΝ (Μανδενάκη κ.ά., 2025), μόνον οι τέσσερις στους δέκα συμμετέχοντες απάντησαν ότι η ΤΝ θα κάνει τη ζωή μας καλύτερη, ενώ οι οκτώ στους δέκα κρίνουν ότι δεν πρέπει να επιτρέπεται στα παιδιά να χρησιμοποιούν προγράμματα ΤΝ. Επιπλέον, η πρόθεση μελλοντικής χρήσης εμφανίζεται πολύ συγκρατημένη ανάμεσα στους Έλληνες και τις Ελληνίδες, που δήλωσαν ότι δεν έχουν χρησιμοποιήσει ποτέ προγράμματα της λεγόμενης Δημιουργικής, Παραγωγικής ή Γενετικής Νοημοσύνης (Generative AI). Σχεδόν οκτώ στους δέκα (78,8%) θεωρούν «καθόλου» ή «όχι πολύ» πιθανό να χρησιμοποιήσουν τέτοια εργαλεία τους επόμενους έξι μήνες. Με άλλα λόγια, παρά την αυξημένη προβολή και συζήτηση για την ΤΝ, η πραγματική υιοθέτηση φαίνεται χαμηλή, υποδηλώνοντας είτε έλλειψη ενδιαφέροντος είτε έλλειψη γραμματισμού. Οι πρόσφατες μελέτες δείχνουν ότι ο γραμματισμός στην Τεχνητή Νοημοσύνη γίνεται πιο αποτελεσματικός όταν συνδυάζεται με τον *γραμματισμό για το μέλλον* (futures literacy), δηλαδή την ικανότητα να «χρησιμοποιεί» κανείς το μέλλον για να κατανοήσει καλύτερα το παρόν, να εξερευνήσει εναλλακτικά σενάρια, να

¹⁶⁶ Γούτα, Α. (2025). Οι τεχνοθηκοί πανικοί της ανθρωπότητας, από τον Πλάτωνα μέχρι την Τεχνητή Νοημοσύνη [Συνέντευξη με τον Χαράλαμπο Τσέκερη]. Αθηναϊκό-Μακεδονικό Πρακτορείο Ειδήσεων. <https://www.amna.gr/home/article/940011/Oi-technoithikoi-panikoi-tis-anthropotitas--apo-ton-Platona-mechritin-Techniti-Noimosuni>

ανακαλύψει κρυμμένες δυνατότητες, να αξιοποιήσει την αβεβαιότητα και να προετοιμαστεί για το αύριο.

Τέλος, η ανθρωποκεντρική προσέγγιση της ψηφιακής μετάβασης έρχεται να μας υπενθυμίσει ότι η τεχνολογία δύναται να λειτουργήσει ως σύμμαχος της κοινωνικής και εκπαιδευτικής προόδου, χωρίς να υποκαθιστά τις θεμελιώδεις ανθρώπινες γνωστικές λειτουργίες και αρετές. Η περιεκτική ηθική και δημοκρατική διακυβέρνηση της τεχνολογίας προϋποθέτει ότι οι αποφάσεις παραμένουν στον άνθρωπο, ενώ ο γραμματισμός στην Τεχνητή Νοημοσύνη αποτελεί βασικό όρο για τη μείωση των πάσης φύσεως ψηφιακών ανισοτήτων και τη συγκρότηση μιας κοινωνίας που δεν παρακολουθεί παθητικά τις τεχνολογικές εξελίξεις, αλλά συμμετέχει ενεργά και με αυτοπεποίθηση στη διαμόρφωση του ψηφιακού της μέλλοντος.

Βιβλιογραφία

- Brynjolfsson, E. (2022). The Turing trap: The promise & peril of human-like artificial intelligence. *Daedalus*, 151(2), 272-287.
- Chrysanthopoulou, K. (2025). Media literacy: A bridge between communication and education. *Envisioning the Future of Communication*, 2(1), 458-470.
- Council of Europe (2024). *Council of Europe framework convention on artificial intelligence and human rights, democracy and the rule of law* (Council of Europe Treaty Series No. 225).
- European Parliamentary Research Service (EPRS) (2020). *Digital sovereignty for Europe*. (EPRS Briefing No. 651992). European Parliament. [https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI\(2020\)651992](https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2020)651992)
- European Commission (2019). *Ethics guidelines for trustworthy AI*. High-Level Expert Group on Artificial Intelligence. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- International Energy Agency (2024). *Electricity 2024: Analysis and forecast to 2026*.
- Organisation for Economic Co-operation and Development (OECD) (2022). *Measuring the environmental footprint of artificial intelligence* (OECD Digital Economy Papers No. 341).
- Royal Society (2012). *Science as an open enterprise*.
- Tsekeris, C. (2019). Surviving and thriving in the Fourth Industrial Revolution: Digital skills for education and society. *Homo Virtualis*, 2(1), 34-42. <https://doi.org/10.12681/homvir.20192>
- UN Trade and Development (UNCTAD) (2021). *Digital economy report 2021: Cross-border data flows and development: For whom the data flow*. United Nations. <https://unctad.org/publication/digital-economy-report-2021>

United Nations Educational, Scientific and Cultural Organization (UNESCO) (2021). *Recommendation on open science*. <https://unesdoc.unesco.org/ark:/48223/pf0000379949>

United Nations Children's Fund (UNICEF) (2021). *Policy guidance on AI for children*. UNICEF Office of Global Insight and Policy.

World Economic Forum (WEF) (2020). *AI procurement in a box: A toolkit for governments to buy artificial intelligence*.

Εθνική Επιτροπή Βιοηθικής και Τεχνοηθικής (ΕΕΒΤ) (2025). *Ετήσια Έκθεση 2025*. <https://bioethics.gr/announcements-26/ethsia-ek8esh-2025-3295>

Εθνική Επιτροπή Βιοηθικής και Τεχνοηθικής (ΕΕΒΤ) (2025). *Γνώμη για τις εφαρμογές τεχνητής νοημοσύνης στο ελληνικό σχολείο*. <https://bioethics.gr/announcements/nea-gnwmh-gia-tis-efarmoges-texnhths-nohmosynhs-sto-ellhniko-sxoleio-18-martioy-2025-3219>

Μανδενάκη, Κ., Τσέκερης, Χ., Κληρονόμος, Ν., Δεμερτζής, Ν., Λιναρδής, Α., Παπαδούδης, Γ., & Φρέντζου, Χ. (2025). *Έκθεση – World Internet Project Greece 2025: Διείσδυση του Διαδικτύου στην Ελλάδα – Στάσεις και απόψεις για την Παραγωγική Τεχνητή Νοημοσύνη*. JustReDI: Ανθεκτικότητα, Συμπερίληψη και Ανάπτυξη: Προς μια Δίκαιη Πράσινη και Ψηφιακή Μετάβαση των Ελληνικών Περιφερειών, TAEDR-0537352, Ελλάδα 2.0 – NextGenerationEU. ΕΚΚΕ.

Τσέκερης, Χ., Δεμερτζής, Ν., Παπαδούδης, Γ., Λιναρδής, Α., Μανδενάκη, Κ., & Χριστοφιλόπουλος, Ε. (2023). *Το διαδίκτυο στην Ελλάδα: Το 4ο κύμα του World Internet Project Greece*. ΕΚΚΕ & Ειδική Γραμματεία Μακροπρόθεσμου Σχεδιασμού.

Χρυσανθοπούλου, Κ. (2025). *Εγγράμματοι σε έναν ψηφιακό κόσμο: Παιδεία στα Μέσα και την Πληροφορία*. Τόπος.

[ΠΩΣ] ΜΠΟΡΕΙ Η ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΝΑ ΕΙΝΑΙ
ΗΘΙΚΗ ΚΑΙ ΣΥΝΝΟΜΗ;
Ιούλιος 2026

Τεχνητή Νοημοσύνη, κράτος δικαίου και διοίκηση



Λίλιαν Μήτρου

6.1. Η συνεισφορά της Τεχνητής Νοημοσύνης στην ανάπτυξη μιας αποτελεσματικής διοίκησης

Η Τεχνητή Νοημοσύνη μπορεί καταρχάς να επιφέρει ή να συνεισφέρει σε σημαντικά υψηλότερα επίπεδα σχεδιασμού, αποτελεσματικότητας και απόδοσης. Πολλές δημόσιες υπηρεσίες και οργανισμοί αναζητούν τη χρήση νέων μορφών ανάλυσης δεδομένων, έτσι ώστε να παρέχουν καλύτερες υπηρεσίες. Οι τεχνικές ΤΝ μπορούν να κάνουν προβλέψεις πολύ περισσότερο περιεκτικές και ακριβείς, καθώς βασίζονται στην ανάλυση τεράστιων ποσοτήτων δεδομένων, υπό την απαραίτητη προϋπόθεση της διαθεσιμότητας ψηφιοποιημένων/επεξεργάσιμων δεδομένων. Η ανάλυση αυτή επιτρέπει περαιτέρω τη διάγνωση «ανωμαλιών» και αποκλίσεων, και την επικέντρωση σε ειδικές περιπτώσεις.¹⁶⁷ Η ΤΝ μπορεί να ενισχύσει τη Διοίκηση στην επαύξηση και αυτοματοποίηση των ακόλουθων δραστηριοτήτων: προσωποποιημένες υπηρεσίες,¹⁶⁸ στοχευμένη προώθηση πληροφόρησης και υπηρεσιών, διερεύνηση της απάτης, εφοδιαστικές αλυσίδες και οργάνωση, ανάλυση ως προς την παροχή δανείων, κατασκευές, αυτόνομα οχήματα, παροχή υπηρεσιών υγείας.

Τεχνολογίες ΤΝ υιοθετούνται όλο και συχνότερα πλέον και στον δημόσιο τομέα, όπως στον τομέα της εκπαίδευσης ή της κοινωνικής πολιτικής, με σκοπό την υποστήριξη κοινωνικών υπηρεσιών ή τον έλεγχο αυτών, τον εντοπισμό της ανάγκης παρέμβασης¹⁶⁹ ή την προδιάγνωση κινδύνου απάτης. Αξίζει να παρατηρήσουμε ότι συστήματα ΤΝ χρησιμοποιούνται και στο πλαίσιο της εφαρμογής του νόμου, καθώς και της διαχείρισης αιτημάτων που σχετίζονται με τη μετανάστευση ή την παροχή ασύλου.¹⁷⁰ Καθώς το κόστος πέφτει και η ΤΝ κερδίζει σε δυνατότητα αναπαραγωγής, οικονομία κλίμακας και αποτελεσματικότητα, αυτά τα εργαλεία φαίνονται από πολλές απόψεις ιδιαίτερα ελκυστικά. Λόγω του υπερενθουσιασμού που επικρατεί κατά το τελευταίο διάστημα δεν λείπουν αναγγελίες και επαγγελίες χρήσης της ΤΝ ως μαγικής λύσης για την επίλυση προβλημάτων ελέγχου και σχεδιασμού.

¹⁶⁷ Ιδιαίτερα χρήσιμο στις περιπτώσεις φορολογικής απάτης ή απάτης σε σχέση με κοινωνικές παροχές.

¹⁶⁸ Αν και προσιδιάζει σε περιβάλλοντα ιδιωτικού τομέα, πολλοί αναλυτές επισημαίνουν ότι η κατάρτιση προφίλ και οι κατηγοριοποιήσεις στις οποίες μπορεί να προβεί η διοίκηση με τη χρήση συστημάτων ΤΝ θα επιτρέπουν τη στοχευμένη παροχή υπηρεσιών σε πολίτες με παρόμοιες ανάγκες ή ενδιαφέροντα.

¹⁶⁹ Για παράδειγμα, η φλαμανδική υπηρεσία Kind en Gezin είχε αναπτύξει ήδη το 2014 ένα τέτοιο σύστημα για να προδιαγνώσει περιπτώσεις παροχής κοινωνικών υπηρεσιών σε παιδιά και οικογένειες που απαιτούν περαιτέρω ελέγχους και επιτήρηση.

¹⁷⁰ Η μεγάλη πλειοψηφία αυτών των συστημάτων κατηγοριοποιείται ως συστήματα «υψηλού κινδύνου» στην Πράξη για την Τεχνητή Νοημοσύνη – Παράρτημα III.

6.2. Μπορεί η Τεχνητή Νοημοσύνη να εφαρμόσει τον νόμο;

Ένα θεμελιακό ζήτημα αφορά την εναρμόνιση μεταξύ των απαιτήσεων του κράτους δικαίου και της χρηστής διοίκησης, και των πλεονεκτημάτων που προσφέρουν αυτές οι «μηχανές». Όταν ένας υπάλληλος ή ένα διοικητικό όργανο προβαίνει σε ενέργειες ή λαμβάνει αποφάσεις (που απαιτούν νοημοσύνη!), υφίστανται –σε ένα κράτος δικαίου– αυτονόητες απαιτήσεις αναφορικά με την ακρίβεια, την αξιοπιστία, την ορθότητα της κρίσης τους, ενώ αναμφίβολα τίθενται τα ζητήματα της δίκαιης και ισότιμης μεταχείρισης, της αιτιολογίας, της λογοδοσίας και της ευθύνης, παράλληλα προς τις επιφυλάξεις για τη διαχειρισσιμότητα, την κοινωνική αποδοχή και την πολιτική νομιμοποίηση της «απόφασης διά της ΤΝ», που τίθενται με επιτακτικό τρόπο.

Αμφισβητείται έντονα κατά πόσον η διοικητική διαδικασία μπορεί να «αυτοματοποιηθεί» και να «ποσοτικοποιηθεί», καθώς και εάν τα συστήματα ΤΝ είναι σε θέση να εντοπίσουν τον εφαρμοστέο κανόνα ή να ερμηνεύσουν αόριστες νομικές έννοιες, όπως «δημόσιο συμφέρον», «σπουδαίος λόγος», «σοβαρός κίνδυνος», «αναγκαία μέτρα», «επείγουσα και απρόβλεπτη ανάγκη». Αν ορισμένες αόριστες νομικές έννοιες μπορούν να τυποποιηθούν λόγω της συχνής εφαρμογής και της παγιωμένης ερμηνείας, δεν συμβαίνει το ίδιο με άλλες, στις οποίες η ερμηνεία πρέπει να γίνει υπό το φως των συγκεκριμένων περιστάσεων. Υπό την προϋπόθεση της εκπαίδευσης με ποιοτικά σύνολα δεδομένων και υπό τον περιορισμό της ελλιπούς επεξηγησιμότητας, συστήματα μηχανικής μάθησης μπορεί να παράγουν συνεκτικά αποτελέσματα, αλλά ταυτόχρονα πιθανολογείται ότι δεν μπορούν να λαμβάνουν υπ' όψιν, τουλάχιστον επαρκώς, τις ιδιαιτερότητες κάθε περίπτωσης.

Ωστόσο, όλο και περισσότερο διατυπώνεται η ανάγκη να υιοθετηθούν συστήματα λήψης αποφάσεων ή έστω υποστήριξης της λήψης αποφάσεων. Παρατηρείται ότι η χρήση εργαλείων ΤΝ για τη λήψη μιας διοικητικής απόφασης (ή μέρους αυτής) γίνεται δεκτή, όταν βασίζεται σε σαφή, σταθερά και περιορισμένα κριτήρια και το νομικό καθεστώς είναι τόσο λεπτομερές, ώστε να μην απαιτείται καμία ευχέρεια εκ μέρους της. Δεκτικές σε διεκπεραίωση μέσω εργαλείων ΤΝ είναι και δραστηριότητες της διοίκησης που είναι απλές και επαναλαμβανόμενες. Η δυνατότητα ανάθεσης διοικητικών εργασιών σε εργαλεία ΤΝ περιπλέκεται ακόμη περισσότερο όταν ένα τέτοιο σύστημα

καλείται να παράγει μία απόφαση που δεν βασίζεται σε δεδομένα, αλλά και σε αξίες και αξιολογήσεις.

Το πρόβλημα τίθεται με ένταση όταν η διοικητική Αρχή πρέπει να προβεί σε σταθμίσεις μεταξύ αντιτιθέμενων δικαιωμάτων και συμφερόντων. Η υποχρέωση της διοίκησης να λαμβάνει υπ' όψιν και να εξετάζει προσεκτικά και αμερόληπτα το σύνολο των δεδομένων και στοιχείων που σχετίζονται με μία περίπτωση, αποδίδοντας στο καθένα από αυτά το ανάλογο βάρος επί της απόφασης που πρόκειται να ληφθεί,¹⁷¹ απορρέει και από την αρχή της χρηστής διοίκησης. Μία γενική αρχή που υιοθετεί τόσο το εθνικό όσο και το Ευρωπαϊκό Δίκαιο και την οποία ο Χάρτης των Θεμελιωδών Δικαιωμάτων της ΕΕ κατοχυρώνει ως «δικαίωμα στη χρηστή διοίκηση» (άρθρο 41).

Η κανονιστική αντίληψη που αρχίζει να διαμορφώνεται λαμβάνει υπ' όψιν το συγκεκριμένο, αλλά και το γενικότερο πλαίσιο της διάκρισης μεταξύ δέσμιας αρμοδιότητας και άσκησης διακριτικής ευχέρειας, αποκλείοντας καταρχήν τη χρήση ΤΝ στη δεύτερη περίπτωση ή υπάγοντάς τη στην υποχρέωση νομοθετικής ρύθμισης και στην εισαγωγή εγγυήσεων για τα δικαιωμάτων των προσώπων. Η ύπαρξη νομικού θεμελίου για την εισαγωγή συστημάτων ΤΝ στη διοίκηση συνιστά απαίτηση που απορρέει και από την αρχή της νομιμότητας. Σύμφωνα με την αρχή αυτή, βάση της διοικητικής δράσης είναι ο νόμος που θέτει τις προϋποθέσεις και τα όριά της. Η διοικητική δράση πρέπει αφενός να εδράζεται στον νόμο και αφετέρου να είναι σύμφωνη με αυτόν.

¹⁷¹ Βλ. Court of Justice of the European Union. (1991). *Judgment of the Court of 21 November 1991* [European Court Reports 1991 I-05469, Document 61990CJ0269, ECLI:EU:C:1991:438]. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:61990CJ0269>

6.3. Η Τεχνητή Νοημοσύνη και οι απαιτήσεις του κράτους δικαίου

Η χρήση εργαλείων Τεχνητής Νοημοσύνης στη διαδικασία υποστήριξης ή λήψης αποφάσεων από τη διοίκηση θέτει, όμως, γενικότερα ζητήματα σε σχέση με ό,τι αντιλαμβανόμαστε ως απαιτήσεις του κράτους δικαίου (rule of law). Οι εξ αυτού απορρέουσες προβλεψιμότητα (foreseeability) και συνεκτικότητα (coherence), όχι μόνο της νομοθεσίας, αλλά και της κυβερνητικής/διοικητικής δράσης, διακυβεύονται όταν χρησιμοποιούνται συστήματα ΤΝ, που –ακριβώς λόγω των εγγενών χαρακτηριστικών τους– δεν είναι διαφανή και ελέγξιμα. Συστήματα ΤΝ που βασίζονται στη χρήση δεδομένων τα οποία είναι δυναμικά και μπορούν να αλλάξουν με τον χρόνο, καταλήγοντας σε διαφορετικά και ενδεχομένως ανεπιθύμητα αποτελέσματα, ή συστήματα ΤΝ που παράγουν αποφάσεις βασισμένα στη δική «τους» αξιακή εκτίμηση και στα δικά «τους» ερμηνευτικά σχήματα μπορεί να βρεθούν ακριβώς στον αντίποδα του κράτους δικαίου.

Ζητήματα εγείρονται και αναφορικά με τις αρχές της ισότητας, της αμεροληψίας και της απαίτησης για δίκαιη διοίκηση. Η αρχή της αμεροληψίας των οργάνων της δημόσιας διοίκησης αποτελεί ειδικότερη έκφραση της αρχής του κράτους δικαίου και επιβάλλει στη δημόσια διοίκηση να λειτουργεί ως χρηστή διοίκηση, βάσει στόχων που καθορίζονται από την προστασία του δημοσίου συμφέροντος, χωρίς εμπάθεια και προκατάληψη έναντι του διοικουμένου. Ο αυτοματισμός και η χρήση συστημάτων ΤΝ μπορεί να ενισχύσουν την ισότητα απέναντι στον νόμο, εξορθολογίζοντας τη λήψη αποφάσεων, περιορίζοντας προκαταλήψεις, διασφαλίζοντας μια μη μεροληπτική απόφαση και εξαλείφοντας τη διαφθορά. Υπό αυτό το πρίσμα, ο Οργανισμός Θεμελιωδών Δικαιωμάτων της Ευρωπαϊκής Ένωσης (European Union Agency for Fundamental Rights – FRA) φαίνεται να αποδέχεται ότι η αλγοριθμική ανάλυση δεδομένων μπορεί να παράγει αποτελέσματα, τα οποία, υπό συγκεκριμένες περιστάσεις, μπορούν να συνεισφέρουν στον περιορισμό των προκαταλήψεων και των στερεοτύπων ακριβώς μέσω του περιορισμού της εξάρτησης από την ανθρώπινη κρίση.¹⁷²

¹⁷² Fundamental Rights Agency (2020). *Getting the future right: Artificial intelligence and fundamental rights*. Publications Office of the European Union. <https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-rights>

Ωστόσο, θα ήταν λάθος να θεωρήσουμε ότι αυτές οι τεχνικές ανάλυσης είναι αντικειμενικές, απλώς και μόνο επειδή καθοδηγούνται από δεδομένα. Η «αλγοριθμική ουδετερότητα» κάθε άλλο παρά δεδομένη είναι, καθώς η χρήση των δεδομένων που εισάγονται προς ανάλυση, ο σκοπός των αλγορίθμων μηχανικής μάθησης (κατηγοριοποίηση, ταξινόμηση, διαχωρισμός) και η ίδια η λειτουργία των αλγορίθμων μπορούν, κάλλιστα, να οδηγήσουν σε αποφάσεις που χαρακτηρίζονται από άνιση μεταχείριση και διακρίσεις. Η επιλογή και η ποιότητα των δεδομένων (έλλειψη αντιπροσωπευτικότητας και ακρίβειας), με τα οποία εκπαιδεύονται και ελέγχονται οι αλγόριθμοι, αλλά και η επιλογή των χαρακτηριστικών και των δομών ανάλυσης που μπορεί να αναπαράγουν ή έστω να αντανakλούν τις αντιλήψεις και τις προκαταλήψεις των σχεδιαστών μπορεί να προδιαγράψουν το αποτέλεσμα που θα παραχθεί. Οι προκαταλήψεις μπορεί να παρεισφρέουν στις επιλογές των συνόλων δεδομένων. Μια αναδυόμενη προκατάληψη μπορεί να ανακύπτει ακριβώς από την εσφαλμένη ερμηνεία των αποτελεσμάτων μιας εφαρμογής ΤΝ, ιδίως στις περιπτώσεις ανάλυσης στατιστικών μεγεθών. Πέραν των επιφυλάξεων ως προς την ποιότητα της απόφασης, καθώς τα αποτελέσματα αποδίδουν ό,τι είναι δεδομένο και όχι ό,τι θα μπορούσε ή θα έπρεπε να είναι, το πρόσθετο πρόβλημα έγκειται ακριβώς στο ότι τα δεδομένα κληροδοτούν τις προκαταλήψεις του παρελθόντος.

Τέτοια ζητήματα ενδέχεται να ανακύψουν στην περίπτωση της χρήσης συστημάτων ΤΝ που προορίζονται να χρησιμοποιηθούν από τις δημόσιες Αρχές ή για λογαριασμό των δημόσιων Αρχών για την αξιολόγηση της επιλεξιμότητας φυσικών προσώπων ως προς βασικές παροχές και υπηρεσίες δημόσιας αρωγής, συμπεριλαμβανομένων των υπηρεσιών υγειονομικής περίθαλψης, καθώς και για τη χορήγηση, μείωση, ανάκληση ή ανάκτηση των εν λόγω παροχών και υπηρεσιών. Όπως αναφέρεται στην αιτιολογική σκέψη 58 της AI Act, πρόκειται για υπηρεσίες υγειονομικής περίθαλψης, παροχές κοινωνικής ασφάλισης, κοινωνικές υπηρεσίες που παρέχουν προστασία σε περιπτώσεις, μεταξύ άλλων, μητρότητας, ασθένειας, εργατικού ατυχήματος, εξάρτησης ή γήρατος και απώλειας εργασίας, καθώς και κοινωνική και στεγαστική αρωγή.

Δεν είναι παράδοξο που η Πράξη για την Τεχνητή Νοημοσύνη κατατάσσει τα συστήματα αυτά στα συστήματα «υψηλού κινδύνου». Τα πρόσωπα που ζητούν ή λαμβάνουν τέτοιες βασικές παροχές και υπηρεσίες δημόσιας υποστήριξης από δημόσιες Αρχές συνήθως εξαρτώνται κατά κανόνα από τις εν λόγω παροχές και υπηρεσίες, και βρίσκονται σε ευάλωτη θέση σε σχέση με τις αρμόδιες Αρχές. Μια απόφαση για την απόδοση ή άρνηση τέτοιων παροχών μπορεί να έχει σημαντικές έως κρίσιμες επιπτώσεις στη ζωή των προσώπων, που συχνά ανήκουν σε ευάλωτες κατηγορίες πληθυσμού. Άλλωστε, η πρόσφατη εμπειρία από τη χρήση τέτοιων συστημάτων δικαιολογεί την άποψη που ενστερνίζεται και ο ενωσιακός νομοθέτης, ότι τα συστήματα ΤΝ που χρησιμοποιούνται για τους σκοπούς αυτούς μπορούν να

οδηγήσουν σε διακρίσεις μεταξύ προσώπων ή ομάδων και να διαιωνίσουν ιστορικές νοοτροπίες διακρίσεων, όπως αυτές που βασίζονται στη φυλετική ή εθνοτική καταγωγή, στο φύλο, στις αναπηρίες, στην ηλικία ή στον γενετήσιο προσανατολισμό, ή να δημιουργήσουν νέες μορφές διακρίσεων.¹⁷³ Μια απόφαση, ειλημμένη επί τη βάσει συστημάτων ΤΝ, μπορεί να παράγει διακρίσεις στο μέτρο που η ανάλυση εστιάζεται ή περιορίζεται σε μεθόδους αναγνώρισης προτύπων (pattern) και δεν συμπεριλαμβάνει τη διάγνωση και αξιολόγηση αιτίων και αποτελεσμάτων. Μια προκατάληψη μπορεί να ανακύπτει ακριβώς από την εσφαλμένη ερμηνεία των αποτελεσμάτων μιας εφαρμογής ΤΝ, ιδίως στις περιπτώσεις ανάλυσης στατιστικών μεγεθών.

Στις περιπτώσεις αυτές η διακριτική αντιμετώπιση μπορεί να οφείλεται στον σχεδιασμό του συστήματος, στα δεδομένα που χρησιμοποιούνται ή σε παραμέτρους για τη λήψη μιας απόφασης που συνδέουν έναν τόπο κατοικίας με την εθνοτική καταγωγή ή την κοινωνική θέση, δημιουργώντας, μάλιστα, την εντύπωση μιας αιτιώδους σχέσης, η οποία ωστόσο είναι πολύ πιθανό να μην υφίσταται. Οι προκαταλήψεις μπορεί να εκδηλώνονται μέσω των χαρακτηριστικών ενός συστήματος, π.χ. οι μετρικές και οι δομές ανάλυσης, οι οποίες επιλέγονται από τους σχεδιαστές για τα συστήματα ΤΝ και ενδέχεται να αναπαράγουν τις προϋδεάσεις και τις προκαταλήψεις των εμπνευστών ή των δημιουργών τους. Κρίσιμη είναι η διάκριση μεταξύ διαφορών σε επίπεδο ομάδων και της συμπεριφοράς σε ατομικό επίπεδο, με αποτέλεσμα να μην μπορεί να συναχθούν ασφαλείς προγνώσεις για τη συμπεριφορά ενός ανθρώπου επί τη βάσει μίας συλλογικής συμπεριφοράς μίας ομάδας. Ωστόσο, η κατηγοριοποίηση σε ομάδες απολήγει σε δυσμενή μεταχείριση σε ατομικό επίπεδο.

Δεν εγείρονται μόνο σοβαροί προβληματισμοί από την άποψη της ηθικής, αλλά τίθενται και ουσιαστικά νομικά ζητήματα, καθώς μπορεί να παραβιάζονται θεμελιώδη δικαιώματα των προσώπων αυτών, όπως το δικαίωμα στην κοινωνική προστασία, στη μεταχείριση χωρίς διακρίσεις, στην ανθρώπινη αξιοπρέπεια ή σε πραγματικά και αποτελεσματικά μέσα έννομης προστασίας από μια τέτοια απόφαση.

¹⁷³ Χαρακτηριστικό παράδειγμα είναι το σύστημα SyRI (Systeem Risico Indicatie), το οποίο χρησιμοποιούταν για τον εντοπισμό λήψης/ληπτών κοινωνικών παροχών χωρίς να υπάρχει προς τούτο δικαίωμα. Η χρήση αυτού του συστήματος προκάλεσε, ωστόσο, μεγάλες αντιδράσεις. Το Περιφερειακό Δικαστήριο της Χάγης (District Court) έκρινε (5/2/2020) ότι το συγκεκριμένο σύστημα παραβίαζε το άρθρο 8 της ΕΣΔΑ.

6.4. Τα ζητήματα της διαφάνειας και της αιτιολογίας

Ένα από τα κρίσιμότερα ζητήματα που θέτει η υιοθέτηση συστημάτων Τεχνητής Νοημοσύνης αφορά την αδιαφάνεια. Η διαφάνεια της διοικητικής δράσης είναι βασική παράμετρος του κράτους δικαίου και της δημοκρατικής αρχής, στον βαθμό που επιτρέπει τον έλεγχο της δράσης των κρατικών οργάνων και Αρχών από τους πολίτες. Η διαδικασία λήψης μιας απόφασης θα πρέπει να είναι «γενικώς ορατή και κατανοητή», ώστε να διασφαλίζεται η αιτιολογία και κατ' επέκταση η ελεγχσιμότητα και η λογοδοσία, ιδίως όταν οι Αρχές ασκούν διακριτική ευχέρεια. Λόγω της εγγενούς αδιαφάνειας των συστημάτων ΤΝ, οι απαιτήσεις ενημέρωσης είναι ιδιαίτερα σημαντικές. Σε κάθε περίπτωση, φαίνεται ότι το πρώτο και λιγότερο αμφισβητούμενο βήμα/στάδιο είναι η ενημέρωση για την ύπαρξη ενός συστήματος ΤΝ και της επικείμενης χρήσης του, χωρίς να μπορούμε να διαβεβαιώσουμε ότι επαρκεί αυτού του είδους η ενημέρωση για να επιλύσει το μείζον ζήτημα της «αλγοριθμικής αδιαφάνειας», να ανοίξει δηλαδή το «μαύρο κουτί» (black box) μιας διοικητικής πράξης/απόφασης.¹⁷⁴ Θα μπορούσαμε να υποστηρίξουμε ότι μια εκτενής υποχρέωση ενημέρωσης για τα συστήματα αυτά, που προϋποθέτει, άλλωστε, την εποπτεία της διοίκησης επί της διαδικασίας, είναι καθοριστική και για την επίτευξη της διαφάνειας, αλλά και την προστασία των δικαιωμάτων των προσώπων. Καθώς είναι σχεδόν αδύνατον για έναν άνθρωπο να ανασυνθέσει και να κατανοήσει τις περίπλοκες λειτουργίες του αλγορίθμου μηχανικής μάθησης, κάθε προσπάθεια να διατυπώσει κανείς τους τρόπους για να παρατηρήσει ένα συγκεκριμένο προδιαγνωσθέν αποτέλεσμα υπόκειται σε σοβαρούς περιορισμούς. Είναι εξαιρετικά δύσκολο να μπορεί να επιτευχθεί γενική διαφάνεια ως προς τη λειτουργία του συστήματος, συμπεριλαμβανομένων των λόγων και των κριτηρίων που οδηγούν σε μια απόφαση και της κατανόησης των περιορισμών ως προς τα σύνολα δεδομένων που χρησιμοποιούνται για την εκμάθηση.

Η απαίτηση της αιτιολογίας σχετίζεται στενά με το ζήτημα της διαφάνειας και του «μαύρου κουτιού», δηλαδή την αδυναμία παρακολούθησης της

¹⁷⁴ Ειδικότερα Vorras, A., & Mitrou, L. (2021). Unboxing the Black Box of Artificial Intelligence: Algorithmic Transparency and/or a Right to Functional Explainability. In *EU Internet Law in the Digital Single Market* (p. 247 ff.). Springer. Επίσης, Βόρρας, Α. (2026). *Η εφαρμογή της αρχής της διαφάνειας στην Τεχνητή Νοημοσύνη* (σ. 23 επ.). Νομική Βιβλιοθήκη.

διαδικασίας της επεξεργασίας. Αυτό συμβαίνει ιδίως όταν μια απόφαση βασίζεται σε στατιστική αξιολόγηση και τα κριτήρια που έχουν εισαχθεί δεν είναι σαφή, ενώ ασαφής παραμένει και η βαρύτητά τους. Το πρόβλημα –αν και λιγότερο έντονο– υφίσταται και σε μερικώς αυτοματοποιημένες αποφάσεις, όταν έχουν «προετοιμαστεί» από μία εφαρμογή ΤΝ. Στην περίπτωση αυτήν, η δυνατότητα ουσιαστικής αιτιολογίας εξαρτάται από το εάν το πρόσωπο που αποφασίζει είναι σε θέση να παρακολουθήσει –έστω εκ των υστέρων– τη διαδικασία και να μπορεί να αιτιολογήσει την απόφαση.

Αξίζει να επισημανθεί ότι το Συμβούλιο της Επικρατείας με την απόφαση 1206/2024 απέρριψε τον ισχυρισμό της διοίκησης, με τον οποίο κατ' ουσίαν προβαλλόταν ότι «δεν νοείται αιτιολόγηση των ατομικών διοικητικών πράξεων οι οποίες εκδίδονται βάσει ηλεκτρονικής αυτοματοποιημένης επεξεργασίας δεδομένων, διότι δεν είναι δυνατή η αμφισβήτηση της επιστημονικής αξιοπιστίας της όλης διαδικασίας», κι επιβεβαίωσε ότι η αιτιολογία είναι συστατικό στοιχείο του κράτους δικαίου συναπτόμενη με τις αρχές της διαφάνειας και της νομιμότητας της διοικητικής δράσης, αλλά και της αποτελεσματικής δικαστικής προστασίας.

Η διαφάνεια και αιτιολογία (και αιτιολογησιμότητα) της χρήσης ΤΝ και των αποτελεσμάτων της, αλλά και οι άλλες αρχές και απαιτήσεις κατατείνουν και συγκλίνουν στην αρχή της λογοδοσίας. Η έννοια της λογοδοσίας δεν μπορεί να προσδιοριστεί μεν με ακρίβεια, ωστόσο εν προκειμένω εμπεριέχει τουλάχιστον τα ακόλουθα στοιχεία:

- α) υπευθυνότητα για τις πράξεις και τις επιλογές,
- β) δυνατότητα και βούληση εξήγησης των λόγων που οδηγούν σε αυτές και
- γ) ευθύνη υπό την έννοια της πρόβλεψης κυρώσεων.

Η λογοδοσία ενέχει και τη διάσταση της ευθύνης. Εν προκειμένω, διαφαίνεται ένα –καταρχάς– «κενό»: Αν οι πολίτες υπόκεινται στα αποτελέσματα ενός συστήματος ΤΝ που αυτοματοποιεί γνωστικές λειτουργίες και νομικές διεργασίες που προσιδιάζουν σε πρόσωπα, σε ποιον αποδίδεται η ευθύνη για το «αλγοριθμικό αποτέλεσμα»; Η πρόβλεψη σαφών και υποκείμενων σε καταλογισμό διαδικασιών ανθρώπινης παρέμβασης και αιτιολογίας είναι απαραίτητη για να ικανοποιηθεί η απαίτηση της λογοδοσίας από την πλευρά της διοίκησης.

Καθώς η εισαγωγή συστημάτων ΤΝ φαίνεται να οδηγεί σε έναν περιορισμό του δικαιώματος προηγούμενης ακρόασης, περιορισμός που επιτείνεται στην περίπτωση της απόλυτης αυτοματοποίησης, έχει υποστηριχθεί ότι πλήρως αυτοματοποιημένες αποφάσεις είναι επιτρεπτές μόνο εφόσον είναι προς όφελος του διοικουμένου. Η νομιμότητα της χρήσης συστημάτων, όπως

η αυτόματη επιβολή προστίμων σε περίπτωση τροχαίων παραβάσεων, τελεί υπό την αίρεση της διασφάλισης του δικαιώματος αμφισβήτησης της απόφασης.

Το κράτος δικαίου και η ποιότητά του συνδέονται με την ανταπόκριση σε μεταβαλλόμενες συνθήκες και ανάγκες. Καθώς η τεχνολογία και οι χρήσεις της αναπλάθουν την κοινωνία και αναδιατάσσουν τις σχέσεις που αναπτύσσονται, είναι αναμενόμενο να δοκιμάζεται ο τρόπος με τον οποίο προσλαμβάνουμε, αντιλαμβανόμαστε και συνακόλουθα ερμηνεύουμε και εφαρμόζουμε αρχές, όπως η διαφάνεια, η λογοδοσία, η προβλεψιμότητα, η συνεκτικότητα, η ισότητα. Ο νομοθέτης καλείται να διασφαλίσει την υιοθέτηση συγκεκριμένων απαιτήσεων και εγγυήσεων αναφορικά με την εξηγησιμότητα, τη διαφάνεια, τη δίκαιη μεταχείριση και τη λογοδοσία αναφορικά με τα συστήματα ΤΝ, αναλαμβάνοντας και το πρόσθετο θεσμικό καθήκον να εισάγει αποτελεσματικούς μηχανισμούς διακυβέρνησης αυτών των συστημάτων, ώστε να ελαχιστοποιήσει τους κινδύνους χωρίς να ανακόψει την –επωφελή για τη διοίκηση και κατ’ επέκταση για την κοινωνία– ανάπτυξή τους.

Η επίτευξη της νομιμοποίησης απαιτεί αξιολόγηση και στάθμιση των επιπτώσεων της υιοθέτησης αυτοματοποιημένων συστημάτων απόφασης. Η εισαγωγή των Εκτιμήσεων Αντικτύπου στα Θεμελιώδη Δικαιώματα (Fundamental Rights Impact Assessments – FRIAs) με το άρθρο 27 της AI Act επεκτείνει τη λογική της ρύθμισης βάσει κινδύνου σε ένα ευρύτερο σύνολο δικαιωμάτων, συμπεριλαμβανομένων της μη διάκρισης, της αξιοπρέπειας, της ελευθερίας της έκφρασης, αλλά και της ύπαρξης διαδικαστικών εγγυήσεων. Όπως επισημαίνουν η Άννα Θωμαΐδου και ο Κωνσταντίνος Λημνιώτης,¹⁷⁵ η έμφαση δίνεται στις επιπτώσεις σε κοινωνικό και δημοκρατικό επίπεδο, αντανακλώντας το γεγονός ότι τα συστήματα ΤΝ μπορούν να επηρεάσουν όχι μόνο ατομικά αποτελέσματα, αλλά και συλλογικές διαδικασίες λήψης αποφάσεων, την κοινωνική εμπιστοσύνη και τη δημόσια διακυβέρνηση. Ταυτόχρονα, η εκτίμηση των κινδύνων και των επιπτώσεων λειτουργεί ως κρίσιμος μηχανισμός για την εναρμόνιση μεταξύ τεχνικής καινοτομίας και προστασίας των θεμελιωδών δικαιωμάτων, διασφαλίζοντας ότι η ανάπτυξη και χρήση της ΤΝ είναι σύμφωνη τόσο με τις κανονιστικές υποχρεώσεις όσο και με τις δημοκρατικές αξίες.

Η πιθανότητα επέλευσης βλάβης πρέπει να αξιολογείται τόσο με βάση τεχνικούς παράγοντες –όπως η ακρίβεια του μοντέλου– όσο και με βάση κοινωνικούς παράγοντες –όπως η ευαλωτότητα των επηρεαζόμενων ομάδων.¹⁷⁶

¹⁷⁵ Thomaïdou, A., & Limniotis, K. (2025). Navigating through human rights in AI: exploring the interplay between GDPR and fundamental rights impact assessment. *Journal of Cybersecurity and Privacy*, 5(1), Article 7. <https://doi.org/10.3390/jcp5010007>

¹⁷⁶ Malgieri, G., & Santos, C. (2025). Assessing the (severity of) impacts on fundamental rights. *Computer Law & Security Review*, 56. <https://doi.org/10.1016/j.clsr.2025.106113>

Η Εκτίμηση Αντικτύπου για τα Θεμελιώδη Δικαιώματα απαιτεί διεπιστημονική εξειδίκευση, η οποία περιλαμβάνει νομικές, ηθικές και τεχνικές προσεγγίσεις. Η Εκτίμηση Αντικτύπου που εισάγεται, άλλωστε, με το άρθρο 27 της AI Act θα πρέπει να εκτείνεται στις επιπτώσεις τέτοιων συστημάτων στην κοινωνία, στις προσδοκίες και τις αντιλήψεις των πολιτών για τον τρόπο με τον οποίο η διοίκηση παρέχει υπηρεσίες με γνώμονα και την προστασία των δικαιωμάτων και των θεμελιακών αξιών της συνταγματικής τάξης.¹⁷⁷

¹⁷⁷ Εκτενέστερη ανάλυση στο Μήτρου, Λ. (2023). Μπορεί ο αλγόριθμος να διοικεί; Στο Λ. Μήτρου (Επιμ.), *Μπορεί ο αλγόριθμος...* (σ. 253 επ.). Πανεπιστημιακές Εκδόσεις Κρήτης.

6.5. Ο «ανθρώπινος παράγοντας» και η «μεροληψία αυτοματισμού»

Υποστηρίζεται ότι πολλοί από τους παραπάνω προβληματισμούς παρέλκουν, όταν ένα σύστημα Τεχνητής Νοημοσύνης χρησιμοποιείται όχι για τη λήψη αποφάσεων, αλλά για την υποστήριξη αυτών. Αξίζει να σημειωθεί, μάλιστα, ότι ορισμένα συστήματα ΤΝ, τα οποία καταρχήν εμπίπτουν στις περιπτώσεις χρήσης «υψηλού κινδύνου», μπορούν, ωστόσο, να εξαιρεθούν από την υπαγωγή στις απαιτήσεις που ισχύουν για τα συστήματα ΤΝ «υψηλού κινδύνου», εφόσον δεν συνεπάγονται σημαντικό κίνδυνο πρόκλησης βλάβης στην υγεία, στην ασφάλεια ή στα θεμελιώδη δικαιώματα των προσώπων. Η εξαίρεση αυτή υφίσταται, μεταξύ άλλων, όταν το σύστημα ΤΝ προορίζεται να επιτελέσει «ένα στενά διαδικαστικό έργο» ή «ένα προπαρασκευαστικό έργο».

Ο Οργανισμός Θεμελιωδών Δικαιωμάτων της Ευρωπαϊκής Ένωσης¹⁷⁸ επισημαίνει, ωστόσο, ότι θα πρέπει να υιοθετηθεί μια σαφής και περιοριστική ερμηνευτική προσέγγιση ως προς το «φίλτρο» της Πράξης για την Τεχνητή Νοημοσύνη. Μια τέτοια συστατική ερμηνεία κρίνεται, κατά τον FRA, αναγκαία, δεδομένου ότι η χρήση συστημάτων ΤΝ από τη διοίκηση είναι δυνατόν να επιφέρει ουσιώδεις επιπτώσεις στη ζωή των προσώπων, ενώ ακόμη και τεχνολογίες που εμφανίζονται εκ πρώτης όψεως ως απλώς υποστηρικτικές μπορούν, στην πράξη, να ασκήσουν καθοριστική επιρροή στη διαδικασία λήψης αποφάσεων. Πράγματι, συστήματα που χρησιμοποιούνται σε συγκεκριμένους τομείς, όπως οι διαδικασίες πρόσληψης προσωπικού ή η χορήγηση δημόσιων παροχών, ενσωματώνουν από τη φύση τους στοιχεία υψηλού κινδύνου.

Χωρίς αμφιβολία, η ένταση της «επιρροής» της ΤΝ φαίνεται και είναι καταρχάς περιορισμένη όταν η τελική απόφαση εναπόκειται στον ανθρώπινο παράγοντα και ο ρόλος του συστήματος ΤΝ εξαντλείται στη σύσταση, στην πρόταση κάποιου σχεδίου. Πόσο, όμως, ο άνθρωπος μπορεί να αποστασιοποιηθεί από τη σύσταση του συστήματος ΤΝ;

¹⁷⁸ Πρόκειται για τον οργανισμό της ΕΕ που παρέχει ανεξάρτητη εμπειρογνώμοσύνη και συμβουλές σχετικά με τα θεμελιώδη δικαιώματα στα θεσμικά όργανα της Ένωσης και τα κράτη-μέλη.

Όταν οι άνθρωποι αλληλεπιδρούν με συστήματα ΤΝ, δεν είναι σπάνιο το φαινόμενο της λεγόμενης μεροληψίας αυτοματισμού ή προκατάληψης υπέρ της αυτοματοποίησης (automation bias). Αναγόμενη στη Συμπεριφορική Ψυχολογία, η μεροληψία αυτοματισμού συνιστά την ανθρώπινη τάση να εμπιστεύεται κανείς άκριτα τις συστάσεις ή να ακολουθεί τις προτάσεις ενός αυτοματοποιημένου συστήματος. Η μεροληψία αυτοματισμού οδηγεί τον άνθρωπο να ακολουθήσει τη σύσταση, μη υποβάλλοντάς τη σε έλεγχο ώστε να εντοπίσει λάθη ή ελλείψεις. Ήδη μελέτες έχουν αναδείξει την αποτυχία των ανθρώπων να παρεμβαίνουν σε αποτελέσματα/«στοιχεία εξόδου» που έχουν παραχθεί από αλγοριθμικά συστήματα, είτε επειδή θεωρούν ότι το «σύστημα ξέρει καλύτερα» είτε λόγω κόπωσης και αδυναμίας διευθέτησης του όγκου της εργασίας μέσω της ανθρώπινης εργασίας.¹⁷⁹

Ο Έλληνας νομοθέτης φαίνεται να έχει επίγνωση των δυσκολιών που ανακύπτουν όταν οι άνθρωποι ενδέχεται να αποκλίνουν από αποφάσεις που προτείνονται από την ΤΝ, καθώς κάθε απόκλιση απαιτεί επαρκή αιτιολόγηση σε σχέση με τα «δεδομένα εισόδου», τους χρησιμοποιούμενους αλγορίθμους, την κατανόηση του πλαισίου, την ερμηνεία των κανονιστικών διατάξεων, καθώς και άλλες πτυχές που δεν αποτυπώνονται από την ΤΝ. Στην περίπτωση του Ελληνικού Κτηματολογίου, το αποφαινόμενο όργανο δεν υποχρεούται να αιτιολογεί οποιαδήποτε απόφαση αποκλίνει από τη σύσταση του συστήματος ΤΝ (άρθρο 11 του Ν. 5142/2024).

Το ζήτημα της επάρκειας της ανθρώπινης παρέμβασης ανακύπτει και στις περιπτώσεις που η Πράξη για την Τεχνητή Νοημοσύνη ανάγει την ανθρώπινη εποπτεία σε εγγύηση για την προστασία δικαιωμάτων και αντίστοιχα υποχρέωση των παρόχων συστημάτων ΤΝ «υψηλού κινδύνου» (άρθρο 14 της AIAct), τα οποία θα πρέπει να σχεδιάζονται και να αναπτύσσονται κατά τρόπο ώστε τα φυσικά πρόσωπα να μπορούν να επιβλέπουν τη λειτουργία τους και να διασφαλίζουν ότι χρησιμοποιούνται με τον προβλεπόμενο τρόπο, αλλά και ότι οι επιπτώσεις τους αντιμετωπίζονται στη διάρκεια του κύκλου ζωής του συστήματος. Για τον σκοπό αυτόν, ο πάροχος του συστήματος θα πρέπει να προσδιορίζει κατάλληλα μέτρα ανθρώπινης εποπτείας πριν από τη διάθεση του συστήματος στην αγορά ή τη θέση του σε λειτουργία (αιτιολογική σκέψη 73).

Ωστόσο, ο ενωσιακός νομοθέτης έχει επίγνωση του κινδύνου της «πιθανή[ς] τάση[ς] αυτόματης εξάρτησης ή υπέρμετρης εξάρτησης από τα “στοιχεία εξόδου” που παράγονται από σύστημα ΤΝ “υψηλού κινδύνου” (μεροληψία αυτοματισμού)». Για τον λόγο αυτόν απαιτείται όχι μόνο τα συστήματα ΤΝ να σχεδιάζονται κατά τρόπο ώστε να υπόκεινται σε ενσωματωμένους

¹⁷⁹ Lazcoz, G., & de Hert, P. (2023). Humans in the GDPR and AIA governance of automated and algorithmic systems. Essential pre-requisites against abdicating responsibilities. *Computer Law & Security Review*, 50. <https://doi.org/10.1016/j.clsr.2023.105833>

λειτουργικούς περιορισμούς, οι οποίοι δεν μπορούν να παρακαμφθούν από το ίδιο το σύστημα και διασφαλίζουν την ανταπόκρισή του στον άνθρωπο που το χειρίζεται, αλλά και τα φυσικά πρόσωπα στα οποία έχει ανατεθεί η ανθρώπινη εποπτεία να είναι σε θέση να ερμηνεύουν σωστά τα «στοιχεία εξόδου» του συστήματος. Επιπλέον, πρέπει να έχουν τη δυνατότητα, σε κάθε συγκεκριμένη περίπτωση, να αποφασίζουν τη μη χρήση του συστήματος ΤΝ «υψηλού κινδύνου» ή να αγνοούν, να παρακάμπτουν ή να αντιστρέφουν με οποιονδήποτε άλλον τρόπο τα «στοιχεία εξόδου» του.

Η ανθρώπινη εποπτεία ή παρέμβαση κινδυνεύει να εκφυλιστεί σε επίφαση ελέγχου εάν δεν υφίστανται οι προϋποθέσεις για ουσιαστική, επαρκή και διαρκή διαδικασία.¹⁸⁰ Αυτό σημαίνει ότι πρέπει να διασφαλίζεται ότι οι εποπτεύοντες άνθρωποι έχουν την «απαραίτητη επάρκεια, κατάρτιση, εξουσία και υποστήριξη» για την εκτέλεση αυτού του ρόλου. Αυτό περιλαμβάνει την εκπαίδευση των χρηστών, αλλά και τον εφοδιασμό τους με μεθοδολογίες και εργαλεία για την κριτική αξιολόγηση των αποτελεσμάτων της ΤΝ. Ειδικά για τη δημόσια διοίκηση αυτό σημαίνει ότι απαιτούνται –πέραν του γενικού γραμματισμού στην ΤΝ– ειδικά προγράμματα κατάρτισης.

¹⁸⁰ Haitsma, L., & Brink, B. (2025). From Human Intervention to Human Involvement: A Critical Examination of the Role of Humans in (Semi-) Automated Administrative Decision-Making. *Digital Government: Research and Practice*, 6(3), 1-17. <https://doi.org/10.1145/3716173>

6.6. Ζητήματα από τη χρήση Παραγωγικής Τεχνητής Νοημοσύνης από τη/στη διοίκηση

Το φαινόμενο της διαδεδομένης χρήσης της Παραγωγικής Τεχνητής Νοημοσύνης (Generative AI)¹⁸¹ αφορά και τη διοίκηση. Ήδη τους πρώτους μήνες της εκρηκτικής εμφάνισης του ChatGPT, η Ευρωπαϊκή Επιτροπή δήλωνε ότι «η εμφάνιση της Παραγωγικής ΤΝ σηματοδοτεί μια ριζική τομή: πρωτοφανείς δυνατότητες για την υποστήριξη του προσωπικού μας, τη μείωση του διοικητικού φόρτου και την ενίσχυση της ποιότητας και του αντικτύπου του έργου της Επιτροπής».¹⁸² Η Παραγωγική ΤΝ φαίνεται να προσφέρει ένα νέο εργαλείο για την αναζωογόνηση της δημόσιας διοίκησης, το οποίο έχει τη δυνατότητα να συμβάλει σε μια νέα καμπή του εγχειρήματος της «επανεφεύρεσης της διακυβέρνησης» (reinventing government) μέσω της τεχνολογίας.¹⁸³

Παρά τα διάφορα πλεονεκτήματά της, όπως η ανάλυση μεγάλου όγκου δεδομένων, ο αυτοματοποιημένος έλεγχος εγγράφων, η παραμετροποίηση σε συγκεκριμένες διοικητικές ανάγκες, η χρήση αυτή ενέχει κινδύνους. Οι κίνδυνοι αυτοί αφορούν την επαρκή αιτιολογία των διοικητικών πράξεων και αποφάσεων, την εξάρτηση από τη χρήση τέτοιων συστημάτων με αποτέλεσμα την υποβάθμιση των ικανοτήτων και δεξιοτήτων των υπαλλήλων, αλλά και ζητήματα που αφορούν την προστασία και την ασφάλεια δεδομένων, των προσωπικών δεδομένων συμπεριλαμβανομένων, και την προστασία απορρήτων και εν γένει της διανοητικής ιδιοκτησίας.

Λαμβάνοντας υπ' όψιν και τα προβλήματα που απορρέουν από λάθη, ανακρίβειες ή και τις «παραισθήσεις» (hallucinations) που χαρακτηρίζουν τέτοια συστήματα, θα πρέπει να υπάρξουν απαγορεύσεις, περιορισμοί και οδηγίες για τη σύννομη και υπεύνη χρήση τέτοιων συστημάτων. Οι

¹⁸¹ Πρόκειται για συστήματα ΤΝ που έχουν σχεδιαστεί και εκπαιδευτεί σε ευρείας κλίμακας δεδομένα και έχουν τη δυνατότητα να «μαθαίνουν» από τα δεδομένα εκπαίδευσής τους και να παράγουν πρωτότυπο περιεχόμενο εικόνας, ήχου, βίντεο ή άλλου τύπου με βάση συγκεκριμένα ερωτήματα, αιτήματα ή εντολές (prompts) που εισάγει στο σύστημα ο χρήστης. Υποκατηγορία των μοντέλων Παραγωγικής ΤΝ αποτελούν τα, πολύ διαδεδομένα στις μέρες μας, Μεγάλα Γλωσσικά Μοντέλα, τα οποία έχουν σχεδιαστεί και εκπαιδευτεί σε ευρείας κλίμακας δεδομένα, με σκοπό την ανάπτυξη ικανότητας κατανόησης και παραγωγής λόγου.

¹⁸² European Commission. (2024). *Communication to the Commission, artificial intelligence in the European Commission (AI@EC): A strategic vision to foster the development and use of lawful, safe and trustworthy Artificial Intelligence systems in the European Commission*. (C(2024) 380 final).

¹⁸³ Weerts, S. (2025). Generative AI in public administration in light of the regulatory awakening in the US and EU. *Cambridge Forum on AI: Law and Governance*, 1, e3, 1-19. <https://doi.org/10.1017/cfl.2024.10>

σχετικές απαιτήσεις, συστάσεις και οδηγίες θα πρέπει να απευθύνονται διακριτά και κατά περίπτωση τόσο στις διοικητικές Αρχές ως τέτοιες όσο και στους μεμονωμένους υπαλλήλους, καθώς δεν είναι σπάνιο το φαινόμενο να γίνεται μεμονωμένη χρήση των συστημάτων αυτών. Ειδική πρόνοια θα πρέπει να ληφθεί για τη χρήση Παραγωγικής Τεχνητής Νοημοσύνης α) από τις δημόσιες Αρχές και β) από μεμονωμένους δημόσιους υπαλλήλους και λειτουργούς.¹⁸⁴

¹⁸⁴ Η Ειδική Γραμματεία Τεχνητής Νοημοσύνης και Δεδομένων έχει προαναγγείλει την έκδοση ενός τέτοιου Οδηγού.

[ΠΩΣ] ΜΠΟΡΕΙ Η ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΝΑ ΕΙΝΑΙ
ΗΘΙΚΗ ΚΑΙ ΣΥΝΝΟΜΗ;
Ιούλιος 2026

Για μια ηθική Τεχνητή Νοημοσύνη στο πεδίο της Δικαιοσύνης



Παναγιώτης Περάκης

7.1. Σημείο εκκίνησης-Εισαγωγική αναγκαία επισήμανση

Οι σχέσεις ηθικής και Δικαιοσύνης είναι σχέσεις ιστορικά φορτισμένες. Στην πορεία του ανθρώπινου πολιτισμού προς έναν καλύτερο, δικαιότερο και δημοκρατικότερο κόσμο, οι απόψεις για τη Δικαιοσύνη ήταν πάντα μοιρασμένες σε δύο στρατόπεδα, με επίκεντρο την ηθική. Η υποκειμενικότητα (που αφήνει χώρο για αυθαιρεσία) και η μη δεσμευτικότητα των κανόνων υπήρξαν συνήθως ο κύριος στόχος της μίας πλευράς, ενώ το φυσικό Δίκαιο, το οποίο επικαλούνται αυτοί οι κανόνες, ο αντίλογος της άλλης. Τελικά, οι αντιλήψεις που επικράτησαν και οδήγησαν τις εξελίξεις προς τη σύγχρονη εποχή προτάσσουν την κοινή για όλους βεβαιότητα των κανόνων του θετικού Δικαίου έναντι αυτών ενός αβέβαιου και αμφισβητούμενου φυσικού Δικαίου. Ήτοι, χωρίς να παραλείπεται εντελώς η σημασία της αντιστοίχισης των εκάστοτε θεσπιζόμενων κανόνων προς τις βασικές ηθικές αρχές που πρέπει να διέπουν την κοινωνία μας, αυτό που αναγορεύει τον κανόνα δικαίου σε τέτοιο είναι τα εξωτερικά χαρακτηριστικά του, εμφανή και αναμφισβήτητα για όλους, δηλαδή η διαδικασία θέσπισής του σύμφωνα με τα προβλεπόμενα από τη δημοκρατική αρχή, και όχι από την ηθική του περιεχομένου του, η οποία είναι αμφισβητήσιμη ανάλογα με τους τόπους, τις εποχές και, κυρίως, την προσωπική θεώρηση και συγκρότηση του κάθε ανθρώπου, ο οποίος δικαιούται να αναπτύσσει ελεύθερα την προσωπικότητά του.

Η συζήτηση αυτή, που έδειχνε να έχει κλείσει οριστικά, τουλάχιστον στις δυτικές κοινωνίες, επανέρχεται ξανά σήμερα λόγω της Τεχνητής Νοημοσύνης, με αφορμή τις ανησυχίες και τα ερωτηματικά για το εύρος και τα όρια των δυνατοτήτων της. Είναι πραγματικά εντυπωσιακό ότι σε μια εποχή άκρατου σχετικισμού και πανσπερμίας απόψεων, η ανάγκη για μια ηθική ΤΝ, που θα υπηρετεί το γενικό καλό της ανθρωπότητας και θα λειτουργεί σύμφωνα με τις θεμελιώδεις αρχές που πρέπει να διέπουν την κοινωνία μας, είναι σχεδόν ομόφωνη.

Χρειάζεται, όμως, μια βασική επισήμανση, που ταυτίζεται και με τη θέση που η ΕΕ έλαβε στη σχετική ευρεία παγκόσμια συζήτηση: Ο στόχος για μια ηθική ΤΝ δεν αφορά στη θέσπιση *ηθικών κανόνων* που να τη διέπουν, διότι με αυτούς υπάρχει ο βάσιμος κίνδυνος, κυρίως λόγω της μη δεσμευτικότητάς τους, να έχουμε απλώς ένα ευχολόγιο, που στην πραγματικότητα

θα επιτρέπει την ασυδοσία στην ανάπτυξη και χρήση των εφαρμογών ΤΝ, ένα «*ethics washing*», όπως χαρακτηριστικά αναφέρεται. Η ανάγκη για τη λεγόμενη ηθική ΤΝ απαιτεί τη ρύθμισή της, με τη θέσπιση νομικά δεσμευτικών κανόνων, που να διασφαλίζουν με τρόπο επιτακτικό και ελέγξιμο, καθώς και με απειλή κυρώσεων σε περίπτωση παραβίασής τους, ότι θα είναι ηθική, δηλαδή ότι θα είναι σύμφωνη με συγκεκριμένες αρχές, εις όφελος της ανθρωπότητας. Ακόμη κι αν υπάρχουν επιφυλάξεις σε σχέση με την παραπάνω θέση, κυρίως από την πλευρά της οικονομίας ή της γεωστρατηγικής, με βασικό επιχείρημα ότι η θέσπιση ανελαστικού νομικού πλαισίου και λεπτομερειακών («γραφειοκρατικών» για τους επικριτές της) διαδικασιών δημιουργεί ανταγωνιστικό μειονέκτημα για την Ευρώπη, η θέση αυτή ισχύει απόλυτα για τη Δικαιοσύνη, αφού η απονομή της δεν μπορεί να επαφίεται σε ηθικούς κανόνες, αλλά οφείλει να λειτουργεί με βάση προκαθορισμένους και μη αμφισβητήσιμους νομικούς κανόνες, γνωστούς εκ των προτέρων και δεσμευτικούς για όλους.

7.2. Οι βασικές αρχές για μια ηθική Τεχνητή Νοημοσύνη στο πεδίο της Δικαιοσύνης

Το πρώτο διεθνές δεσμευτικό νομικό κείμενο για την Τεχνητή Νοημοσύνη είναι η Σύμβαση-Πλαίσιο του Συμβουλίου της Ευρώπης για την Τεχνητή Νοημοσύνη, τα Ανθρώπινα Δικαιώματα, τη Δημοκρατία και το Κράτος Δικαίου (2024). Εκεί ορίζονται οι παρακάτω γενικές αρχές για την ΤΝ: ανθρώπινη αξιοπρέπεια & προσωπική αυτονομία, διαφάνεια & εποπτεία, λογοδοσία & ευθύνη, ισότητα & μη διακριτική μεταχείριση, προστασία ιδιωτικότητας & προσωπικών δεδομένων, αξιοπιστία & ασφαλής καινοτομία.

Λίγο μετά ακολούθησε η ΕΕ και ο Κανονισμός ΕΕ 2024/1689 (γνωστός ως Πράξη για την Τεχνητή Νοημοσύνη ή AI Act), που όρισε ως θεμελιώδη αρχή η ΤΝ να είναι ανθρωποκεντρική, καθώς και ότι η ανάπτυξή της πρέπει να γίνει σύμφωνα με τις αξίες της Ένωσης και να χρησιμεύει ως εργαλείο με στόχο την αύξηση της ανθρώπινης ευημερίας.¹⁸⁵ Κατά τα λοιπά, παραπέμπει στις επτά αρχές που από το 2019 έχει διατυπώσει η Ομάδα Εμπειρογνομόνων Υψηλού Επιπέδου για την Τεχνητή Νοημοσύνη στις Κατευθυντήριες Γραμμές Δεοντολογίας για Αξιοπιστη Τεχνητή Νοημοσύνη,¹⁸⁶ με στόχο τη διασφάλιση της αξιοπιστίας και της δεοντολογικής ορθότητας της ΤΝ. Αυτές είναι:

- α) ανθρώπινη παρέμβαση και εποπτεία,
- β) τεχνική ευρωστία και ασφάλεια,
- γ) ιδιωτική ζωή και διακυβέρνηση των δεδομένων,
- δ) διαφάνεια,
- ε) ποικιλομορφία, απαγόρευση των διακρίσεων και δικαιοσύνη,

¹⁸⁵ Βλ. Προοίμιο, αιτιολογική σκέψη 6: «Λόγω του σημαντικού αντικτύπου που μπορεί να έχει η ΤΝ στην κοινωνία και δεδομένης της ανάγκης οικοδόμησης εμπιστοσύνης, έχει ζωτική σημασία η ΤΝ και το κανονιστικό της πλαίσιο να αναπτυχθούν με τρόπο σύμφωνο προς τις αξίες της Ένωσης, όπως κατοχυρώνονται στο άρθρο 2 της Συνθήκης για την Ευρωπαϊκή Ένωση (ΣΕΕ), προς τα θεμελιώδη δικαιώματα και τις ελευθερίες που κατοχυρώνονται στις Συνθήκες, και, σύμφωνα με το άρθρο 6 ΣΕΕ, προς τον Χάρτη. Ως προαπαιτούμενο, η ΤΝ θα πρέπει να είναι μια τεχνολογία ανθρωποκεντρική. Θα πρέπει να χρησιμεύει σαν εργαλείο για τους ανθρώπους, με απώτατο στόχο την αύξηση της ανθρώπινης ευημερίας».

¹⁸⁶ High-Level Expert Group on Artificial Intelligence. (2019). *Ethics guidelines for trustworthy AI*. European Commission. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

στ) κοινωνική και περιβαλλοντική ευημερία και

ζ) λογοδοσία.

Οι παραπάνω αρχές ισχύουν για όλες τις εφαρμογές ΤΝ, συνεπώς και στη Δικαιοσύνη.

Περαιτέρω, τα συστήματα ΤΝ που προορίζονται να χρησιμοποιηθούν για την απονομή της Δικαιοσύνης, λόγω του δυνητικά σημαντικού αντικτύπου τους στη δημοκρατία, στο κράτος δικαίου, στις ατομικές ελευθερίες, καθώς και στο δικαίωμα πραγματικής προσφυγής και αμερόληπτου δικαστηρίου, κατατάσσονται από την AI Act στην κατηγορία «υψηλού κινδύνου»,¹⁸⁷ ενώ ξεκαθαρίζεται ότι «η χρήση εργαλείων ΤΝ μπορεί να στηρίξει την εξουσία λήψης αποφάσεων των δικαστών ή τη δικαστική ανεξαρτησία, αλλά δεν θα πρέπει να την υποκαταστήσει: Η τελική λήψη αποφάσεων πρέπει να παραμείνει δραστηριότητα κατευθυνόμενη από τον άνθρωπο» (βλ. Προοίμιο, αιτιολογική σκέψη 61). Συνεπώς, επιπροσθέτως των όσων προαναφέρθηκαν, για τα συστήματα ΤΝ που προορίζονται για τη Δικαιοσύνη ισχύουν όλες οι απαιτήσεις που προβλέπονται για τα συστήματα «υψηλού κινδύνου», για τις οποίες σκόπιμο είναι να λαμβάνονται υπ' όψιν οι πρόσφατες (Δεκέμβριος 2025) συστάσεις του Οργανισμού Θεμελιωδών Δικαιωμάτων της Ευρωπαϊκής Ένωσης με τίτλο «Αξιολόγηση Τεχνητής Νοημοσύνης Υψηλού Κινδύνου: Κίνδυνοι για τα Θεμελιώδη Δικαιώματα» (Assessing High-risk Artificial Intelligence: Fundamental Rights Risks), που αποτελούν συνέχεια προηγούμενων εκθέσεων του FRA για την ΤΝ. Είναι προφανές ότι η εφαρμογή όλων των σχετικών απαιτήσεων στη Δικαιοσύνη επιβάλλει συγκεκριμένο επιμερισμό της ευθύνης, που, για να μην υπάρχουν αμφιβολίες, κενά ή αλληλοκαλύψεις, πρέπει να γίνεται με τρόπο νομικά δεσμευτικό.

Επιπλέον, δε, αυτών, για κάθε σύστημα που προορίζεται για τη Δικαιοσύνη και χαρακτηρίζεται «υψηλού κινδύνου», πρέπει, πριν από την ανάπτυξή του, να διενεργείται εκτίμηση των επιπτώσεων που η χρήση του μπορεί να έχει στα θεμελιώδη δικαιώματα.¹⁸⁸

¹⁸⁷ Συγκεκριμένα, η Πράξη (βλ. άρθρο 6 παρ. 2 και Παράρτημα III, αρ. 8α) ορίζει ως «υψηλού κινδύνου» τα «συστήματα ΤΝ που προορίζονται να χρησιμοποιηθούν από δικαστική Αρχή ή για λογαριασμό της για τη συνδρομή προς δικαστική Αρχή κατά την έρευνα και την ερμηνεία των πραγματικών περιστατικών και του νόμου και κατά την εφαρμογή του νόμου σε συγκεκριμένο σύνολο πραγματικών περιστατικών ή να χρησιμοποιηθούν με παρεμφερή τρόπο σε εναλλακτική επίλυση διαφορών». Το ποια συστήματα ακριβώς εντάσσονται στην παραπάνω πρόβλεψη και ποια όχι δεν είναι καθόλου προφανές και είναι βέβαιο ότι θα απαιτηθούν περαιτέρω εξειδικεύσεις και εξέταση ad hoc κάθε φορά. Συναφώς στο Προοίμιο, αιτιολογική σκέψη 61 αναφέρεται: «Ο χαρακτηρισμός συστημάτων ΤΝ ως “υψηλού κινδύνου” δεν θα πρέπει, ωστόσο, να επεκταθεί σε συστήματα ΤΝ που προορίζονται για αμιγώς βοηθητικές διοικητικές δραστηριότητες, οι οποίες δεν επηρεάζουν στην πράξη την απονομή της Δικαιοσύνης σε μεμονωμένες περιπτώσεις, όπως η ανωνυμοποίηση ή η ψευδωνυμοποίηση δικαστικών αποφάσεων, εγγράφων ή δεδομένων, η επικοινωνία μεταξύ των μελών του προσωπικού ή η εκτέλεση διοικητικών καθηκόντων».

¹⁸⁸ Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), art. 27 (Fundamental Rights Impact Assessment for High-Risk AI Systems). Official Journal of the European Union, L, 2024/1689.

Τέλος, κομβική σημασία για το μέλλον και για οποιαδήποτε τυχόν χρήση της ΤΝ στη Δικαιοσύνη έχει αναμφίβολα η διάταξη του άρθρου 86 της AI Act («Δικαίωμα επεξήγησης της ατομικής λήψης αποφάσεων»). Κατοχυρώνει το δικαίωμα κάθε θιγόμενου προσώπου που υπόκειται σε απόφαση, η οποία λαμβάνεται με βάση τα «στοιχεία εξόδου» συστήματος ΤΝ «υψηλού κινδύνου» και παράγει έννομα αποτελέσματα ή επηρεάζει σημαντικά το εν λόγω πρόσωπο με παρόμοιο τρόπο, ώστε να θεωρεί ότι υφίσταται δυσμενείς συνέπειες για την υγεία, την ασφάλεια και τα θεμελιώδη δικαιώματά του, να λάβει σαφείς και ουσιαστικές επεξηγήσεις σχετικά με τον ρόλο του συστήματος ΤΝ στη διαδικασία λήψης αποφάσεων, καθώς και τα κύρια στοιχεία της ληφθείσας απόφασης.

Αρχές για μια ηθική ΤΝ ειδικά στο πεδίο της Δικαιοσύνης έχουν διατυπώσει πολλοί διεθνείς φορείς, χωρίς ουσιαστικές αποκλίσεις από όσα προαναφέρθηκαν. Ιδιαίτερης βαρύτητας είναι οι θέσεις της Ευρωπαϊκής Επιτροπής για την Αποτελεσματικότητα της Δικαιοσύνης (European Commission for the Efficiency of Justice – CEPEJ) του Συμβουλίου της Ευρώπης, που τον Δεκέμβριο του 2018 διατύπωσε τις ακόλουθες πέντε αρχές για τη χρήση της ΤΝ στα δικαστικά συστήματα.¹⁸⁹

α) Αρχή του σεβασμού των θεμελιωδών δικαιωμάτων (διασφάλιση ότι ο σχεδιασμός και η εφαρμογή εργαλείων και υπηρεσιών ΤΝ είναι συμβατά με τα θεμελιώδη δικαιώματα).

β) Αρχή της μη διακριτικής μεταχείρισης (πρόληψη της ανάπτυξης ή της ενίσχυσης οποιασδήποτε διακριτικής μεταχείρισης μεταξύ ατόμων ή ομάδων ατόμων).

γ) Αρχή της ποιότητας και της ασφάλειας (επεξεργασία δικαστικών αποφάσεων και δεδομένων με χρήση πιστοποιημένων πηγών και άυλων δεδομένων, με βάση μοντέλα που έχουν δημιουργηθεί με διεπιστημονικό τρόπο μέσα σε ένα ασφαλές τεχνολογικό περιβάλλον).

δ) Αρχή της διαφάνειας, της αμεροληψίας και της δίκαιης μεταχείρισης (καθιέρωση προσβάσιμων και κατανοητών μεθόδων επεξεργασίας δεδομένων, επιτρεπτό εξωτερικών ελέγχων).

ε) Αρχή του «ελέγχου από τον χρήστη» (απαγόρευση μιας τυποποιημένης προσέγγισης και διασφάλιση ότι οι χρήστες είναι ενημερωμένοι και έχουν τον έλεγχο των επιλογών τους).

¹⁸⁹ Ευρωπαϊκός Χάρτης Δεοντολογίας σχετικά με τη χρήση Τεχνητής Νοημοσύνης (ΤΝ) στα δικαστικά συστήματα και το περιβάλλον τους – European Ethical Charter on the use of artificial intelligence (AI) in judicial systems and their environment.

Σύμφωνα με την CEPEJ, η εφαρμογή της ΤΝ στον τομέα της Δικαιοσύνης μπορεί να συμβάλει στη βελτίωση της αποτελεσματικότητας και της ποιότητάς της, πρέπει όμως να εφαρμόζεται με υπεύθυνο τρόπο, ο οποίος να συνάδει με τα θεμελιώδη δικαιώματα που κατοχυρώνονται ιδίως στην ΕΣΔΑ και στη Σύμβαση του Συμβουλίου της Ευρώπης για την Προστασία των Δεδομένων Προσωπικού Χαρακτήρα. Για την CEPEJ, είναι απαραίτητο να διασφαλιστεί ότι η ΤΝ θα παραμένει ένα εργαλείο στην υπηρεσία του γενικού συμφέροντος και ότι η χρήση της θα σέβεται τα ατομικά δικαιώματα.¹⁹⁰

Εκτός από την CEPEJ, σημαντική δραστηριότητα για μια ηθική ΤΝ στο πεδίο της Δικαιοσύνης έχει αναπτύξει η UNESCO, με σπουδαιότερη την πρόσφατη δημοσίευση (Δεκέμβριος του 2025) Κατευθυντήριων Γραμμών για τη χρήση συστημάτων ΤΝ σε δικαστήρια (Guidelines for the use of AI systems in courts and tribunals).¹⁹¹ Με βασική θέση ότι η ΤΝ μπορεί να βοηθήσει στην ταχύτερη και δικαιότερη απονομή της Δικαιοσύνης, η οποία, όμως, πάντα πρέπει να βασίζεται στην ανθρώπινη κρίση, δηλαδή να παραμένει ανθρωποκεντρική, προτείνονται οι παρακάτω αρχές για την ανάπτυξη, απόκτηση και χρήση συστημάτων ΤΝ στο δικαστικό σύστημα:

- 1) Προστασία των ανθρωπίνων δικαιωμάτων (με ειδική αναφορά σε κάποια ανθρώπινα δικαιώματα που πλήττονται περισσότερο από την ανάπτυξη της ΤΝ, ιδίως αυτά που σχετίζονται με τα πλέον ευάλωτα άτομα και το δικαίωμα μη διακρίσεων).
- 2) Αναλογικότητα (υιοθέτηση, ανάπτυξη και χρήση συστημάτων ΤΝ για σκοπούς αναλογικούς με το πλαίσιο της χρήσης τους).
- 3) Σκοπιμότητα των οφελών (προκαταβολική αξιολόγηση των πιθανών οφελών σε σχέση με την ικανότητα υλοποίησης).
- 4) Ασφάλεια (για την πρόληψη βλαβών σε πρόσωπα και στη φήμη και την εξουσία της Δικαιοσύνης).
- 5) Ασφάλεια πληροφοριών (προστασία εμπιστευτικών πληροφοριών, μέτρα κατά των απειλών στον κυβερνοχώρο και διασφάλιση του συνεχούς ελέγχου των δεδομένων).
- 6) Ακρίβεια και αξιοπιστία (ακριβή και αξιόπιστα συστήματα ΤΝ).

¹⁹⁰ Για κριτικό σχολιασμό του Χάρτη, καθώς και, γενικότερα, της χρήσης της ΤΝ στη Δικαιοσύνη, βλ. Σεβαστίδης, Χ., Ντόκα, Α., & Ξυλιά, Ι. (2020, 17 Αυγούστου). *Τεχνητή νοημοσύνη στη δικαιοσύνη: Πρόσδος ή προαναγγελία ενός δυστοπικού μέλλοντος*; Ένωση Δικαστών και Εισαγγελέων.

¹⁹¹ Stankovich, M., Feldfeber, I., Quiroga, Y., Ciolfi Felice, M., & Marivate, V. (2023). *Global toolkit on AI and the rule of law for the judiciary*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000387331_eng

7) *Επεξηγησιμότητα* (υιοθέτηση, ανάπτυξη και χρήση συστημάτων ΤΝ που να μπορούν να εξηγήσουν τη λογική πίσω από τα αποτελέσματα και τις αποφάσεις τους, καθώς και τον τρόπο με τον οποίο χρησιμοποιούν τα «δεδομένα εισόδου» τους).

8) *Ελεγκσιμότητα* (λήψη μέτρων για να διασφαλίζεται η επιθεώρηση των συστημάτων).

9) *Διαφάνης και ανοιχτή Δικαιοσύνη* (διαφάνεια ως προς τον τρόπο ανάπτυξης του συστήματος, τον τρόπο λειτουργίας του, τα δεδομένα εκπαίδευσής του, τους περιορισμούς του –συμπεριλαμβανομένου του περιθωρίου σφάλματος–, τις δυνατότητες και τον σκοπό του).

10) *Ευαισθητοποίηση και ενημερωμένη χρήση* (κατανόηση λειτουργιών, τύπων χρήσεων, πιθανών επιπτώσεων, περιορισμών και κινδύνων, και ενημέρωση για τις τεχνολογικές εξελίξεις).

11) *Ευθύνη* (η ανάπτυξη και χρήση συστημάτων ΤΝ συνεπάγεται αντίστοιχη ευθύνη για τις αποφάσεις και λοιπές ενέργειες που λαμβάνουν χώρα με την υποστήριξή τους, με επιφύλαξη της πιθανής ευθύνης του παρόχου σε περίπτωση ελαττώματος).

12) *Λογοδοσία και δυνατότητα αμφισβήτησης* (διασφάλιση της λογοδοσίας με ενημέρωση για τα αποτελέσματα που επιτεύχθηκαν ή μη με τη χρήση συστημάτων ΤΝ και εισαγωγή μηχανισμών που να επιτρέπουν στα εμπλεκόμενα μέρη να αμφισβητούν και να εξετάζουν αποτελέσματα που παράγονται από ένα σύστημα ΤΝ και να αμφισβητούν τις αποφάσεις που στηρίζονται σε αυτό, θεωρώντας τους δικαστικούς λειτουργούς ως υπόλογους για τα τυχόν λάθη).

13) *Ανθρώπινη εποπτεία και λήψη αποφάσεων* (δεν επιτρέπεται η εξουσία του δικαστή να μεταβιβάζεται σε συστήματα ΤΝ, ούτε οι αποφάσεις του να βασίζονται αποκλειστικά σε αυτά, ενώ διαδικασίες που ενδέχεται να επηρεάσουν αρνητικά τα δικαιώματα ατόμων ή κοινοτήτων δεν πρέπει αυτοματοποιούνται εξ ολοκλήρου. Αντιθέτως, απαιτείται διατήρηση ανθρώπινης παρέμβασης σε όλες τις διαδικασίες λήψης αποφάσεων, ώστε τα εργαλεία ΤΝ να μην αντικαθιστούν την ανεξάρτητη ανάλυση γεγονότων, νόμων και αποδεικτικών στοιχείων από τον υπεύθυνο λήψης των αποφάσεων).

14) *Σχεδιασμός με επίκεντρο τον άνθρωπο και συμμετοχή* (ανθρωποκεντρικός σχεδιασμός, που πρέπει να περιλαμβάνει μέλη της δικαστικής εξουσίας, αλλά και το κοινό).

15) Πολυμερής διακυβέρνηση και συνεργασία.

Σημαντική, τέλος, και διαρκής παρουσία στη συζήτηση για μια ηθική ΤΝ στον χώρο της Δικαιοσύνης έχει εξ αρχής το Συμβούλιο των Δικηγορικών Συλλόγων της Ευρώπης (Council of Bars and Law Societies of Europe – CCBE),¹⁹² που έχει διατυπώσει αναλυτικές θέσεις για την εισαγωγή της ΤΝ στην καθημερινή νομική πρακτική, για την επίδρασή της στα ανθρώπινα δικαιώματα και για τη χρήση της από τα δικαστήρια (με ιδιαίτερες επιφυλάξεις για τη χρήση ΤΝ στην απονομή της Ποινικής Δικαιοσύνης¹⁹³), καθώς και συστάσεις για τη χρήση των εργαλείων ΤΝ από τους δικηγόρους, όπως και για τις υποχρεώσεις τους.

Για τη χρήση της ΤΝ από τα δικαστήρια, βασική θέση είναι ότι τα εργαλεία ΤΝ πρέπει να προσαρμόζονται στο ιδιαίτερο περιβάλλον της Δικαιοσύνης, λαμβάνοντας υπ' όψιν τις αρχές και τη διαδικαστική αρχιτεκτονική που διέπουν τις δικαστικές διαδικασίες, με επίκεντρο τη μη ανάθεση σε συστήματα ΤΝ της εξουσίας λήψης αποφάσεων του δικαστή. Περαιτέρω, σύμφωνα με το CCBE, θα πρέπει να διασφαλίζονται:

- α) η δυνατότητα επαλήθευσης της εισαγωγής δεδομένων και της συλλογιστικής του εργαλείου ΤΝ,
- β) η δυνατότητα αμφισβήτησης των αποτελεσμάτων της ΤΝ,
- γ) η ουδετερότητα και η αντικειμενικότητα των εργαλείων ΤΝ που χρησιμοποιούνται από το δικαστικό σύστημα, που θα πρέπει και αυτές να είναι εγγυημένες και επαληθεύσιμες,
- δ) η συμμόρφωση με τις αρχές του Γενικού Κανονισμού Προστασίας Δεδομένων.

Από όσα έχουν προεκτεθεί, γίνεται αντιληπτό ότι για τον χώρο της Δικαιοσύνης υπάρχει ήδη ένα αρκετά συνεκτικό πλαίσιο αρχών με στόχο τη διασφάλιση ηθικής ΤΝ. Η σχετική συζήτηση, βεβαίως, συνεχίζεται σε διεθνές και σε εθνικό επίπεδο. Ιδιαίτερη αναφορά αξίζει στην CEPEJ, που, εκτός των άλλων, εξέδωσε πρόσφατα (19.12.2025) Κατευθυντήριες Γραμμές για τη Χρήση της Παραγωγικής Τεχνητής Νοημοσύνης από τα Δικαστήρια

¹⁹² Το Συμβούλιο των Δικηγορικών Συλλόγων της Ευρώπης (CCBE) θεωρείται ο επίσημος ευρωπαϊκός φορέας των δικηγόρων, εκπροσωπώντας τους εθνικούς δικηγορικούς συλλόγους 45 χωρών και μέσω αυτών περισσότερους από 1,1 εκατ. Ευρωπαίους δικηγόρους.

¹⁹³ Council of Bars and Law Societies of Europe (CCBE) (2020). *CCBE considerations on the legal aspects of Artificial Intelligence*. https://www.ccbe.eu/fileadmin/speciality_distribution/public/documents/IT_LAW/ITL_Guides_recommendations/EN_ITL_20200220_CCBE-considerations-on-the-Legal-Aspects-of-AI.pdf

(Guidelines on the use of Generative Artificial Intelligence for Courts).¹⁹⁴

Σε αυτές, αφού διαπιστώνεται η αλματώδης αύξηση της χρήσης ΤΝ στον τομέα της Δικαιοσύνης (αναφέρεται ότι σε σχετική έρευνα το 46% των ερωτηθέντων επιβεβαίωσε τη χρήση τέτοιων εργαλείων σε δικαστήρια), αλλά και τα διόλου ασήμαντα προβλήματα που έχουν εντοπισθεί, προσδιορίζονται συγκεκριμένοι κίνδυνοι και γίνονται αντίστοιχες επισημάνσεις και συστάσεις, με βασική κι εδώ θέση ότι η διαδικασία λήψης των αποφάσεων πρέπει αυστηρά να παραμένει ανθρώπινη. Μεταξύ των συστάσεων είναι, αντί συστημάτων γενικής χρήσης που βρίσκονται στην αγορά, για τη Δικαιοσύνη να προτιμώνται εξειδικευμένα συστήματα που θα δημιουργούνται επί τούτου (customized AI tools), βάσει αναλυτικών προδιαγραφών και με αυστηρές προβλέψεις για τη διακυβέρνηση των δεδομένων, η ανάπτυξη τέτοιων συστημάτων να γίνεται με στρατηγικό σχεδιασμό, σε φάσεις (*step-by-step approach*), με συνεχή παρακολούθηση της απόδοσής τους και ανάλογη επικαιροποίηση, καθώς και η ανάγκη διαρκούς εκπαίδευσης. Συστήνεται ακόμη η θέσπιση κανόνων που θα επιβάλλουν την υποχρεωτική αποκάλυψη της τυχόν χρήσης εργαλείων ΤΝ στο δικαστήριο (βλ. αναλυτικότερα για αυτό παρακάτω, στην υπο-ενότητα 7.4.).

¹⁹⁴ Στη συνεχιζόμενη σήμερα (Μάρτιος 2026) συστηματική δουλειά της CEPEJ στο αντικείμενο συμπεριλαμβάνονται η εκπόνηση Κατευθυντήριων Γραμμών για τη Δικαστική Ανεξαρτησία, καθώς και για τις αυτοματοποιημένες διοικητικές διαδικασίες στη Δικαιοσύνη.

7.3. Η χρήση Τεχνητής Νοημοσύνης στην Ποινική Δικαιοσύνη

Ξεχωριστό ενδιαφέρον παρουσιάζει το ζήτημα της χρήσης εργαλείων Τεχνητής Νοημοσύνης στην Ποινική Δικαιοσύνη. Για το θέμα αυτό, στο Προοίμιο της AI Act (αιτιολογική σκέψη 42) υπάρχει η εξής σημαντική αναφορά, που θέτει σαφή όρια: «Σύμφωνα με το τεκμήριο της αθωότητας, τα φυσικά πρόσωπα στην Ένωση θα πρέπει πάντα να κρίνονται με βάση την πραγματική συμπεριφορά τους. **Τα φυσικά πρόσωπα δεν θα πρέπει ποτέ να κρίνονται με βάση την προβλεπόμενη από την ΤΝ συμπεριφορά, που βασίζεται αποκλειστικά στην κατάρτιση προφίλ για αυτά ή σε γνωρίσματα ή χαρακτηριστικά της προσωπικότητάς τους, όπως ιθαγένεια, τόπος γέννησης, τόπος διαμονής, αριθμός παιδιών, ύψος χρέους ή είδος αυτοκινήτου, χωρίς εύλογες υπόνοιες εμπλοκής του εν λόγω προσώπου σε εγκληματική δραστηριότητα βάσει αντικειμενικών και επαληθεύσιμων γεγονότων και χωρίς ανθρώπινη αξιολόγησή τους.** Ως εκ τούτου, θα πρέπει να απαγορεύονται οι εκτιμήσεις κινδύνου που διενεργούνται για φυσικά πρόσωπα, προκειμένου να εκτιμηθεί η πιθανότητα να τελέσουν αξιόποινη πράξη ή να προβλεφθεί η τέλεση πραγματικής ή δυνητικής αξιόποινης πράξης με βάση αποκλειστικά την κατάρτιση προφίλ για αυτά ή την αξιολόγηση γνωρισμάτων και χαρακτηριστικών της προσωπικότητάς τους».

Με το θέμα ασχολείται και η τελευταία (2025) έκθεση της ειδικής εισηγήτριας του ΟΗΕ για την ανεξαρτησία δικαστών και δικηγόρων, με τον εύγλωττο τίτλο «ΤΝ στο δικαστικό σύστημα: υποσχέσεις και παγίδες» και με βασική θέση ότι τα ανθρώπινα δικαιώματα μπορούν να προστατεύονται μόνο από ανθρώπους και ειδικότερα από ανθρώπους δικαστές. Για τη χρήση της ΤΝ στο πλαίσιο της Ποινικής Δικαιοσύνης διατυπώνονται πολλές επιφυλάξεις (ιδίως λόγω προκαταλήψεων, κινδύνων για την ισότητα των όπλων, καθώς και έλλειψης επεξηγησιμότητας και διαφάνειας) και οι ακόλουθες συστάσεις:

α) Οι εισαγγελείς θα πρέπει να αποφεύγουν να βασίζονται σε στοιχεία από ΤΝ ως βάση για την επιδίωξη καταδικών, εκτός αν είναι βέβαιοι ότι έχουν ελεγχθεί αυστηρά, δεν είναι διακριτικά και μπορούν να αμφισβητηθούν ουσιαστικά από την υπεράσπιση έπειτα από υποχρεωτική γνωστοποίησή τους.

β) Οι προβλεπτικές αναλύσεις δεν θα πρέπει ποτέ να χρησιμοποιούνται ως βάση για δίωξη ή κράτηση.

γ) Οι δικαστές και οι εισαγγελείς θα πρέπει να ενεργούν ως φρουροί (gatekeepers) και να υποβάλλουν και να αποδέχονται σε ποινικές δίκες μόνο αξιόπιστα στοιχεία που έχουν παραχθεί από ΤΝ.

δ) Οι κανόνες αποδεικτικών μέσων και αποδοχής τους δεν θα πρέπει να κάνουν εξαιρέσεις για τα εργαλεία ΤΝ από όσα γενικώς προβλέπονται ως προς το επιτρεπτό ιδιωτικών αποδεικτικών μέσων.

Για το θέμα της χρήσης εργαλείων ΤΝ στην Ποινική Δικαιοσύνη επιφυλάξεις έχουν διατυπωθεί από πολλές ακόμη πλευρές. Η θέση του CCBE είναι ότι χρειάζεται προσεκτική μελέτη και αξιολόγηση πριν από την εισαγωγή τους. Τονίζεται, δε, ότι, παρόλο που η πρόληψη της διάπραξης εγκλημάτων, η αύξηση του ποσοστού εξιχνιάσής τους και η βελτίωση της ποιότητας των ποινικών αποφάσεων είναι ασφαλώς στόχοι θεμιτοί για το κράτος δικαίου, οι κίνδυνοι προκαταλήψεων και διακρίσεων εις βάρος συγκεκριμένων ομάδων είναι υψηλοί. Η απειλή μαζικής παρακολούθησης από συστήματα ΤΝ αποτελεί, επίσης, υπαρκτό κίνδυνο. Συνεπώς, εφαρμογές ΤΝ θα πρέπει να εισάγονται στην Ποινική Δικαιοσύνη μόνον όταν υπάρχουν επαρκείς εγγυήσεις κατά οποιασδήποτε μορφής προκατάληψης ή διάκρισης. Όλα, δε, τα μέτρα αυξημένης επιτήρησης θα πρέπει να αξιολογούνται προσεκτικά ως προς τον αντίκτυπο που μπορεί να έχουν σε μια ανοιχτή και πλουραλιστική κοινωνία.

7.4. Το ζήτημα της υποχρεωτικής αποκάλυψης της χρήσης Τεχνητής Νοημοσύνης σε δικαστήριο

Με αιτιολόγηση τη διασφάλιση της διαφάνειας και της λογοδοσίας στη χρήση εργαλείων Τεχνητής Νοημοσύνης στις δικαστικές διαδικασίες, από ορισμένες πλευρές προτείνεται το τελευταίο διάστημα η θεσμοθέτηση δικονομικής υποχρέωσης αποκάλυψης της χρήσης τους. Προς την κατεύθυνση αυτήν, ιδιαίτερα σημαντική είναι η σύσταση που περιέχεται στις προαναφερθείσες από 19.12.2025 Κατευθυντήριες Γραμμές για τη Χρήση της Παραγωγικής Τεχνητής Νοημοσύνης από τα Δικαστήρια της CEPEJ, όπου, εκτός των άλλων, συστήνεται στα δικαστήρια να θεσπίσουν πρωτόκολλα για την παρακολούθηση της χρήσης ΤΝ από τα μέρη, συμπεριλαμβανομένης της υποχρεωτικής γνωστοποίησης χρήσης εργαλείων Παραγωγικής (Generative) ΤΝ στην προετοιμασία ή στην έρευνα της υπόθεσης, με απειλή κυρώσεων για μη γνωστοποίηση ή υπερβολική εξάρτηση από μη επαληθευμένο περιεχόμενο.¹⁹⁵ Η σύσταση αυτή ακολουθεί μια τάση που από τα τέλη του 2025 παρατηρείται διεθνώς (κυρίως στις ΗΠΑ¹⁹⁶ και στο Ηνωμένο Βασίλειο, αλλά και αλλού¹⁹⁷) προς την κατεύθυνση η χρήση της

¹⁹⁵ «Τα δικαστήρια θα πρέπει να είναι ιδιαίτερα επιφυλακτικά όσον αφορά πολλαπλές μορφές κακής συμπεριφοράς που σχετίζονται με την ΑΙ, συμπεριλαμβανομένων: της υποβολής πλαστών αποδεικτικών στοιχείων, ανύπαρκτων νομικών αναφορών ή περιεχομένου που δημιουργήθηκε από ΑΙ, το οποίο δεν έχει επαρκώς ελεγχθεί από κατάλληλους επαγγελματίες των προσπάθειών χειραγώγησης συστημάτων ΑΙ μέσω κρυφών εντολών ή οδηγιών ενσωματωμένων στα έγγραφα, όπως αόρατο κείμενο, ασυνήθιστη μορφοποίηση ή μικροσκοπικές ενσωματωμένες οδηγίες σχεδιασμένες για να υπονομεύσουν την ανάλυση με τη βοήθεια ΑΙ και της χρήσης Γενετικής Τεχνητής Νοημοσύνης για τη δημιουργία πλαστών βίντεο».

¹⁹⁶ Στις ΗΠΑ πολλοί ομοσπονδιακοί δικαστές έχουν εκδώσει μόνιμες εντολές που απαιτούν την αποκάλυψη πληροφοριών σχετικά με τη χρήση ΤΝ στα δικαστήρια, ενώ ορισμένοι, όπως, λ.χ., στη Βόρεια Περιφέρεια της Καλιφόρνια, έχουν ορίσει συγκεκριμένες λεπτομερείς απαιτήσεις.

¹⁹⁷ Ιδιαίτερο ενδιαφέρον παρουσιάζει η περίπτωση του Δικαστηρίου Κτηματολογικών Διαφορών του Οντάριο (Ontario Land Tribunal) στον Καναδά, όπου υπάρχει Οδηγός για τη Χρήση ΤΝ στις δικαστικές διαδικασίες. Ο Οδηγός αυτός, μετά την τελευταία αναθεώρησή του την 29η.1.2025, προβλέπει ότι από την 30ή.3.2026 οι διάδικοι πρέπει να υποβάλλουν μια δήλωση για τη χρήση ΤΝ. Ειδικότερα, προβλέπεται ότι, «ενώ το δικαστήριο δεν απαγορεύει τη χρήση Τεχνητής Νοημοσύνης (ΤΝ) για τη δημιουργία ή παραγωγή περιεχομένου σε ένα έγγραφο που προετοιμάζεται για σκοπούς διαδικασίας, τα άτομα υποχρεούνται να γνωστοποιούν οποιαδήποτε τέτοια χρήση της ΑΙ μέσω δήλωσης». Και συνεχίζει σημειώνοντας: «Η πρώτη παράγραφος κάθε εγγράφου στο οποίο χρησιμοποιήθηκε ΤΝ για τη δημιουργία ή παραγωγή περιεχομένου πρέπει να περιέχει την εξής δήλωση: Τεχνητή νοημοσύνη (ΑΙ) χρησιμοποιήθηκε για τη δημιουργία περιεχομένου σε αυτό το έγγραφο [στις παραγράφους x έως ψ]. Όλο το περιεχόμενο που δημιουργήθηκε από την ΤΝ, καθώς και η αυθεντικότητα όλων των πηγών που αναφέρονται σε αυτό το έγγραφο έχουν εξεταστεί και επαληθευτεί από [συμπεριλάβετε το όνομα του ατόμου]». Τέτοια δήλωση απαιτείται μόνον όταν χρησιμοποιείται Γενετική ΤΝ για τη δημιουργία περιεχομένου σε ένα έγγραφο. Δεν απαιτείται δήλωση εάν η ΤΝ χρησιμοποιήθηκε μόνο για να προτείνει αλλαγές ή να αξιολογήσει περιεχόμενο που ήδη δημιουργήθηκε από έναν άνθρωπο, ο οποίος στη συνέχεια εξέτασε και υλοποίησε χειροκίνητα τις αλλαγές.

TN να γνωστοποιείται όταν επηρεάζει ουσιαστικά το υλικό που υποβάλλεται σε δικαστήριο.

Η υποχρέωση μιας τέτοιας γνωστοποίησης μπορεί να αφορά τον προσδιορισμό του συγκεκριμένου εργαλείου TN που χρησιμοποιήθηκε, τον προσδιορισμό των τμημάτων του εγγράφου που αυτό δημιούργησε και την πιστοποίηση ότι ένας άνθρωπος έχει επαληθεύσει την ακρίβεια του αποτελέσματος. Το τελευταίο σχετίζεται με τα περιστατικά αναφοράς ανύπαρκτης νομολογίας προερχόμενης από TN –που μπορεί να οφείλονται στο φαινόμενο των λεγόμενων «παραισθήσεων» (hallucinations)–, που συνήθως θεωρούνται περιφρόνηση του δικαστηρίου. Στις περιπτώσεις αυτές, τα δικαστήρια απηύθυναν αυστηρές προειδοποιήσεις ή προέβησαν σε κυρώσεις, όπως επιβολή προστίμου, καταβολές εξόδων, αλλά και παραπομπές δικηγόρων στις αρμόδιες ρυθμιστικές/πειθαρχικές Αρχές.¹⁹⁸

Θέση των δικηγόρων ως προς τη θέσπιση ή μη μιας τέτοιας υποχρέωσης δεν έχει ακόμη διαμορφωθεί. Μία κάπως σχετική αναφορά υπάρχει στις προαναφερθείσες Considerations on the legal aspects of Artificial Intelligence του CCBE του 2020, όπου διατυπώνεται η άποψη πως όλα τα εμπλεκόμενα μέρη σε μια δικαστική διαδικασία θα πρέπει πάντα να είναι σε θέση να εντοπίζουν, στο πλαίσιο μιας δικαστικής απόφασης, τα στοιχεία που προκύπτουν από την εφαρμογή ενός εργαλείου TN (άρα η θέση αυτή αφορά τη χρήση εργαλείου TN από τον δικαστή και όχι από τα μέρη). Προστίθεται, όμως, αμέσως μετά η γενικότερη τοποθέτηση ότι θα πρέπει να υπάρχει αυστηρός διαχωρισμός μεταξύ των δεδομένων ή αποτελεσμάτων που προέρχονται από τη λειτουργία ενός συστήματος TN και των υπόλοιπων δεδομένων στην επίδικη διαφορά. Ένας, όμως, τέτοιος διαχωρισμός μάλλον προϋποθέτει υποχρέωση γνωστοποίησης όσων αποτελούν προϊόν TN.

Το ζήτημα είναι βέβαιο ότι θα μας απασχολήσει έντονα στο άμεσο μέλλον, αφού η χρήση εργαλείων TN στη δικαστηριακή πρακτική αυξάνεται με πολύ ταχείς ρυθμούς και, συνεπώς, η τυχόν θεσμοθέτηση μια τέτοιας υποχρέωσης κατ' αποδοχήν της σύστασης της CEPEJ θα έχει οπωσδήποτε σοβαρές επιπτώσεις. Βάσιμα, πάντως, μπορεί να πιθανολογηθεί ότι η εφαρμογή της στην πράξη εξαρτάται από την επίλυση αρκετά δυσεπίλυτων προβλημάτων ερμηνείας και οριοθέτησης, που θα απαιτήσουν περαιτέρω επεξεργασίες και εξειδικεύσεις, χωρίς βεβαίως αυτό καθόλου να υποτιμά τη σημασία της.

¹⁹⁸ Ενδιαφέρον παρουσιάζει και η περίπτωση της Αυστραλίας, όπου σε κάποια δικαστήρια (the Supreme Court, Land and Environment Court και District Courts) έχουν θεσπιστεί απαγορεύσεις χρήσης TN για ορισμένες πράξεις, ενώ για τις υπόλοιπες απαιτείται γνωστοποίηση τυχόν χρήσης της, αλλού προβλέπεται ότι για ορισμένες πράξεις (όπως για σύνταξη περιεχομένου πραγματογνωμοσύνης) η χρήση TN είναι επιτρεπτή μόνο έπειτα από άδεια του δικαστηρίου (αυτή είναι η θέση του Ομοσπονδιακού Δικαστηρίου), ενώ, τέλος, σε άλλα δικαστήρια (Victoria) συστήνεται απλώς η αυξημένη επιμέλεια –«particular caution»– στη χρήση εργαλείων TN χωρίς συγκεκριμένους περιορισμούς ή υποχρεώσεις.

7.5. Συμπέρασμα

Το πλαίσιο αρχών για μια ηθική Τεχνητή Νοημοσύνη στον χώρο της Δικαιοσύνης έχει ήδη προδιαγραφεί, χωρίς βεβαίως αυτό να σημαίνει ότι δεν θα υπάρξουν πρακτικές δυσκολίες για τη διασφάλισή του σε κάθε δικαστική διαδικασία, σε κάθε δικηγορική υπηρεσία. Κεντρική θέση στο πλαίσιο αυτό, τουλάχιστον με βάση τις σημερινές δυνατότητες και το επίπεδο αξιοπιστίας της ΤΝ, είναι η διατήρηση του *ανθρωποκεντρικού χαρακτήρα της Δικαιοσύνης*. Ότι, δηλαδή, η εξουσία απονομής της πρέπει να παραμείνει ολόκληρη στα χέρια του δικαστή και η ΤΝ να περιοριστεί μόνο σε υποστηρικτικό ρόλο, ελεγχόμενο και διαφανή. Η εισαγωγή της ΤΝ στις δικαστικές διαδικασίες πρέπει να γίνει με αργά βήματα, σύνεση και στρατηγικό σχεδιασμό. Όχι μόνο για να διασφαλίζονται όλες οι απαραίτητες εγγυήσεις, αλλά και για να κερδίζεται η απαραίτητη εμπιστοσύνη¹⁹⁹ της κοινωνίας, η οποία ευλόγως διατηρεί στάση καχυποψίας.²⁰⁰ Οι εφαρμογές ΤΝ πρέπει να προσαρμόζονται στις χρήσεις και τα περιβάλλοντα για τα οποία προορίζονται και όχι το αντίθετο. Και σε κάθε δίλημμα ορθή να θεωρείται η αυστηρή εκδοχή.

Για τους δικηγόρους, απαιτείται άμεσα ενημέρωση και εκπαίδευσή τους, ώστε να μπορούν να κατανοούν και να χρησιμοποιούν υπεύθυνα τα διάφορα εργαλεία ΤΝ, προκειμένου να είναι σε θέση να τηρούν τις υποχρεώσεις τους απέναντι στα δικαστήρια, στους πελάτες και τους συναδέλφους τους, σύμφωνα με τη δεοντολογία και τις θεμελιώδεις αρχές του δικηγορικού επαγγέλματος. Ουσιαστικό ρόλο και ευθύνη έχουν να διαδραματίσουν οι Δικηγορικοί Σύλλογοι, οι οποίοι μπορούν ίσως να καλύψουν και το χάσμα διαπραγματευτικής ισχύος μεταξύ της συντριπτικής πλειονότητας των μελών τους και των παρόχων εφαρμογών ΤΝ, για τη σύναψη συμβατικών όρων που θα διασφαλίζουν διαφάνεια και συμβατότητα της λειτουργίας τους με τις επαγγελματικές υποχρεώσεις των δικηγόρων.

¹⁹⁹ Μήτρου, Λ. (2023). *Μπορεί ο αλγόριθμος...* Πανεπιστημιακές Εκδόσεις Κρήτης. Κοφίνης, Σ. (2021). Από την ψηφιακή δικαιοσύνη στον ψηφιακό δικαστή: Μπορεί η Τεχνητή Νοημοσύνη να αντικαταστήσει τους δικαστές; Εθνική Σχολή Δικαστικών Λειτουργιών.

²⁰⁰ διαΝΕΟσις (2026). *Τι πιστεύουν οι Έλληνες για την Τεχνητή Νοημοσύνη Α΄ Μέρος – Έκθεση αποτελεσμάτων*. https://www.dianeosis.org/wp-content/uploads/2026/05/AI_Dimoskopiki_Apotelesmata_190526.pdf

Η μετάβαση στην εποχή της ΤΝ μας υποχρεώνει να αναστοχαστούμε τις βασικές αρχές που διέπουν τον θεσμό της Δικαιοσύνης και τους ρόλους δικαστών και δικηγόρων. Το αίτημα για μια ηθική ΤΝ τις επιβεβαιώνει. Οι προκλήσεις για τις προσαρμογές και επικαιροποιήσεις που θα απαιτηθούν μπορούν να γίνουν ευκαιρία για να κάνουμε τη Δικαιοσύνη καλύτερη.

Βιβλιογραφία

- Κανέλλος, Λ. (2021). *Εφαρμογές Τεχνητής Νοημοσύνης στο δίκαιο και στη δικαστική πρακτική*. Νομική Βιβλιοθήκη.
- Κοφίνης, Σ. (2021). Από την ψηφιακή δικαιοσύνη στον ψηφιακό δικαστή: Μπορεί η Τεχνητή Νοημοσύνη να αντικαταστήσει τους δικαστές; Εθνική Σχολή Δικαστικών Λειτουργών.
- Μήτρου, Λ. (Επιμ.). (2023). *Μπορεί ο αλγόριθμος...* Πανεπιστημιακές Εκδόσεις Κρήτης.
- Σηφάκης, Ι. (2020). *Κατανοώντας και αλλάζοντας τον κόσμο*. Αρμός.
- Συμεωνίδης, Ι. (2019). Τεχνητή νοημοσύνη και δικαιοσύνη. *Εφημερίδα Διοικητικού Δικαίου (ΕφημΔΔ)*, 4, 462-469.
- Τσινόρεμα, Σ. (2024). Τεχνητή νοημοσύνη με ανθρώπινο πρόσωπο. Προς μια τεχνοθική της ευθύνης. *διά-Λογος*, 14, 163-192.
- Council of Bars and Law Societies of Europe (CCBE) (2014). *CCBE Comparative Study on Governmental Surveillance of Lawyers' Data in the Cloud*. <https://www.ccbe.eu>
- Council of Bars and Law Societies of Europe (CCBE) (2019). *CCBE Comments on the Stakeholders' Consultation on Draft Artificial Intelligence Ethics Guidelines*. <https://www.ccbe.eu>
- Council of Bars and Law Societies of Europe (CCBE) (2020). *CCBE Considerations on the Legal Aspects of AI*. <https://www.ccbe.eu>
- Council of Bars and Law Societies of Europe (CCBE) (2022). *CCBE Guide on the use of AI-based tools by lawyers and law firms in the EU*. <https://www.ccbe.eu>
- Council of Bars and Law Societies of Europe (CCBE) (2025a). *CCBE Guidelines on the use of cloud computing by Bars and Lawyers*. <https://www.ccbe.eu>

www.ccbe.eu/fileadmin/speciality_distribution/public/documents/IT_LAW/ITL_Guides_recommendations/EN_ITL_20250227_CCBE-guidelines-on-the-use-of-cloud-computing-by-lawyers.pdf

Council of Bars and Law Societies of Europe (CCBE) (2025b). *CCBE Guide on the use of generative AI by lawyers*. https://www.ccbe.eu/fileadmin/speciality_distribution/public/documents/IT_LAW/ITL_Guides_recommendations/EN_ITL_20251002_CCBE-guide-on-the-use-of-the-use-of-generative-AI-for-lawyers.pdf

Council of Bars and Law Societies of Europe (CCBE) (2026). *CCBE Technical Guide on the use of AI tools and models by lawyers*. https://www.ccbe.eu/fileadmin/speciality_distribution/public/documents/IT_LAW/ITL_Guides_recommendations/EN_ITL_20260327_CCBE-technical-guide-on-the-use-of-AI-tools-and-models-by-lawyers.pdf

Danks, D., & London, A. J. (2017). *Algorithmic bias in autonomous systems*. In C. Sierra (Ed.), *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17)* (pp. 4691-4697). International Joint Conferences on Artificial Intelligence Organization. <https://doi.org/10.24963/ijcai.2017/654>

European Commission for the Efficiency of Justice (CEPEJ) (2018). *European Ethical Charter on the use of artificial intelligence (AI) in judicial systems and their environment*. Council of Europe. <https://www.coe.int/en/web/cepej>

European Commission for the Efficiency of Justice (CEPEJ) (2025). *Guidelines on the use of generative Artificial Intelligence (AI) in courts*. Council of Europe. <https://www.coe.int/en/web/cepej/cepej-tools-on-cyberjustice-and-artificial-intelligence>

European Union Agency for Fundamental Rights (FRA) (2025). *Assessing High-risk Artificial Intelligence: Fundamental Rights Risks*. Publications Office of the European Union.

European Union Agency for Fundamental Rights (FRA) (n.d.). *Facial recognition technology: Fundamental rights considerations in the context of law enforcement*. Publications Office of the European Union.

High-Level Expert Group on Artificial Intelligence (2019). *Ethics guidelines for trustworthy AI*. European Commission. <https://digital-strategy.ec.europa.eu>

Organisation for Economic Co-operation and Development (OECD) (2019). *OECD AI principles*. <https://oecd.ai/en/ai-principles>

Terzidou, K. (2025). *Generative AI systems in legal practice offering quality legal services while upholding legal ethics*. *International Journal of Law in Context*, 21(3), 431-452. <https://doi.org/10.1017/S1744552325000047>

United Nations General Assembly (2024). *Resolution on the promotion of safe, secure and trustworthy artificial intelligence systems (A/RES/78/265)*. <https://documents.un.org>

United Nations Educational, Scientific and Cultural Organization (UNESCO) (2023). *Global toolkit on AI and the rule of law for the judiciary*. <https://unesdoc.unesco.org>

United Nations Educational, Scientific and Cultural Organization (UNESCO) (2025). *Guidelines for the use of AI systems in courts and tribunals*. <https://unesdoc.unesco.org/ark:/48223/pf0000396582>

[ΠΩΣ] ΜΠΟΡΕΙ Η ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΝΑ ΕΙΝΑΙ
ΗΘΙΚΗ ΚΑΙ ΣΥΝΝΟΜΗ;
Ιούλιος 2026

Οι επιπτώσεις της Τεχνητής Νοημοσύνης στον κόσμο της εργασίας και τη διαμόρφωση των εργασιακών σχέσεων



**Ηρακλής Βογιατζής, Λίλιαν Μήτρου,
Κλειώ Σγουροπούλου**

Συγγραφική συνεισφορά ανά υποκεφάλαιο:

8.1. – Ηρακλής Βογιατζής

8.2., 8.4., 8.5., 8.6. – Λίλιαν Μήτρου

8.3. – Κλειώ Σγουροπούλου

Μια επιφύλαξη που συχνά εκφράζεται σε σχέση με τη ραγδαία «εισβολή» της Τεχνητής Νοημοσύνης σε πολλά κοινωνικά και επαγγελματικά πεδία είναι αν «θα μας πάρουν τα ρομπότ τις δουλειές».²⁰¹ Αντίστοιχοι φόβοι εκδηλώνονταν κάθε φορά που η τεχνολογία και τα επιτεύγματά της υποστήριζαν και ακόμη περισσότερο όταν αντικαθιστούσαν τον άνθρωπο κατά την εκτέλεση εργασιών. Ανεξάρτητα από τη βασιμότητα ή την υπερβολή ως προς την έκταση και την ένταση τέτοιων επιφυλάξεων, η ΤΝ φαίνεται πως αλλάζει ή μπορεί να αλλάξει ριζικά τον κόσμο της εργασίας, εγείροντας ηθικά και νομικά ζητήματα ως προς τα δικαιώματα των εργαζομένων, τις κοινωνικές επιπτώσεις από τον μετασχηματισμό της εργασίας, τα νέα «χάσματα».

Πιο συγκεκριμένα, η χρήση συστημάτων ΤΝ μπορεί να έχει εκτενείς συνέπειες για την εργασία και την απασχόληση που αφορούν πολυάριθμες κατηγορίες επαγγελματιών, επηρεάζοντας εργοδότες και εργαζομένους. Η χρήση συστημάτων ΤΝ θα μπορούσε να δυσχεράνει την πρόσβαση στην εργασία, να αυξήσει την ένταση της εργασίας, να ενισχύσει ή να επιδεινώσει τις ανισορροπίες ισχύος μεταξύ εργοδοτών και εργαζομένων, να μειώσει την ανθρώπινη συμμετοχή σε αποφάσεις σχετικά με την πρόσληψη, την αξιολόγηση και την απόλυση, και να υπονομεύσει θεμελιώδεις αρχές και δικαιώματα στην εργασία. Ο Ηρακλής Βογιατζής ανιχνεύει τα ζητήματα αυτά στη συμβολή του στην παρούσα μελέτη.

201 Μπαλαμπανίδης, Γ., Φαναράς, Σ., & Αποστολοπούλου, Π. (2026). *Τι πιστεύουν οι Έλληνες για την Τεχνητή Νοημοσύνη Α΄ Μέρος*. διαΝΕΟσις. https://www.dianeosis.org/wp-content/uploads/2026/05/AI_METRON_ANALYSIS_essay_190526.pdf Όπως σημειώνουν οι Γιάννης Μπαλαμπανίδης, Στράτος Φαναράς και Πέννυ Αποστολοπούλου στον σχολιασμό της έρευνας αυτής που διεξήχθη για λογαριασμό της διαΝΕΟσις αναφορικά με το τι πιστεύουν οι Έλληνες για την ΤΝ, όσον αφορά το πεδίο της εργασίας, ειδικότερα, αξίζει να σημειωθεί το εξής: Η επίδραση της ΤΝ αναμένεται να είναι μάλλον θετική στη διάσταση της παραγωγικότητας, και ακόμη περισσότερο από τους άνδρες παρά από τις γυναίκες (75,5% έναντι 58,9%), από τα άτομα υψηλότερου εκπαιδευτικού επιπέδου (79,8% στους κατόχους μεταπτυχιακού/διδακτορικού έναντι 56,4% στους αποφοίτους Δευτεροβάθμιας) και υψηλότερης κοινωνικής τάξης (76,3% στην ανώτερη/μέση ανώτερη τάξη, αλλά 55,4% στη μέση κατώτερη), καθώς και από όσους κατοικούν στην Αττική παρά στην Περιφέρεια – δηλαδή, σε στρώματα του πληθυσμού που είναι περισσότερο insiders της αγοράς εργασίας. Αντίθετα, η αντίληψη ότι η ΤΝ θα έχει αρνητική επίδραση στις θέσεις εργασίας, δηλαδή ότι θα υποκαταστήσει ανθρώπινη εργασία, φαίνεται να ενισχύεται σε κατηγορίες που από αυτήν τη σκοπιά αισθάνονται περισσότερο επισφαλείς: στις γυναίκες, στις μεγαλύτερες παραγωγικές ηλικίες 40-64 ετών, στους άνεργους, που είναι εξ ορισμού outsiders της αγοράς εργασίας, και στα άτομα χαμηλότερης κοινωνικής τάξης.

8.1. Το «τέλος της εργασίας» ή νέες μορφές ψηφιακής εργασίας

8.1.1. Οι επαναλαμβανόμενες προβλέψεις

Τον Φεβρουάριο του 2026, ο διευθύνων σύμβουλος της Microsoft AI, Mustafa Suleyman, δήλωσε: «Νομίζω πως θα έχουμε [συστήματα Τεχνητής Νοημοσύνης με] αποδοτικότητα στο επίπεδο του ανθρώπου, στις περισσότερες, εάν όχι όλες, επαγγελματικές εργασίες. Οι περισσότερες δουλειές γραφείου που γίνονται μπροστά από μια οθόνη, είτε αυτή είναι του δικηγόρου είτε αυτή είναι του λογιστή είτε του πρότζεκτ μάνατζερ, είτε του μάρκετινγκ, θα αυτοματοποιηθούν πλήρως από την Τεχνητή Νοημοσύνη τους επόμενους 12 με 18 μήνες. Μπορούμε να δούμε ήδη τι συμβαίνει στην περίπτωση των προγραμματιστών» (Khalaf, 2026).

Η απειλή εκτοπισμού του ανθρώπου από την παραγωγή αποτελεί μέρος του δημόσιου διαλόγου τις τελευταίες δεκαετίες (Rifkin, 1995), ωστόσο η άνοδος της ΤΝ ενίσχυσε περισσότερο τους φόβους για το «τέλος της εργασίας». Σε έναν κόσμο όπου τα μοντέλα μηχανικής μάθησης μπορούν να κάνουν τη δουλειά ενός δικηγόρου ή ενός γραφίστα, οι θέσεις εργασίας της μεσαίας τάξης, που είχαν δημιουργηθεί στις μεταπολεμικές οικονομίες, φαίνεται πως κινδυνεύουν από την αυτοματοποίηση.

Οι μελέτες που πραγματοποιήθηκαν τα προηγούμενα χρόνια είναι αποκάλυπτικές, όπως η διάσημη έρευνα της Οξφόρδης (Frey & Osborne, 2017), η οποία υποστηρίζει πως το 47% των θέσεων εργασίας που υπάρχουν σήμερα κινδυνεύει από την πλήρη αυτοματοποίηση. Ένα πέπλο καλύπτει την ΤΝ, δίνοντάς της υπερφυσικές δυνάμεις, οι οποίες «ανατρέπουν» τα δεδομένα των μοντέλων παραγωγής αξίας. Επιχειρήσεις (κυρίως από τον τομέα των ψηφιακών τεχνολογιών) απολύουν μάζες εργαζομένων (Abril, 2026), αναμένοντας την αντικατάστασή τους από Μεγάλα Γλωσσικά Μοντέλα και λογισμικά Αυτενεργούσας Τεχνητής Νοημοσύνης (Agentic AI). Μπορεί τα συστήματα που τώρα διαθέτουν αυτές οι εταιρείες να είναι ημιτελή ή να λειτουργούν σε καθεστώς «παραισθήσεων» (hallucinations), παρ' όλα αυτά μεγάλοι όμιλοι επιταχύνουν τις διαδικασίες αποδέσμευσης του προσωπικού τους (Davenport & Srinivasan, 2026). Η ΤΝ δεν κυριαρχεί με όρους αποτελεσματικότητας, αλλά με αυτούς της υπόσχεσης ενός τεράστιου κέρδους σε καθεστώς απόλυτου ανταγωνισμού. Σε αυτό το κλίμα, ο διευθύνων σύμβουλος

της OpenAI, Sam Altman, βάζει στοίχηματα στην ιδιωτική συνομιλία που έχει με φίλους του, διευθυντές εταιρειών τεχνολογίας. Ποντάρουν στο έτος που θα εμφανιστεί η πρώτη εταιρεία με αξία δισεκατομμυρίων, στην οποία θα εργάζεται ένας και μόνο άνθρωπος (Ashley, 2025).

Το μόνο που μπορούμε να πούμε με βεβαιότητα είναι ότι το τοπίο της εργασίας αλλάζει, ίσως με διαφορετικό τρόπο από αυτόν που περιγράφει ο Mustafa Suleyman. Η ΤΝ δύναται να αυτοματοποιήσει μεγάλο αριθμό εργασιακών δραστηριοτήτων. Πράγματι, οι τεχνολογίες Παραγωγικής Τεχνητής Νοημοσύνης (Generative AI) επιδρούν στις θέσεις που απαιτούν διανοητική εργασία, οι οποίες μέχρι τώρα σχετιζόνταν με την αύξηση της μεσαίας τάξης.

Σε μια πρόσφατη μελέτη, η Διεθνής Οργάνωση Εργασίας (International Labor Organization) εξέτασε τα εργασιακά καθήκοντα που δύναται να εκτελέσει το Μεγάλο Γλωσσικό Μοντέλο τέταρτης γενιάς της OpenAI, GPT-4. Σύμφωνα με αυτήν, παρουσιάζεται μια μετατόπιση της αυτοματοποίησης. Η ΤΝ δεν απειλεί μόνο τις χειρωνακτικές ή διανοητικές εργασίες, που χαρακτηρίζονται από επαναλαμβανόμενες δραστηριότητες. Οι προβλέψεις της δείχνουν μεγάλη πιθανότητα αντικατάστασης εργασιών που σχετίζονται με γνωσιακές δραστηριότητες, όπως η κατανόηση και επεξεργασία κειμένου, η σύνταξη αναφορών κ.ά. (Gmyrek et al., 2023).

Άλλες μελέτες εντοπίζουν πιο συγκεκριμένες διαταραχές στις παραδοσιακές αγορές εργασίας, όπως για παράδειγμα τη μείωση της απασχόλησης των νεοεισερχομένων σε αυτές, ειδικότερα στους τομείς της εξυπηρέτησης πελατών και της ανάπτυξης λογισμικού (Brynjolfsson et al., 2025). Ταυτόχρονα, το αφήγημα περί του «τέλους της εργασίας» δεν μένει μόνο στο επίπεδο εξαγγελιών που στοχεύουν να πείσουν τους επενδυτές για τις μελλοντικές δυνατότητες του ενός ή του άλλου μοντέλου μηχανικής μάθησης. Τα διευθυντικά στελέχη της Amazon, σε μια ενημέρωση του διοικητικού συμβουλίου, υποστήριξαν πως επεξεργάζονται συστήματα ρομποτικής αυτοματοποίησης βασισμένα στην ΤΝ, τα οποία θα τους δώσουν τη δυνατότητα να εξοικονομήσουν 600.000 θέσεις εργασίας στις αλυσίδες της εφοδιαστικής (Kantor, 2025).

Το ενδιαφέρον εδώ είναι πως η εταιρεία δεν εξετάζει απολύσεις, αλλά διπλασιασμό των δραστηριοτήτων της, χωρίς τις αντίστοιχες νέες προσλήψεις. Η πολυπλοκότητα του φαινομένου απαγορεύει τις εύκολες μαθηματικές αναγωγές για τη μέτρηση των θέσεων εργασίας, καθώς αυτό που πρέπει να μετρηθεί δεν είναι η μείωση ή αύξηση των θέσεων στην αγορά εργασίας (οι οποίες μπορεί να αυξάνονται ισχνά ή να μένουν σταθερές), αλλά οι αλλαγές σε σχέση με την αντίστοιχη επένδυση κεφαλαίων. Μια τέτοια αρνητική τάση παρουσιάζει το φαινόμενο της «μεγάλης αποσύνδεσης» μεταξύ της παραγωγικότητας (που προέρχεται από την είσοδο νέων τεχνολογιών) και

των μισθών, το οποίο, παρότι δεν σχετίζεται μόνο με τις νέες τεχνολογίες, φαίνεται να συνοδεύει την ψηφιακή εποχή (Paternesi & Stirati, 2023).

Η εικόνα της αντικατάστασης του ανθρώπου από τις μηχανές είναι, σε έναν βαθμό, παραπλανητική. Η αυτοματοποίηση δεν έχει ως αποτέλεσμα μόνο την αντικατάσταση θέσεων εργασίας και τη μείωση ορισμένων από αυτές σε τομείς όπου οι δραστηριότητες είναι σαφώς δομημένες και ενέχουν μεγάλα επίπεδα επανάληψης, αλλά αυξάνει τις θέσεις εργασίας σε τομείς όπου τα καθημερινά καθήκοντα επαυξάνονται από τη χρήση της ΤΝ (Srinivasan, 2026). Η «επαύξηση» (augmentation) της εργασίας είναι ένα φαινόμενο που εμφανίζεται σε δουλειές όπου η ανθρώπινη δράση και η κριτική σκέψη ενισχύονται από τη δυνατότητα επεξεργασίας μεγάλου όγκου δεδομένων των μοντέλων μηχανικής μάθησης. Τέτοια παραδείγματα υπάρχουν σε πολλούς τομείς, όπως ο Χρηματοπιστωτικός και ο τομέας της Φαρμακευτικής Έρευνας. Ένα άλλο παράδειγμα είναι η ανάθεση στην ΤΝ όλων των χρονοβόρων και ανιαρών καθηκόντων που επιβάλλει η διαχείριση της γραφειοκρατίας της έρευνας στα πανεπιστήμια. Κάτι τέτοιο θα μπορούσε να «επαυξήσει» την ικανότητα των ατόμων που εργάζονται στην έρευνα, ενισχύοντας πολύπλευρα όχι μόνο τα ίδια, αλλά και το πανεπιστήμιο και την οικονομία.

Η αναλογία ανάμεσα στις εργασίες που καταστρέφονται και σε αυτές που δημιουργούνται αποτελεί κρίσιμο δείκτη των εξελίξεων. Εάν οι πολιτικές που ακολουθούνται από κράτη και επιχειρήσεις είναι έτοιμες να απορροφήσουν τη ζημιά που προκαλεί η αυτοματοποίηση, δίνοντας χρόνο προσαρμογής στη νέα πραγματικότητα, τότε η κρίση μπορεί να αποσοβηθεί μέσω της επανειδίκευσης των εργαζομένων και της αναβάθμισης των δεξιοτήτων τους.

Ο νομπελίστας οικονομολόγος Daron Acemoglu μαζί με τον Pasqual Restrepo έχουν χαρτογραφήσει τον διάλογο για το «τέλος της εργασίας», επιχειρώντας να υπερβούν το δίπολο ανάμεσα στην κοινωνία της αφθονίας και στην κοινωνία χωρίς δουλειές. Προκρίνουν τη μελέτη του ζητήματος πάνω στη βάση δύο παράλληλων φαινομένων στον χώρο της παραγωγής. Αφενός παρατηρούν τα φαινόμενα εκτόπισης του ανθρώπου από την παραγωγή, τα οποία παρουσιάσαμε σύντομα μέχρι τώρα. Αφετέρου, εντοπίζουν αυτό που ονομάζουν «αντισταθμιστικές δυνάμεις» (Acemoglu & Restrepo, 2018), δηλαδή τις νέες μορφές εργασίας, την ενίσχυση της παραγωγικότητάς της και τα νέα παραγωγικά δίκτυα που δημιουργεί η ΤΝ. Μια χαρακτηριστική τέτοια δύναμη εμφανίζεται με τη μείωση του κόστους της παραγωγής, η οποία συντελεί στη δημιουργία νέων αναγκών προς ικανοποίηση, επεκτείνοντας τις παραγωγικές αλυσίδες. Για παράδειγμα, ευρήματα όπως αυτά του Information Technology and Innovation Foundation (ITIF) τονίζουν τη δημιουργία 119.900 νέων θέσεων εργασίας στις ΗΠΑ που προκάλεσε η ΤΝ, σε αντίθεση με τις 12.700 που αυτοματοποιήθηκαν (Ostertag, 2025). Μια τέτοια παρατήρηση δεν πρέπει να προκαλεί εφησυχασμό, όμως προσφέρει στην αξιολόγηση

των δυνατοτήτων που εμφανίζουν τα μοντέλα της μηχανικής στο πλαίσιο της ευρύτερης παραγωγικής διαδικασίας.

Οι αντισταθμιστικές δυνάμεις των Daron Acemoglu και Pasqual Restrepo εμφανίζονται με ειρωνικό τρόπο στα προβλήματα που αντιμετωπίζουν τα επενδυτικά εγχειρήματα της υποδομής της ΤΝ, όπως τα data centers. Μία από τις προκλήσεις της ανάπτυξης των υποδομών σε μαζική κλίμακα, την οποία προκαλούν οι ξέφρενες κεφαλαιακές επενδύσεις, είναι η έλλειψη εργατικού δυναμικού. Ειδικότητες όπως οικοδόμοι, ηλεκτρολόγοι και ψυκτικοί βρίσκονται σε αναντιστοιχία με τις προβλεπόμενες ανάγκες για τη δημιουργία και τη συντήρηση αυτών των κέντρων. Τα κενά σε αυτές τις ειδικότητες στις ΗΠΑ κυμαίνονται από 45% έως 80% (Girishankar & Smith, 2025).

Η σημασία των πολιτικών διαχείρισης αναδεικνύεται ως το επίδικο μιας κοινωνικής και τεχνολογικής διαδικασίας, η οποία πρέπει να απομακρυνθεί από αφηγήματα περί τεχνολογικής λαίλαπας, που το μόνο που επιτυγχάνουν είναι η αποφυγή ρύθμισης και λογοδοσίας των ολιγοπωλίων της τεχνολογίας. Το ερώτημα δεν αφορά, λοιπόν, μόνο την απώλεια θέσεων εργασίας, αλλά τις μετατοπίσεις που λαμβάνουν χώρα στις αγορές εργασίας.

Οι μετατοπίσεις δεν εμφανίζονται μόνο στο πλήθος των θέσεων εργασίας, αλλά και στις μορφές εργασίας. Η άνοδος των Μεγάλων Γλωσσικών Μοντέλων επέδρασε καθοριστικά στην παραγωγή κώδικα, δηλαδή στην εργασία των προγραμματιστών. Σε μια, σχεδόν, βιβλική ανατροπή, οι μηχανικοί λογισμικού και οι προγραμματιστές βρίσκονται τα τελευταία χρόνια σε καθεστώς αντικατάστασης από τις μεγάλες εταιρείες λογισμικού. Οι νέες εφαρμογές της ΤΝ μπορούν, πλέον, να παράγουν χιλιάδες γραμμές κώδικα με απλές προτροπές. Η μηχανική λογισμικού αντικαθίσταται από τη μηχανική προτροπών (prompting engineering). Η επαύξηση της παραγωγικότητας των προγραμματιστών είναι αδιαμφισβήτητη, κρίνοντας εκ πρώτης όψεως, ωστόσο τα πράγματα δεν είναι πάντα ρόδινα. Μια διεθνής έρευνα της ελβετικής εταιρείας ανοικτού λογισμικού Sonar, στην οποία συμμετείχαν πάνω από 1.000 προγραμματιστές και προγραμματίστριες, διαπίστωσε πως η εξοικονόμηση χρόνου, εξαιτίας της ΤΝ, δεν είναι αυτονόητη (SonarSource, 2025). Οι πιο συστηματικοί χρήστες υποστηρίζουν ότι η κοπιαστική δουλειά δεν εξαφανίζεται, αλλά μετατοπίζεται σε άλλες δραστηριότητες. Τα εργαλεία, ορισμένες φορές, παράγουν κώδικα «σπαγγέτι», ο οποίος απαιτεί από τον προγραμματιστή χρόνο και κόπο για την κατανόηση, την αποσφαλμάτωση και την επαναχρησιμοποίησή του.

8.1.2. Ο κόσμος ανάποδα

Το 2021, η Amazon παρουσίασε την καινοτομία «Just Walk Out» στα φυσικά καταστήματα λιανικής πώλησης Amazon Go. Τα καταστήματα λειτουργούσαν χωρίς ταμεία: Ο πελάτης αγόραζε αυτά που επιθυμούσε

και στη συνέχεια περνούσε την έξοδο του καταστήματος. Ένα σύστημα ΤΝ επιτηρούσε, με τη χρήση καμερών και αισθητήρων, τις αγορές και τις κινήσεις στους εσωτερικούς χώρους, στέλνοντας αυτόματα τη χρέωση στον λογαριασμό Amazon του πελάτη. Το σύστημα αυτό είχε ως σκοπό την παράκαμψη της διαδικασίας «σκαναρίσματος» και ζύγισης των προϊόντων από υπαλλήλους ή τη χρήση αυτόματων ταμείων από τους πελάτες. Ωστόσο, πίσω από την κουρτίνα της αυτοματοποίησης, η πραγματικότητα ήταν διαφορετική.

Οι κάμερες και οι αισθητήρες της ΤΝ μπορούσαν να καταγράψουν ένα ποσοστό των αγορών. Το 70% των πωλήσεων εξεταζόταν από ένα κρυφό εργατικό δυναμικό, το οποίο επιτηρούσε τις κινήσεις των πελατών μέσω των καμερών, κάνοντας τις απαραίτητες παρεμβάσεις και καταχωρίσεις. Οι «ψηφιακοί» ταμίες βρίσκονταν στην Ινδία και αριθμούσαν πάνω από χίλια άτομα. Η Amazon ονομάζει αυτό το εργατικό δυναμικό «ανθρώπινους αξιολογητές» (human reviewers), αποσιωπώντας λεπτομέρειες για τη φύση και την πληρωμή της εργασίας τους (Bridle, 2024).

Η ιστορία του Amazon Go δεν αποτελεί εξαίρεση ή στραβοπάτημα. Το 2005, ο ιδρυτής της εταιρείας, Jeff Bezos, παρουσίασε μια πραγματικά καινοτόμα ιδέα (αυτήν τη φορά), την πλατφόρμα «Mechanical Turk», την οποία ονόμασε «τεχνητή» Τεχνητή Νοημοσύνη. Οι εργαζόμενοι σε αυτήν την πλατφόρμα εκτελούν μικροεργασίες (microwork) για μερικά σεντς του δολαρίου, οι οποίες δεν μπορούν να αυτοματοποιηθούν από τις μηχανές. Η ιδέα που θέλει τους ανθρώπους να εργάζονται στην περιφέρεια του μηχανικού υπολογισμού, συμπληρώνοντάς τον, αποτελεί μια βασική επιχειρηματική δραστηριότητα της Amazon εδώ και δύο δεκαετίες. Η ιστορία των καταστημάτων της όχι μόνο δεν προκαλεί έκπληξη, αλλά είναι συνεπής με τον τρόπο που λύνονται πολλά τεχνολογικά προβλήματα.

Εργαζόμενοι σε πλατφόρμες ή υπάλληλοι σε μεγάλες εταιρείες τεχνικής υποστήριξης εργάζονται πυρετωδώς για την παραγωγή, επισημείωση (annotation) και εκπαίδευση των αλγορίθμων της ΤΝ. Η έρευνα των Paola Tubaro et al. ανέδειξε τους τρόπους με τους οποίους η ανθρώπινη εργασία υπάρχει πίσω από αυτά τα συστήματα (Tubaro et al., 2020). Οι άνθρωποι εμφανίζονται ως «εκπαιδευτές» των μοντέλων μηχανικής μάθησης, επιμελούμενοι τα δεδομένα τους. Άλλες φορές είναι «ελεγκτές»: Στέκονται μπροστά στα αποτελέσματα και τις προβλέψεις του μοντέλου, βοηθώντας το να αυτοβελτιωθεί. Η τρίτη κατηγορία είναι οι «μιμητές», όπως οι αόρατοι εργαζόμενοι της Amazon Go.

Η εργασία που είναι απαραίτητη για την παραγωγή και συντήρηση της ΤΝ βρίσκεται στο παρασκήνιο των αφηγήσεων για το μέλλον της εργασίας. Εκατομμύρια εργαζόμενοι βρίσκονται διασκορπισμένοι στα μήκη και τα πλάτη της υφηνλίου. Ψηφιακές πλατφόρμες εργασίας στη Λατινική Αμερική

εκπαιδεύουν τις ρομποτικές σκούπες, επισημαίνοντας φωτογραφίες από εσωτερικούς χώρους σπιτιών (Guo, 2023). Κενυάτισσες εκπαιδεύουν αλγορίθμους να αναγνωρίζουν βίαιο και αποτρόπαιο περιεχόμενο, ή επιβλαβείς εντολές, καθαρίζοντας τις εφαρμογές της OpenAI (Rowe, 2023). Εργαζόμενοι στη Μαδαγασκάρη εργάζονται ανά ομάδες σε δωμάτια σπιτιών για γαλλικές εταιρείες ΤΝ, οι οποίες αναπαράγουν τις ροές μεταφοράς εργασίας της εποχής της αποικιοκρατίας, αυτήν τη φορά μέσω της ψηφιακής δικτύωσης (Le Ludec et al., 2023).

Η κοινωνιολογική και ανθρωπολογική έρευνα γύρω από την ψηφιακή εργασία έχει φέρει στο προσκήνιο, εδώ και πολλά χρόνια, τις μορφές εργασίας «πίσω από την κουρτίνα» της αυτοματοποίησης. Η εργασία σε αυτά τα περιβάλλοντα γίνεται, συχνά, με το κομμάτι, απομακρυσμένα, εκτός των παραδοσιακών θεσμών που προστατεύουν την εργασία (Rani & Dhir, 2024). Όπως συμβαίνει και σε άλλους τομείς της ψηφιακής εργασίας και στην οικονομία της περιστασιακής απασχόλησης, το μοντέλο διάρθρωσης της παραγωγής βασίζεται στον αλγοριθμικό έλεγχο και την απουσία θεσμών διαιτησίας. Ενώ η οργάνωση της εργασίας χαρακτηρίζεται από την απουσία συλλογικών συμβάσεων και την παράκαμψη των νομικών πλαισίων ασφάλισης και ιατροφαρμακευτικής περίθαλψης των εργαζομένων.

Τα δεδομένα του ομικλώδους αυτού τοπίου δεν είναι σαφή. Οι εταιρείες ψηφιακών πλατφορμών αποκρύπτουν τα πλήθη των εργαζομένων που προσφέρουν στην παραγωγή της ΤΝ, πίσω από εταιρικά σύμφωνα εμπιστευτικότητας και πατέντες. Τα δίκτυα παραγωγής βασίζονται στην εξωτερική ανάθεση, περιπλέκοντας ακόμη περισσότερο τη λογοδοσία των εταιρειών, οι οποίες πολλές φορές αδυνατούν να γνωρίζουν τις ακριβείς συνθήκες εργασίας ή την τοποθεσία των εργαζομένων, αφού η διαχείρισή τους γίνεται αλγοριθμικά. Το διάσπαρτο πλήθος είναι πιθανό να εργάζεται σε περισσότερες από μία πλατφόρμες, χωρίς να έχει επαφή με την εταιρεία που αναθέτει την εργασία και, φυσικά, χωρίς να απολαμβάνει τους μισθούς και τα προνόμια των εργαζομένων της Silicon Valley.

Οι «υποβολείς» της ΤΝ βρίσκονται καλά κρυμμένοι στο κέντρο της σκηνής, μακριά από το οπτικό πεδίο των χρηστών των εφαρμογών. Ωστόσο, αν στοχαστούμε την αλληλεπίδρασή μας με αυτές, θα βρούμε πολλά κοινά ως προς τον τρόπο που η ανθρώπινη δραστηριότητα παράγει αξία για τα μοντέλα μηχανικής μάθησης. Εκτός από την αόρατη εργασία των εργαζομένων στις χώρες του παγκόσμιου Νότου, η εργασία των χρηστών είναι κομβικής σημασίας για τη λειτουργία αυτών των συστημάτων (Casilli, 2025). Σκεφτείτε πόσες φορές ασυναίσθητα έχουμε επιβραβεύσει ή επικρίνει μια απάντηση που δίνει ένα Μεγάλο Γλωσσικό Μοντέλο, πατώντας «thumbs up». Ή τις αμέτρητες φορές που η επίσκεψή μας σε μια ιστοσελίδα μάς παραπέμπει σε ένα σύνολο από εικόνες, από τις οποίες πρέπει να επιλέξουμε «πού υπάρχει μια γέφυρα». Σε κάθε μία από αυτές τις περιπτώσεις, η

αλληλεπίδρασή μας με τη μηχανή προσφέρει μια πληροφορία ή ενισχύει τη μάθηση του μοντέλου της. Τη στιγμή εκείνη, πραγματοποιούμε κάτι αντίστοιχο με τους εργαζομένους στη Βενεζουέλα ή στην Κένυα, χωρίς μάλιστα η εταιρεία που διανέμει το μοντέλο να πληρώσει δεκάρα.

Το επιχειρηματικό μοντέλο των ολιγοπωλίων της TN δημιουργεί περιβάλλοντα εργασίας εντός των οποίων πραγματοποιείται ο καταμερισμός της: Η μαζική παραγωγή και η επισημείωση δεδομένων ανατίθενται σε εργαζομένους ψηφιακών πλατφορμών, ενώ ένα άλλο τμήμα, αυτό που αφορά την ευθυγράμμιση του μοντέλου με τις ανάγκες των χρηστών, ανατίθεται στους χρήστες. Απέναντι στις ψηφιακές εφαρμογές όλοι είμαστε ίσοι, τμήματα μιας παραγωγικής διαδικασίας σε καθεστώς αλγοριθμικής διοίκησης. Η πλατφόρμα φέρνει εις πέρας ένα σύνθετο καθήκον: τη συγκέντρωση των ανθρώπινων δραστηριοτήτων και την ενσωμάτωσή τους στα μοντέλα μηχανικής μάθησης.

Τα αριθμητικά δεδομένα που διαθέτουμε για τον υπολογισμό των αλλαγών που φέρνει η TN στις αγορές εργασίας χαρακτηρίζονται από σημαντικές απώλειες. Οι άτυπες μορφές εργασίας στις ψηφιακές πλατφόρμες, τα ακατογράφητα μονοπάτια μεταφοράς εργασίας από χώρες του παγκόσμιου Νότου, αλλά και η περιφρόνηση προς τις σύντομες αλληλεπιδράσεις με την TN δημιουργούν μια μαύρη τρύπα στα στατιστικά μας. Η εργασία που βλέπουμε να εξαφανίζεται στις χώρες με ισχυρή μεσαία τάξη μετατοπίζεται σε διαφορετικά μήκη και πλάτη ή και σε διαφορετικά κοινωνικά στρώματα της ίδιας γεωγραφικής περιοχής. Αυτή η μετατόπιση συνοδεύεται από την «αορατοποίηση» της εργασίας, όχι με σκοπό να την εξαλείψει ή απλά να την αγνοήσει, αλλά να την αποζημιώσει λιγότερο. Την ίδια στιγμή που η TN συγκεντρώνει αμύθητα επενδυτικά ποσά, η εργασία που τη δημιουργεί συνθλίβεται και αποσπάται από τους θεσμούς προστασίας που είχαν θεσπιστεί τους προηγούμενους αιώνες.

Η παρούσα συνθήκη μάς προκαλεί να ξεφύγουμε από τον εγκλωβισμό της τεκνοφοβίας ή του τεχνολογικού οπτιμισμού και να αντιμετωπίσουμε το ζήτημα στη ρίζα του, κοινωνικά. Είναι δυνατή μια εκ νέου νοηματοδότηση των ανθρώπινων δραστηριοτήτων, έτσι ώστε αυτές να ενσωματώνονται στα μοντέλα της μηχανικής μάθησης με τρόπο διαφανή, αποδίδοντάς τους την αξία που τους αρμόζει; Είναι σίγουρο ότι η TN μάς φέρνει μπροστά από αυτό το ερώτημα, το οποίο για να απαντηθεί απαιτεί από εμάς να οραματιστούμε νέους τρόπους οργάνωσης όχι μόνο της παραγωγής, αλλά και της κοινωνίας. Απαιτεί πολιτικές αντίστασης, αλλά και δημιουργίας, κάτι που οι κοινωνίες μας έκαναν στο πρόσφατο παρελθόν.

Το 1945, στην παρουσίαση του πρώτου ψηφιακού υπολογιστή από τον στρατό των ΗΠΑ, κατέφθασαν δημοσιογράφοι, προκειμένου να καταγράψουν το νέο τεχνούργημα. Στις φωτογραφίες που υπάρχουν από εκείνη τη

μέρα, βλέπουμε άντρες αξιωματικούς να στέκονται μπροστά από τον κεντρικό υπολογιστή. Στο φόντο των φωτογραφιών υπήρχαν κάποιες γυναίκες. Ήταν εργαζόμενες, με σημαντικές σπουδές στα μαθηματικά και τη μηχανική, οι οποίες ήταν επιφορτισμένες να «μεταφράζουν» στη γλώσσα της μηχανής τις μαθηματικές εξισώσεις που έγραφαν οι άντρες που ηγούνταν του εγχειρήματος. Οι πρώτες γυναίκες «υπολογίστριες» ή προγραμματίστριες έλαβαν μέρος στον καταμερισμό εργασίας, γιατί οι άνδρες μηχανικοί θεωρούσαν υποδεέστερη εργασία τον προγραμματισμό, ο οποίος σε εκείνη την περίοδο απαιτούσε τη «βρόμικη δουλειά» ρύθμισης καλωδίων, λυχνιών και άλλων μηχανικών μερών του υπολογιστή (Τύμπας & Μεργούπη-Σαβαΐδου, 2021). Σήμερα η εργασία του προγραμματιστή και της προγραμματίστριας δεν βρίσκεται στο περιθώριο μιας φωτογραφίας, αλλά στο κέντρο της συζήτησης για την τεχνολογική καινοτομία, με αποτέλεσμα να αναγνωρίζεται και να αποζημιώνεται αναλόγως.

Τα ερωτήματα για το «τέλος της εργασίας» μετασχηματίζονται σε αναγκαίους προβληματισμούς για την ψηφιακή εργασία. Επανέρχεται στο προσκήνιο η κρισιμότητα της ανθρώπινης εργασίας την εποχή της ΤΝ. Η προστασία της εργασίας και η ανάδειξη της σημασίας της στην παραγωγική διαδικασία δεν είναι πρωτοφανές κοινωνικό φαινόμενο, σε αντίθεση με την τεχνολογία της ΤΝ. Έχουμε αρκετή εμπειρία ως κοινωνίες, την οποία οφείλουμε σε αυτήν τη φάση να ανακαλέσουμε και να επαναφέρουμε στον κοινωνικό και πολιτικό διάλογο.

Βιβλιογραφία

- Abril, D. (2026, April 6). Tech companies are cutting jobs and betting on AI. The payoff is far from guaranteed. *The Guardian*. <https://www.theguardian.com/technology/2026/apr/06/tech-layoffs-ai-work>
- Acemoglu, D., & Restrepo, P. (2018). Artificial intelligence, automation and work (NBER Working Paper No. 24196). *National Bureau of Economic Research*.
- Ashley, M. (2025, February 17). The future is solo: AI is creating billion-dollar one-person companies. *Forbes*. <https://www.forbes.com/sites/michaelashley/2025/02/17/the-future-is-solo-ai-is-creating-billion-dollar-one-person-companies/>
- Bridle, J. (2024, April 10). So, Amazon's 'AI-powered' cashier-free shops use a lot of... humans. Here's why that shouldn't surprise you. *The Guardian*. <https://www.theguardian.com/commentisfree/2024/apr/10/amazon-ai-cashier-less-shops-humans-technology>
- Brynjolfsson, E., Chandar, B., & Chen, R. (2025). *Canaries in the coal mine? Six facts about the recent employment effects of artificial intelligence* [Working paper]. Stanford Digital Economy Lab. <https://digitaleconomy.stanford.edu/publications/canaries-in-the-coal-mine/>
- Casilli, A. A. (2025). *Waiting for robots: The hired hands of automation*. University of Chicago Press.
- Davenport, T. H., & Srinivasan, L. (2026, January 29). Companies are laying off workers because of AI's potential—Not its performance. *Harvard Business Review*. <https://hbr.org/2026/01/companies-are-laying-off-workers-because-of-ais-potential-not-its-performance>
- Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114, 254-280. <https://doi.org/10.1016/j.techfore.2016.08.019>
- Girishankar, N., & Smith, K. (2025, September 16). *GenAI's human infrastructure challenge – Can the United States meet skilled trade labor*

demand through 2030? Center for Strategic and International Studies (CSIS). <https://www.csis.org/analysis/genais-human-infrastructure-challenge-can-united-states-meet-skilled-trade-labor-demand>

Gmyrek, P., Berg, J., & Bescond, D. (2023). *Generative AI and jobs: A global analysis of potential effects on job quantity and quality* [ILO Working Paper 96]. International Labour Organization. <https://doi.org/10.54394/FHEM8239>

Guo, E. (2023, February 13). A Roomba recorded a woman on the toilet. How did screenshots end up on Facebook? *MIT Technology Review*. <https://www.technologyreview.com/2022/12/19/1065306/roomba-irobot-robot-vacuums-artificial-intelligence-training-data-privacy/>

Kantor, J., Weise, K., & Yaffe-Bellany, D. (2025, Oktober 21). Amazon Plans to Replace More Than Half a Million Jobs With Robots. *The New York Times*. <https://www.nytimes.com/2025/10/21/technology/inside-amazons-plans-to-replace-workers-with-robots.html>

Khalaf, R. (2026, February 12). Mustafa Suleyman sets out Microsoft AI's goal of 'humanist superintelligence'. *Financial Times*. <https://www.ft.com/content/f1ec830c-2f08-4b1a-b70f-7330f260753c>

Le Ludec, C., Cornet, M., & Casilli, A. A. (2023). The problem with annotation. Human labour and outsourcing between France and Madagascar. *Big Data & Society*, 10(2). <https://doi.org/10.1177/20539517231188723>

Ostertag, M. (2025). *AI's job impact: Gains outpace losses*. Information Technology and Innovation Foundation (ITIF). <https://itif.org/publications/2025/12/18/ais-job-impact-gains-outpace-losses/>

Paternes, M. W., & Stirati A. (2023). The decoupling between labour compensation and productivity in high-income countries: Why is the nexus broken? *British Journal of Industrial Relations*, 61(2), 425-463. <https://doi.org/10.1111/bjir.12713>

Rani, U., & Dhir, R. K. (2024). AI-enabled business model and human-in-the-loop (deceptive AI): implications for labor. In *Handbook of artificial intelligence at work* (pp. 47-75). Edward Elgar Publishing.

Rifkin, J. (1995). *The end of work: The decline of the global labor force and the dawn of the post-market era*. G. P. Putnam's Sons.

- Rowe, N. (2023, August 2). 'It's destroyed me completely': Kenyan moderators decry toll of training of AI models. *The Guardian*. <https://www.theguardian.com/technology/2023/aug/02/ai-chatbot-training-human-toll-content-moderator-meta-openai>
- Schwellnus, C., Kappeler, A. & Pionnier, P. (2017). *Decoupling of wages from productivity: Macro-level facts* [OECD Economics Department Working Papers, No. 1373]. OECD Publishing. <https://doi.org/10.1787/d4764493-en>.
- SonarSource (2025). *State of code: Developer survey report*. <https://www.sonarsource.com/state-of-code-developer-survey-report.pdf>
- Srinivasan, S. (2026). *Enhance or Eliminate? How AI Will Likely Change These Jobs*. Harvard Business School Working Knowledge. <https://hbswk.hbs.edu/item/enhance-or-eliminate-how-ai-will-likely-change-these-jobs>
- Tubaro, P., Casilli, A. A., & Coville, M. (2020). The trainer, the verifier, the imitator: Three ways in which human platform workers support artificial intelligence. *Big Data & Society*, 7(1). <https://doi.org/10.1177/2053951720919776>
- Τύμπας, Τ., & Μεργούπη-Σαβαΐδου, Ε. (Επιμ.) (2021). *Ιστορίες της τεχνολογίας του εικοστού αιώνα: Ηλεκτρικά αυτοκίνητα, ξύλινα αεροπλάνα, γαλλικοί αντιδραστήρες, γυναίκες υπολογιστές*. Πανεπιστημιακές Εκδόσεις Κρήτης.

8.2. Ευκαιρίες και κίνδυνοι της χρήσης Τεχνητής Νοημοσύνης στην αγορά εργασίας

Όπως επισημαίνεται σε σχετική μελέτη του ΟΟΣΑ, η Τεχνητή Νοημοσύνη διαφέρει από τις προηγούμενες τεχνολογίες αναφορικά με τις επιπτώσεις στην εργασία τουλάχιστον σε τρεις σημαντικές πτυχές:

α) Επεκτείνει τα είδη των εργασιών που μπορούν να αυτοματοποιηθούν και σε πολλές μη συστηματικές γνωσιακές εργασίες και, επομένως, εκθέτει εργαζομένους που μέχρι πρότινος ήταν σχετικά προστατευμένοι από την αυτοματοποίηση (π.χ. τους υψηλής ειδίκευσης εργαζομένους) στους κινδύνους του εκτοπισμού.

β) Υφίσταται ο κίνδυνος επηρεασμού όλων των επαγγελμάτων και των τομέων, σε αντίθεση, π.χ., με την εισαγωγή των ρομπότ που επηρέασαν κυρίως τον μεταποιητικό τομέα.

γ) Υφίσταται ο κίνδυνος αύξησης της προσωρινής ανεργίας (frictional unemployment), καθώς η ταχύτητα ανάπτυξης και υιοθέτησης της ΤΝ στην αγορά εργασίας αφήνει ελάχιστο χρόνο προσαρμογής.²⁰²

Συνακόλουθα οι εργαζόμενοι ενδέχεται να αντιμετωπίσουν σοβαρά προβλήματα εξαιτίας της εισαγωγής συστημάτων ΤΝ. Η έκθεση στους κινδύνους αυτούς ποικίλλει σε ένταση ανάλογα με τις δεξιότητές τους, το επάγγελμά τους ή το μέγεθος της επιχείρησης στην οποία εργάζονται. Οι παράγοντες αυτοί φαίνεται να επηρεάζουν και τον βαθμό έκθεσής τους στον κίνδυνο μεροληψίας και διακρίσεων, αλλά και παραβιάσεων της ιδιωτικότητας. Σύμφωνα με την έκθεση του ΟΟΣΑ, τα σχετικά στοιχεία δείχνουν, επίσης, ότι η ΤΝ μπορεί, σε ορισμένα επαγγέλματα, να αυξήσει την παραγωγικότητα των εργαζομένων χαμηλής ειδίκευσης, μειώνοντας τα χάσματα παραγωγικότητας σε σχέση με τους εργαζομένους υψηλότερης ειδίκευσης, οδηγώντας

²⁰² Ο ΟΟΣΑ εκτιμούσε προ διετίας ότι τα επαγγέλματα που διατρέχουν τον υψηλότερο κίνδυνο αυτοματοποίησης αντιστοιχούν περίπου στο 27% της συνολικής απασχόλησης. Θα είναι σημαντικό να υποστηριχθούν οι εργαζόμενοι ώστε να μετακινηθούν από τομείς και επαγγέλματα που παρακμάζουν προς νέους και αναπτυσσόμενους τομείς και επαγγέλματα. Βλ. Organisation for Economic Co-operation and Development (OECD). (2024). *Using AI in the workplace: Opportunities, risks and policy responses* (OECD Artificial Intelligence Papers, No. 11). <https://doi.org/10.1787/73d417f9-en>

σε νέες αλλαγές στον καταμερισμό εργασίας και στις δυνατότητες εισόδου στην αντίστοιχη αγορά.

Η υιοθέτηση της ΤΝ επιφέρει σημαντικές μεταβολές στις δεξιότητες, αλλά και στις απαιτήσεις για νέες δεξιότητες. Σε σχετική έρευνα του ΟΟΣΑ αναφέρεται ότι, ενώ οι εργοδότες δηλώνουν ότι η ΤΝ έχει αυξήσει τη σημασία των εξειδικευμένων δεξιοτήτων στην ΤΝ, υποστηρίζουν ότι έχει αυξήσει ακόμη περισσότερο τη σημασία των ανθρώπινων δεξιοτήτων και την ανάγκη για εργαζομένους υψηλού μορφωτικού επιπέδου. Είναι συνεπώς πιθανό να απαιτηθεί ένα ευρύ φάσμα δεξιοτήτων, καθώς η ΤΝ καθίσταται ολοένα και πιο διάχυτη στην οικονομία.²⁰³ Τα ζητήματα αυτά αναλύει η Κλειώ Σγουροπούλου στη μελέτη της.

Η καινοτομία δεν παράγεται σε θεσμικό, πολιτικό, κοινωνικο-οικονομικό κενό και η επιδίωξη της ευημερίας δεν μπορεί παρά να περιλαμβάνει πολιτικές που διασφαλίζουν ότι η νέα οικονομική και εργασιακή πραγματικότητα δεν διευρύνει το ψηφιακό, κοινωνικό και οικονομικό χάσμα, και δεν επιτείνει τις ανισότητες. Η εργασία είναι μια πύλη εισόδου σε ένα ολόκληρο σύστημα δικαιωμάτων, υποχρεώσεων, προστασίας και πρόνοιας.

Η ραγδαία διείσδυση της ΤΝ στην αγορά εργασίας, αλλά και στην καθημερινότητα έχει καταστήσει την απόκτηση δεξιοτήτων χρήσης της έναν νέο και καθοριστικό παράγοντα κοινωνικής διαστρωμάτωσης. Όσοι διαθέτουν πρόσβαση σε εκπαίδευση, τεχνολογικές υποδομές και ψηφιακό γραμματισμό αποκτούν ανταγωνιστικά πλεονεκτήματα που ενισχύουν την επαγγελματική και οικονομική τους θέση. Αντίθετα, όσοι στερούνται αυτές τις γνώσεις και δεξιότητες κινδυνεύουν να περιθωριοποιηθούν σε έναν κόσμο που εξελίσσεται με ταχύτατους ρυθμούς. Ο κίνδυνος αυτός δεν αφορά μόνο τους οικονομικά αδύναμους, αλλά έχει γεωγραφικές, ηλικιακές, φυλετικές και μορφωτικές διαστάσεις, διαμορφώνοντας ένα πολυεπίπεδο ψηφιακό χάσμα που απειλεί να αναπαράγει, να εμβαθύνει και να εντείνει προϋπάρχουσες ανισότητες. Η εκπόνηση πολιτικών για τις δεξιότητες δεν συνιστά μόνο αναγκαία και εκ των ων ουκ άνευ συνθήκη για την αποδοτική χρήση της ΤΝ, αλλά και απαραίτητη προϋπόθεση για να διασφαλιστούν η κοινωνική συνοχή και η συνολική ευημερία.

²⁰³ Lane, M., Williams, M., & Broecke, S. (2023). *The impact of AI on the workplace: Main findings from the OECD AI surveys of employers and workers* (OECD Social, Employment and Migration Working Papers, No. 288). OECD Publishing. <https://doi.org/10.1787/ea0a0fe1-en>

8.3. Από την αυτοματοποίηση στην επαύξηση: Η Τεχνητή Νοημοσύνη ως καταλύτης δομικού μετασχηματισμού της εργασίας και των δεξιοτήτων στην Ελλάδα

Η ελληνική αγορά εργασίας βρίσκεται μπροστά σε μια δομική μεταβολή, καθώς η Τεχνητή Νοημοσύνη μετασχηματίζει την παραγωγή και επιβάλλει ένα νέο εκπαιδευτικό μοντέλο για τη χώρα. Μέσα από το πρίσμα της δημιουργικής καταστροφής, αναλύεται η μετάβαση από την απλή ψηφιοποίηση στην αυτοματοποίηση γνωστικών καθηκόντων ανώτερου επιπέδου, αναδεικνύοντας το λεγόμενο «Ελληνικό Παράδοξο»: τη δυσαρμονία μεταξύ του υψηλού επιστημονικού κεφαλαίου και της χαμηλής ψηφιακής ωριμότητας των εγχώριων επιχειρήσεων. Μέσα από το πρίσμα των εργασιακών καθηκόντων (task-based approach), διακρίνονται τρεις βασικές επιδράσεις: η υποκατάσταση τυποποιημένων εργασιών, η επαύξηση της απόδοσης σε σύνθετα αντικείμενα και η δημιουργία ζήτησης για νέες, υβριδικές ειδικότητες.

Ιδιαίτερη έμφαση δίνεται στο αναδυόμενο χάσμα δεξιοτήτων (skills gap), το οποίο καθιστά αναγκαία τη μετατόπιση από τη στατική τεχνική γνώση προς έναν συνδυασμό αλγοριθμικού γραμματισμού (AI Literacy) και ανθρωποκεντρικών δεξιοτήτων. Η εξέλιξη αυτή αποτελεί βασική προϋπόθεση για τη διασφάλιση της κοινωνικής συνοχής και της οικονομικής ανθεκτικότητας στην εποχή της Τεχνητής Νοημοσύνης. Παράλληλα, εξετάζονται τα δομικά εμπόδια, όπως η δημογραφική γήρανση και η έλλειψη ψηφιακής κυριαρχίας στα δεδομένα, καθώς και οι ηθικοί κίνδυνοι της αλγοριθμικής μεροληψίας. Μια δέσμη προτάσεων πολιτικής εστιάζει στη διεπιστημονική αναδιάρθρωση της Τριτοβάθμιας Εκπαίδευσης, στην υιοθέτηση μικρο-πιστοποιήσεων (micro-credentials) και στον μετασχηματισμό των δημόσιων υπηρεσιών απασχόλησης σε προγνωστικούς μηχανισμούς κατάρτισης. Οι προτάσεις αυτές εναρμονίζονται με τους Άξονες Ανάπτυξης της εθνικής στρατηγικής, η οποία προκρίνει τον συμμετοχικό σχεδιασμό και τη δημιουργία ενός Εθνικού Οικοσυστήματος Ψηφιακών Ικανοτήτων. Η επιτυχία της μετάβασης εξαρτάται από τη μετατόπιση προς την «Ψηφιακή Ηγεσία» (Digital Leadership) και τη «Διαχείριση Αλλαγών» (Change Management), ειδικά για τα στελέχη του δημόσιου και του ιδιωτικού τομέα (MDC, 2022). Στόχος είναι η διαμόρφωση

ενός νέου «κοινωνικού συμβολαίου» που θα διασφαλίζει ότι η τεχνολογική πρόοδος θα λειτουργήσει ως συμπεριληπτικός επιταχυντής ευημερίας για την ελληνική κοινωνία.

8.3.1. Εισαγωγή

Η εμφάνιση και η ραγδαία εξάπλωση της ΤΝ δεν συνιστά απλώς μια γραμμική εξέλιξη στην ιστορία της Πληροφορικής, αλλά μια θεμελιώδη μετατόπιση του τεχνολογικού και οικονομικού παραδείγματος (Acemoglu et al., 2022· Autor, 2015). Ενώ οι προηγούμενες φάσεις του ψηφιακού μετασχηματισμού επικεντρώθηκαν στην αυτοματοποίηση της αποθήκευσης και της μεταφοράς πληροφοριών, η έλευση της Παραγωγικής ΤΝ (Generative AI) και των Μεγάλων Γλωσσικών Μοντέλων μεταφέρει την αυτοματοποίηση στο πεδίο της ανώτερης γνωστικής λειτουργίας. Σε μια χώρα όπως η Ελλάδα, όπου η ενίσχυση της παραγωγικότητας και η ψηφιακή σύγκλιση αποτελούν κεντρικούς εθνικούς στόχους, χαρτογραφείται το νέο τοπίο και προσδιορίζονται οι ευκαιρίες και οι κίνδυνοι που αναδύονται στην αγορά εργασίας.

Στο επίκεντρο βρίσκεται η «Δημιουργική καταστροφή» του Joseph Schumpeter, όπως αυτή εκδηλώνεται στις σύγχρονες συνθήκες της ψηφιακής και αλγοριθμικής εποχής. Η ΤΝ δεν «κλέβει» απλώς θέσεις εργασίας αναδιατάσσει τη δομή των επαγγελματών, αποσυνθέτοντάς τα σε επιμέρους καθήκοντα (tasks) (Arntz et al., 2016). Αυτή η διαφοροποίηση είναι κρίσιμη για τη χάραξη πολιτικής: Αν αντιμετωπίσουμε την ΤΝ ως μια εξωγενή απειλή που πρέπει να περιοριστεί, κινδυνεύουμε με οικονομικό μαρασμό. Αν, αντιθέτως, την αντιληφθούμε ως έναν πολλαπλασιαστή ισχύος του ανθρώπινου κεφαλαίου, η Ελλάδα μπορεί να καλύψει το χαμένο έδαφος δεκαετιών.

Κεντρική επιδίωξη αποτελεί η εξέταση των μηχανισμών μέσω των οποίων η ΤΝ επιδρά στην απασχόληση, καθώς και ο προσδιορισμός των αναγκών προσαρμογών για την αναθεώρηση του εκπαιδευτικού μοντέλου και των πολιτικών επιμόρφωσης. Η μετάβαση απαιτεί ένα νέο «κοινωνικό συμβόλαιο», όπου η τεχνολογική πρόοδος δεν θα οδηγεί σε κοινωνικό αποκλεισμό, αλλά σε μια νέα εποχή ενισχυμένης ανθρώπινης δημιουργικότητας.

8.3.2. Παγκόσμιες τάσεις και το «Ελληνικό Παράδοξο»

Σε παγκόσμιο επίπεδο, η ΤΝ αναμένεται να επηρεάσει έως και το 40% των θέσεων εργασίας παγκοσμίως (OECD, 2024). Πρόσφατα δεδομένα υπογραμμίζουν ότι η αλλαγή δεν αφορά μόνο την κατάργηση θέσεων, αλλά τον μετασχηματισμό τους: Κατά μέσο όρο, το 1/3 των δεξιοτήτων που απαιτούνται για μια θέση εργασίας άλλαξε μέσα στην τελευταία τριετία (2021-2024). Μάλιστα, για το 25% των επαγγελματών με τη μεγαλύτερη αναστάτωση, το ποσοστό αλλαγής των απαιτούμενων δεξιοτήτων αγγίζει το εντυπωσιακό 75% (Lightcast, 2025). Στις προηγμένες οικονομίες, το ποσοστό αυτό αγγίζει

το 60%, καθώς η ΤΝ επηρεάζει κυρίως θέσεις εργασίας υψηλής εξειδίκευσης. Στην Ελλάδα, ωστόσο, παρατηρείται αυτό που ονομάζουμε «Ελληνικό Παράδοξο»: Ενώ διαθέτουμε ένα εξαιρετικά καταρτισμένο δυναμικό επιστημόνων Πληροφορικής και μηχανικών, η απορρόφηση αυτών των δεξιοτήτων από την εγχώρια αγορά παραμένει δυσανάλογα χαμηλή (OECD, 2025c).

Η ελληνική οικονομία χαρακτηρίζεται από έναν κατακερματισμένο επιχειρηματικό ιστό, με συντριπτική πλειονότητα ΜμΕ, οι οποίες συχνά στερούνται τις υποδομές και τα κεφάλαια για να ενσωματώσουν την ΤΝ στον πυρήνα της λειτουργίας τους. Αυτό δημιουργεί έναν κίνδυνο ψηφιακής απόκλισης (digital divergence). Ενώ οι μεγάλες πολυεθνικές και οι εξωστρεφείς ελληνικές εταιρείες ήδη αποκομίζουν οφέλη από τη χρήση αλγορίθμων στη λήψη αποφάσεων και στην εφοδιαστική αλυσίδα, η μέση επιχείρηση κινδυνεύει να μείνει εκτός ανταγωνισμού.

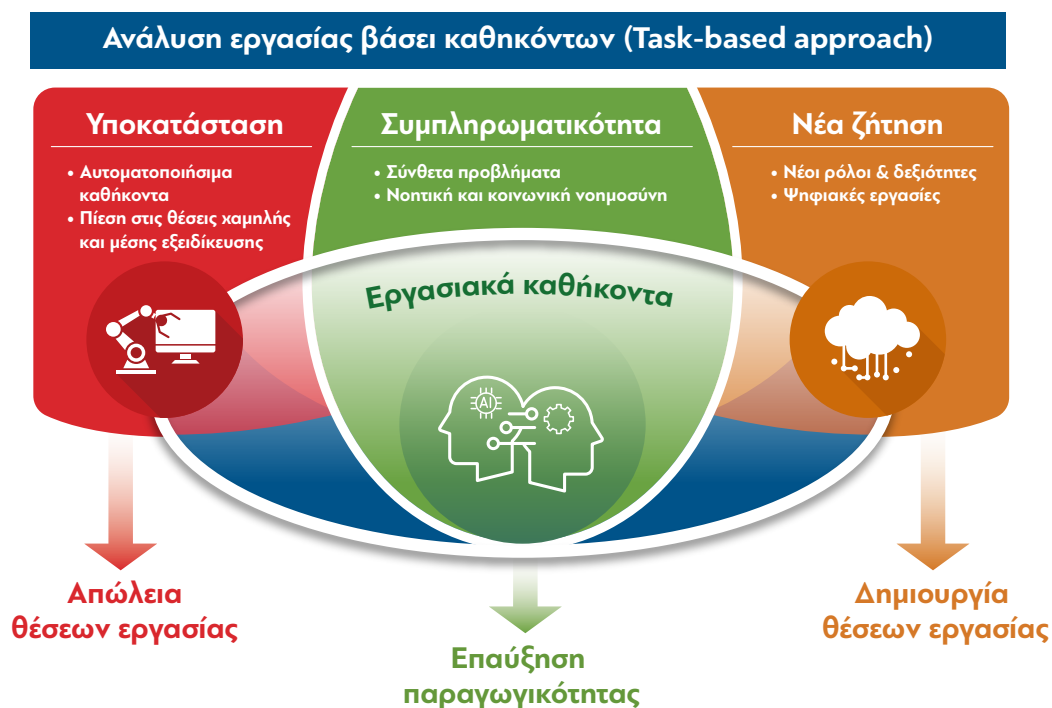
Επιπλέον, η κλαδική διάρθρωση της ελληνικής οικονομίας παρουσιάζει ιδιαιτερότητες. Ο Τουρισμός και η Ναυτιλία, δύο σημαντικοί πυλώνες, βρίσκονται σε διαφορετικά στάδια ψηφιακής ωριμότητας. Η Ναυτιλία έχει ήδη ξεκινήσει τη χρήση ΤΝ για τη βελτιστοποίηση της κατανάλωσης καυσίμων και την αυτόνομη ναυσιπλοΐα, ενώ ο Τουρισμός βρίσκεται ακόμα στα αρχικά στάδια της εξατομικευμένης εμπειρίας πελάτη μέσω ΤΝ. Η οριζόντια εφαρμογή της ΤΝ στη δημόσια διοίκηση αναδεικνύεται ως καταλύτης που μπορεί να απελευθερώσει πόρους και να μειώσει δραστικά τη γραφειοκρατία, ενισχύοντας συνολικά το επιχειρηματικό περιβάλλον και την εθνική ανταγωνιστικότητα. Στο πλαίσιο αυτό, η Ελλάδα έχει ήδη θεσμοθετήσει την «Εθνική Στρατηγική για την Ανάπτυξη των Ψηφιακών Ικανοτήτων», η οποία ευθυγραμμίζεται με την ευρωπαϊκή «Ψηφιακή Δεκαετία 2030». Ο στρατηγικός αυτός σχεδιασμός θέτει ως στόχο το 80% των ενηλίκων να διαθέτει βασικές ψηφιακές δεξιότητες έως το 2030, ενώ δίνει έμφαση στην ενσωμάτωση αναδυόμενων τεχνολογιών, όπως η ΤΝ και το blockchain στον δημόσιο και τον ιδιωτικό τομέα (MDG, 2022).

Η πρόκληση για την Ελλάδα δεν είναι μόνο η υιοθέτηση της τεχνολογίας, αλλά η διατήρηση του ταλέντου. Η «διαρροή εγκεφάλων» (brain drain) της προηγούμενης δεκαετίας στέρψε από τη χώρα το κεφάλαιο που σήμερα θα πρωτοστατούσε στην επανάσταση της ΤΝ. Η δημιουργία ενός ευνοϊκού οικοσυστήματος για την ΤΝ μπορεί να λειτουργήσει ως κίνητρο για τον επαναπατρισμό αυτών των επιστημόνων, μετατρέποντας τη «διαρροή εγκεφάλων» σε «προσέλκυση εγκεφάλων» (brain gain). Κομβικό ρόλο σε αυτήν τη μετάβαση διαδραματίζουν οι Ευρωπαϊκοί Κόμβοι Ψηφιακής Καινοτομίας (European Digital Innovation Hubs – EDIHs), που λειτουργούν ως επιταχυντές για την υιοθέτηση της ΤΝ από τις ΜμΕ, καθώς και η δημιουργία του ελληνικού εργοστασίου ΤΝ «Pharos», το οποίο αναμένεται να προσφέρει την απαραίτητη υπολογιστική ισχύ και τεχνογνωσία για την ανάπτυξη εγχώριων λύσεων υψηλής κλίμακας.

8.3.3. Η ανατομία της αγοράς εργασίας – Υποκατάσταση, συμπληρωματικότητα και η νέα ζήτηση

Η επίδραση της Τεχνητής Νοημοσύνης στην απασχόληση δεν μπορεί να γίνει κατανοητή μέσα από μια απλοϊκή προσέγγιση απώλειας θέσεων εργασίας. Η ανάλυση βασίζεται στην προσέγγιση των καθηκόντων (task-based approach), η οποία αποτελεί το καθιερωμένο πλαίσιο για την εξέταση των επιπτώσεων της αυτοματοποίησης στην εργασία (Autor, 2015· Arntz et al., 2016). Σύμφωνα με αυτό, κάθε επάγγελμα αποτελείται από ένα σύνολο δραστηριοτήτων που μπορούν να ταξινομηθούν σε τρεις κατηγορίες: τις αυτοματοποιήσιμες, τις συμπληρωματικές και τις δημιουργικές.

Σχήμα 1. Η επίδραση της Τεχνητής Νοημοσύνης στην αγορά εργασίας



Πηγή: Επεξεργασία συγγραφικής ομάδας.

Το φαινόμενο της υποκατάστασης (substitution effect)

Η ΤΝ υπερέχει σε καθήκοντα που χαρακτηρίζονται από υψηλή τυποποίηση και βασίζονται στην επεξεργασία δεδομένων. Στην ελληνική αγορά, αυτό μεταφράζεται σε άμεση πίεση προς επαγγέλματα γραφείου χαμηλής και μέσης εξειδίκευσης. Αντιθέτως, τα επαγγέλματα που βασίζονται σε χειρωνακτική εργασία ή έχουν χαμηλότερες εκπαιδευτικές απαιτήσεις (όπως μάγειρες, εργάτες αποθήκης ή τεχνίτες ξηράς δόμησης) παρουσιάζουν τη

μικρότερη μεταβολή δεξιοτήτων, με το ποσοστό αλλαγής να περιορίζεται στο 8% (Lightcast, 2025). Η δυνατότητα των μοντέλων ΤΝ να εκτελούν εργασίες όπως η καταχώριση δεδομένων (Pouliakas, 2021), η βασική σύνταξη κειμένων και η πρωτοβάθμια εξυπηρέτηση πελατών σημαίνει ότι οι θέσεις εργασίας που περιορίζονται σε αυτές τις δραστηριότητες διατρέχουν υψηλό κίνδυνο. Στην ελληνική αγορά, η υποκατάσταση εργασιών αναμένεται να εξελιχθεί με πιο ήπιους ρυθμούς, καθώς οι ΜμΕ ενσωματώνουν σταδιακά τις νέες τεχνολογίες. Αυτό δημιουργεί ένα πολύτιμο περιθώριο χρόνου για την ομαλή προσαρμογή των εργαζομένων και την αναβάθμιση των δεξιοτήτων τους.

Η επαύξηση και η συμπληρωματικότητα (augmentation effect)

Εδώ εντοπίζεται η μεγαλύτερη ευκαιρία για την εθνική οικονομία. Η ΤΝ λειτουργεί ως συμπληρωματικός παράγοντας για εργαζομένους που η εργασία τους απαιτεί επίλυση σύνθετων προβλημάτων, ηθική κρίση ή υψηλή κοινωνική νοημοσύνη (Acemoglu et al., 2022). Για έναν Έλληνα μηχανικό ή έναν ιατρό, η ΤΝ δεν είναι ανταγωνιστής, αλλά ένας νοητικός βοηθός που αναλαμβάνει τον όγκο της ανάλυσης δεδομένων, επιτρέποντας στον άνθρωπο να εστιάσει στη λήψη αποφάσεων. Η συμπληρωματικότητα αυξάνει την οριακή παραγωγικότητα της εργασίας (Mäkelä & Stephany, 2024), γεγονός που θεωρητικά οδηγεί σε αύξηση των μισθών για όσους κατέχουν τις κατάλληλες δεξιότητες χρήσης αυτών των εργαλείων.

Η δημιουργία νέας ζήτησης (reinstatement effect)

Ιστορικά, κάθε τεχνολογική επανάσταση δημιούργησε επαγγέλματα που προηγουμένως δεν μπορούσαμε να φανταστούμε. Στην Ελλάδα αναδύεται μια νέα κατηγορία εργαζομένων: οι «διαμεσολαβητές ΤΝ». Πρόκειται για επαγγελματίες που γεφυρώνουν το χάσμα ανάμεσα στα αλγοριθμικά αποτελέσματα και στην επιχειρηματική πράξη. Επιπλέον, η αυξανόμενη ανάγκη για επισημείωση δεδομένων (data labeling), έλεγχο ηθικής των αλγορίθμων (AI ethics auditing) και κυβερνοασφάλεια (cybersecurity) θα δημιουργήσει μια νέα ψηφιακή ζήτηση που μπορεί να απορροφήσει μέρος του δυναμικού που θα εκτοπιστεί από παραδοσιακούς κλάδους. Ειδικά στον τομέα της κυβερνοασφάλειας, η ζήτηση διαχέεται πλέον οριζόντια: Επαγγέλματα όπως οι UI/UX Designers και οι Analytics Managers έχουν δει τις απαιτήσεις για γνώσεις ασφάλειας δεδομένων να αυξάνονται κατά 256% και 135% αντίστοιχα, καθώς η προστασία της πληροφορίας καθίσταται δομικό στοιχείο κάθε επιχειρηματικής λειτουργίας (Lightcast, 2025).

8.3.4. Το ζήτημα των δεξιοτήτων και το χάσμα της γνώσης (skills gap)

Η μετάβαση στην οικονομία της ΤΝ αναδεικνύει το ζήτημα των δεξιοτήτων ως την κρίσιμότερη παράμετρο της εθνικής επιτυχίας. Συντελείται μια ριζική

επανεκτίμηση της εξειδίκευσης, καθώς η αξία μετατοπίζεται πλέον από τη στατική τεχνική γνώση στην κριτική κατανόηση. Σε έναν κόσμο όπου η πληροφορία είναι άφθονη και η επεξεργασία της αυτοματοποιημένη, η αξία μετατοπίζεται από τη γνώση του «πώς» (know-how) στη γνώση του «γιατί» (know-why).

Η απαξίωση των «σκληρών δεξιοτήτων»

Παρατηρείται το φαινόμενο της ταχείας απαξίωσης (skills obsolescence) (OECD, 2025b). Τεχνικές δεξιότητες που απαιτούσαν χρόνια εκμάθησης (π.χ. συγγραφή βασικού κώδικα, μετάφραση τεχνικών κειμένων, λογιστική ανάλυση) μπορούν πλέον να εκτελεστούν από την ΤΝ σε δευτερόλεπτα. Αυτό δημιουργεί μια τεράστια πρόκληση για το ελληνικό εκπαιδευτικό σύστημα, το οποίο παραμένει προσανατολισμένο στην παροχή στατικών, τεχνικών γνώσεων.

Η αρχιτεκτονική των «υβριδικών δεξιοτήτων»

Ως απάντηση στις νέες προκλήσεις μπορεί να διαμορφωθεί μια «τριάδα ψηφιακής ανθεκτικότητας» (OECD, 2025b· OECD, 2025c), ένας συνδυασμός ικανοτήτων που θωρακίζει τον εργαζόμενο στην αλγοριθμική εποχή:

α) Αλγοριθμική Παιδεία (AI Literacy): Η ικανότητα κατανόησης των περιορισμών και των μεροληψιών της ΤΝ.

β) Εργασιακή Ευελιξία (Adaptability): Η ικανότητα του εργαζομένου να μεταπηδά ανάμεσα σε διαφορετικά εργαλεία και περιβάλλοντα εργασίας.

γ) Ανθρωπιστική Υπεροχή (Critical Thinking): Δεξιότητες όπως η ενσυναίσθηση, η ηθική συλλογιστική και η διαπραγμάτευση – πεδία όπου η ΤΝ, παρά τις προόδους της, παραμένει μια στατιστική προσομοίωση χωρίς πραγματική κατανόηση.

Οι παραπάνω δεξιότητες δεν συγκροτούν ένα αυτόνομο νέο επάγγελμα, αλλά διαχέονται οριζόντια σε ένα ευρύ φάσμα ρόλων. Η ελληνική Πολιτεία προκρίνει τη χρήση μικρο-διαπιστευτηρίων (micro-credentials) και ανοιχτών ετικετών ποιότητας (open badges) ως εργαλεία για την πιστοποίηση αυτών των υβριδικών ικανοτήτων, επιτρέποντας στους εργαζομένους να χτίζουν εξατομικευμένα μονοπάτια μάθησης. Κεντρικό ρόλο σε αυτήν τη διαδικασία παίζει η Εθνική Ακαδημία Ψηφιακών Ικανοτήτων (ΕΑΨΙ), η οποία λειτουργεί ως πύλη για την ανάπτυξη δεξιοτήτων βάσει συγκεκριμένων επαγγελματικών προφίλ (MDG, 2025). Παράλληλα με την αλγοριθμική παιδεία, αναδύεται η ανάγκη για πράσινες δεξιότητες (green skills). Η μετάβαση στην καθαρή ενέργεια έχει οδηγήσει σε αύξηση κατά 31% της ζήτησης για τέτοιες δεξιότητες, οι οποίες πλέον απαιτούνται ακόμη και σε

μη τεχνικούς ρόλους, όπως οι αναλυτές δεδομένων (data scientists), που καλούνται να διαχειριστούν την ενεργειακή αποδοτικότητα των συστημάτων τους (Lightcast, 2025). Αναδύεται έτσι η ανάγκη για επαγγελματίες που λειτουργούν ως «διαμεσολαβητές» μεταξύ αλγοριθμικών συστημάτων και ανθρώπινης κρίσης, μεταφράζοντας στατιστικές προβλέψεις σε επιχειρησιακές, νομικές ή κοινωνικά αποδεκτές αποφάσεις.

Σχήμα 2. Η τριάδα της ψηφιακής ανθεκτικότητας



Πηγή: Επεξεργασία συγγραφικής ομάδας.

Το χάσμα στην ελληνική αγορά

Το πρόβλημα στην Ελλάδα εντείνεται από το γεγονός ότι οι μεγάλες ηλικιακές ομάδες του εργατικού δυναμικού (45-64 ετών) παρουσιάζουν χαμηλά επίπεδα ψηφιακού γραμματισμού. Τα εθνικά δεδομένα επιβεβαιώνουν αυτήν την απόκλιση: Ενώ η Ελλάδα είναι πρωτοπόρος στην ηλικιακή ομάδα 16-24 ετών, με το 88% των νέων να κατέχει βασικές ψηφιακές δεξιότητες (έναντι 71% του μέσου όρου της ΕΕ), η συνολική κατάταξη της χώρας στο ανθρώπινο κεφάλαιο παραμένει κάτω από τον μέσο όρο (MDG, 2025). Η στρατηγική απάντηση εστιάζει πλέον στην έννοια της ικανότητας ως συνδυασμό γνώσεων, δεξιοτήτων και στάσεων, και ευθυγραμμίζεται με το ευρωπαϊκό πλαίσιο DigComp 2.3 (MDG, 2022· OECD, 2025c). Αν δεν υπάρξει άμεση παρέμβαση, η ΤΝ θα λειτουργήσει ως παράγοντας ηλικιακών και κοινωνικών διακρίσεων, αποκλείοντας ένα μεγάλο μέρος του

πληθυσμού από τις θέσεις εργασίας υψηλής ποιότητας. Το χάσμα αυτό δεν αφορά αποκλειστικά την έλλειψη τεχνικών γνώσεων, αλλά κυρίως την αδυναμία μετασχηματισμού της συσσωρευμένης επαγγελματικής εμπειρίας σε αξία ενισχυμένη από την ΤΝ, γεγονός που καθιστά την επανειδίκευση περισσότερο γνωστική και πολιτισμική παρά τεχνολογική πρόκληση.

Η ανάλυση των δεξιοτήτων δεν αρκεί από μόνη της. Η ελληνική περίπτωση αναδεικνύει μια σειρά δομικών και ηθικών προκλήσεων που, αν αγνοηθούν, μπορούν να ακυρώσουν τα αναπτυξιακά οφέλη της ΤΝ.

8.3.5. Προβλήματα, δομικές αδυναμίες και ηθικές προκλήσεις

Η ενσωμάτωση της ΤΝ στην Ελλάδα δεν προσκρούει μόνο σε τεχνικά εμπόδια, αλλά σε βαθιά ριζωμένες δομικές αδυναμίες που απειλούν να μετατρέψουν μια ευκαιρία ανάπτυξης σε παράγοντα κοινωνικής αποσταθεροποίησης. Στο επίκεντρο της ανάλυσης τίθενται τρεις κεντρικοί άξονες προβληματισμού: το δημογραφικό-ψηφιακό χάσμα, η ανάγκη για περαιτέρω ανάπτυξη υποδομών δεδομένων και οι ηθικές/νομικές προκλήσεις που προκύπτουν από τη χρήση της ΤΝ.

Το δημογραφικό και το ψηφιακό χάσμα

Η Ελλάδα διαθέτει ένα από τα πλέον γηράσκοντα εργατικά δυναμικά στην Ευρώπη. Η δημογραφική αυτή κρίση συνιστά μια επερχόμενη καταιγίδα, καθώς οι εργαζόμενοι μεγαλύτερης ηλικίας (Baby Boomers) αποχωρούν μαζικά από την αγορά εργασίας. Χαρακτηριστικά, στις ΗΠΑ το 80% όσων εγκατέλειψαν το εργατικό δυναμικό μετά το 2020 ήταν άνω των 55 ετών, εντείνοντας τις ελλείψεις σε κρίσιμους κλάδους (Lightcast, 2025). Η εισαγωγή της ΤΝ δημιουργεί ένα χάσμα δύο ταχυτήτων. Από τη μία πλευρά, οι νεότεροι εργαζόμενοι (digital natives) που προσαρμόζονται ταχύτατα και, από την άλλη, μια πολυπληθής ομάδα εργαζομένων άνω των 45-50 ετών, οι οποίοι κατέχουν την εμπειρία, αλλά στερούνται της ψηφιακής ευελιξίας. Χωρίς στοχευμένη κρατική μέριμνα, η ΤΝ θα οδηγήσει σε έναν βίαιο ψηφιακό εξοστρακισμό των μεγαλύτερων ηλικιακά εργαζομένων, αυξάνοντας την ανάγκη για πρόωρες συνταξιοδοτήσεις και επιβαρύνοντας το ασφαλιστικό σύστημα. Για την αντιμετώπιση αυτού του κινδύνου, ο εθνικός σχεδιασμός στοχεύει στη μείωση του ψηφιακού χάσματος για ευάλωτες ομάδες και άτομα τρίτης ηλικίας. Παράλληλα, εισάγεται η έννοια της «ψηφιακής πολιτεότητας» (digital citizenship) στον εθνικό σχεδιασμό, η οποία επιδιώκει την υπεύθυνη και ασφαλή χρήση της τεχνολογίας, πέρα από την απλή τεχνική κατάρτιση (MDG, 2022).

Το ζήτημα της πρόσβασης και της ποιότητας των δεδομένων (data sovereignty)

Η ΤΝ τρέφεται από δεδομένα. Στην Ελλάδα, η έλλειψη δομημένων και ψηφιοποιημένων ιστορικών δεδομένων σε κρίσιμους τομείς, όπως η Δικαιοσύνη, η Υγεία και η δημόσια διοίκηση, καθιστά την εκπαίδευση εγχώριων μοντέλων εξαιρετικά δύσκολη. Αυτό μας αναγκάζει να εισάγουμε λύσεις «μαύρου κουτιού» από το εξωτερικό, οι οποίες συχνά δεν λαμβάνουν υπ' όψιν την ελληνική γλώσσα, τη νομοθεσία ή τις πολιτισμικές ιδιαιτερότητες. Ελλοχεύει ο κίνδυνος της ψηφιακής αποικιοκρατίας, όπου οι ελληνικές επιχειρήσεις θα εξαρτώνται πλήρως από τις υποδομές και τους αλγορίθμους ελάχιστων παγκόσμιων κολοσσών.

Ηθική και αλγοριθμική μεροληψία στην εργασία

Κεντρικό ρόλο κατέχει η ηθική διάσταση της ΤΝ στην εργασία, όπως επισημαίνεται και στην ανάλυση των Oliver Giering και Stefan Kirchner (2025), θέτοντας ερωτήματα για την υπευθυνότητα και τη διαφάνεια των αλγορίθμων. Οι αλγόριθμοι που χρησιμοποιούνται για την πρόσληψη προσωπικού ή την αξιολόγηση της απόδοσης συχνά κληρονομούν τις προκαταλήψεις των δεδομένων πάνω στα οποία εκπαιδεύτηκαν. Στην Ελλάδα, υπάρχει ο κίνδυνος η ΤΝ να διαιωνίσει κοινωνικές ανισότητες (π.χ., έμφυλες διακρίσεις στον κλάδο της τεχνολογίας ή ηλικιακές διακρίσεις).

Στο προσκήνιο έρχεται το ερώτημα της υπευθυνότητας: Ποιος φέρει την ευθύνη όταν ένας αλγόριθμος λαμβάνει μια λανθασμένη ή άδικη απόφαση για έναν εργαζόμενο; Η σταδιακή ενσωμάτωση της Πράξης για την ΤΝ στις εργασιακές σχέσεις αναμένεται να ενισχύσει το αίσθημα ασφάλειας για εργοδότες και εργαζομένους, ιδιαίτερα σε κρίσιμους τομείς, όπως η πρόσληψη και η αξιολόγηση προσωπικού. Η περαιτέρω νομική εμβάθυνση σε αυτές τις πρακτικές θα συμβάλει στην αποφυγή αβεβαιοτήτων και στην προστασία από πιθανές αλγοριθμικές μεροληψίες. Παράλληλα, η διασφάλιση της ηθικής χρήσης της ΤΝ ενισχύεται μέσα από τον κοινωνικό διάλογο και τη συμμετοχή των εργαζομένων στον σχεδιασμό των αλγοριθμικών συστημάτων, διασφαλίζοντας τη διαφάνεια και τη λογοδοσία στις εργασιακές σχέσεις (MDG, 2025).

8.3.6. Πολιτικές εκπαίδευσης, επιμόρφωσης και μετάβασης

Η αντιμετώπιση των παραπάνω προκλήσεων απαιτεί μια ριζική αναδιάρθρωση των πολιτικών εκπαίδευσης και απασχόλησης. Κομβικής σημασίας είναι η υιοθέτηση μιας ολιστικής προσέγγισης, η οποία δεν περιορίζεται σε αποσπασματικές παρεμβάσεις, αλλά καλύπτει ολόκληρο το φάσμα, από τη σχολική εκπαίδευση και την τεχνική-επαγγελματική κατάρτιση, μέχρι την ανώτατη εκπαίδευση και τη διά βίου επαγγελματική μάθηση, με

έμφαση στην αναβάθμιση (upskilling) και στην επανειδίκευση (reskilling) του ανθρώπινου δυναμικού. Στο πλαίσιο αυτό, παρατηρείται μια στρατηγική μετατόπιση προς πολιτικές που θέτουν την Τεχνητή Νοημοσύνη στον πυρήνα του σχεδιασμού τους, όπως αποτυπώνεται στη Βίβλο Ψηφιακού Μετασχηματισμού 2026-2030 (MDG, 2025). Η εξέλιξη αυτή μετασχηματίζει τις παραδοσιακές μεθόδους στήριξης της εργασίας σε ενεργητικές πολιτικές ψηφιακής ετοιμότητας.

Μεταρρύθμιση της Τριτοβάθμιας Εκπαίδευσης: Η διεπιστημονικότητα ως κανόνας

Ως πανεπιστημιακοί, διαπιστώνουμε ότι η στενή εξειδίκευση είναι πλέον τροχοπέδη. Κομβική θεωρείται η ενσωμάτωση της ΤΝ σε κάθε επιστημονικό πεδίο. Η κατανόηση των LLMs από τους νομικούς και η γνώση των περιθωρίων σφάλματος των αλγορίθμων από τους ιατρούς αποτελούν πλέον βασικές προϋποθέσεις για την ορθή άσκηση του επαγγέλματός τους. Το πανεπιστήμιο του μέλλοντος πρέπει να εκπαιδεύει υβριδικούς επιστήμονες.

Το μοντέλο των micro-credentials και η διά βίου μάθηση

Η παραδοσιακή επιμόρφωση μέσω πολύμηνων σεμιναρίων είναι αργή για τους ρυθμούς της ΤΝ. Η υιοθέτηση μικρο-πιστοποιήσεων (micro-credentials) επιτρέπει την εστίαση σε συγκεκριμένες και άμεσα εφαρμόσιμες δεξιότητες. Μέσω αυτών, οι εργαζόμενοι αποκτούν τη δυνατότητα να επικαιροποιούν το ψηφιακό τους προφίλ ανά τακτά χρονικά διαστήματα, ανταποκρινόμενοι στις γρήγορες αλλαγές που φέρνει η Τεχνητή Νοημοσύνη (OECD, 2025a). Η αποτελεσματικότητα των micro-credentials δεν έγκειται στην εκμάθηση συγκεκριμένων εργαλείων, τα οποία απαξιώνονται ταχύτατα, αλλά στην καλλιέργεια διαχρονικών ικανοτήτων συνεργασίας ανθρώπου-αλγορίθμου, εποπτείας αυτοματοποιημένων αποφάσεων και κριτικής αξιολόγησης των αποτελεσμάτων της ΤΝ. Στόχος είναι ο εργαζόμενος, μέσω πλαισίου και δράσεων ενίσχυσης της επανειδίκευσης, να αναβαθμίζει τις δεξιότητές του, διαμορφώνοντας το ψηφιακό του διαβατήριο σύμφωνα με τις σύγχρονες τάσεις. Σε αυτό το πλαίσιο, οι δημόσιες υπηρεσίες απασχόλησης εξελίσσονται σε ενεργούς διαμεσολαβητές, χρησιμοποιώντας προγνωστικά εργαλεία για να προσφέρουν εξατομικευμένη υποστήριξη και να προλαμβάνουν τις επιπτώσεις της αυτοματοποίησης πριν αυτές εκδηλωθούν στην αγορά (MDG, 2025).

8.3.7. Συμπεράσματα

Η ΤΝ δεν αποτελεί μια απλή τεχνολογική προσθήκη, αλλά έναν κοινωνικο-οικονομικό επιταχυντή που θα επανακαθορίσει τη θέση της Ελλάδας στον παγκόσμιο καταμερισμό εργασίας. Η μετάβαση από την οικονομία της πληροφορίας στην οικονομία της νοημοσύνης απαιτεί μια ριζική αλλαγή

νοοτροπίας: Από την προστασία των θέσεων εργασίας πρέπει να περάσουμε στην προστασία των εργαζομένων και την ενίσχυση της ικανότητάς τους για διαρκή προσαρμογή. Η διάκριση αυτή μετατοπίζει το επίκεντρο της πολιτικής από τις στατικές επαγγελματικές κατηγορίες στις δυναμικές δεξιότητες, καθιστώντας τη συνεχή επανεκπαίδευση και τον αλγοριθμικό γραμματισμό βασικούς πυλώνες κοινωνικής συνοχής.

Η πρακτική υλοποίηση αυτής της μετάβασης μπορεί να βασιστεί σε εξειδικευμένα προγράμματα επιδότησης της «επαύξεσης» (augmentation grants) για τις ΜμΕ, τα οποία θα ενισχύουν την αγορά εργαλείων ΤΝ που συμπληρώνουν την ανθρώπινη εργασία αντί να την υποκαθιστούν. Παράλληλα, η καθιέρωση ατομικών μερίδων επανειδίκευσης (reskilling accounts), συνδεδεμένων με το εθνικό πλαίσιο μικρο-πιστοποιήσεων, μπορεί να επιτρέπει στους εργαζομένους να επιλέγουν εξατομικευμένα μονοπάτια μάθησης που ανταποκρίνονται στις ταχύτατες αλλαγές των δεξιοτήτων τους. Για την αντιμετώπιση του δημογραφικού-ψηφιακού χάσματος, η εισαγωγή μοντέλων «μεταφοράς ψηφιακής εμπειρίας» (intergenerational mentoring) εντός των εταιρειών μπορεί να διασφαλίσει ότι η εμπειρία των μεγαλύτερων στελεχών δεν θα απαξιωθεί, αλλά θα ενισχυθεί από την αλγοριθμική παιδεία των νεότερων.

8.3.8. Τελικές σκέψεις

Η επιτυχία της ΤΝ θα κριθεί από την κοινωνική της νομιμοποίηση. Αν οι εργαζόμενοι αντιληφθούν την ΤΝ ως ένα εργαλείο που τους απαλλάσσει από την επαναλαμβανόμενη εργασία, η υιοθέτηση θα είναι οργανική. Η διασφάλιση της εργασιακής ειρήνης συνδέεται άρρηκτα με τη δεοντολογική διακυβέρνηση και την υιοθέτηση πρωτοκόλλων αλγοριθμικής διαφάνειας, όπου ο συμμετοχικός σχεδιασμός των συστημάτων εγγυάται τη λογοδοσία και την προστασία από μεροληψίες στις εργασιακές σχέσεις.

Κλείνοντας, χρειάζεται να τονιστεί ότι η ΤΝ είναι, στον πυρήνα της, μια ανθρώπινη δημιουργία. Οι αλγόριθμοι αντανakλούν τις δικές μας αξίες, τις δικές μας προτεραιότητες και τις δικές μας αδυναμίες. Η Ελλάδα έχει την ιστορική ευκαιρία να ενσωματώσει την τεχνολογία αυτήν με έναν τρόπο που να σέβεται την ανθρωποκεντρική της παράδοση. Προς αυτήν την κατεύθυνση, μια εθνική στρατηγική δεν φοβάται το μέλλον, αλλά το συνδιαμορφώνει. Η γνώση, η ηθική και η συνεργασία είναι τα μόνα εργαλεία που μπορούν να εγγυηθούν ότι η εποχή της ΤΝ θα είναι μια εποχή ευημερίας για όλους τους Έλληνες πολίτες. Σε μια εποχή ραγδαίων τεχνολογικών μεταβολών, η επιλογή δεν είναι αν η Ελλάδα υιοθετεί την ΤΝ, αλλά με ποιους όρους, για ποιους και με ποιο κοινωνικό αποτύπωμα.

Βιβλιογραφία

- Acemoglu, D., Autor, D., Hazell, J., & Restrepo, P. (2022). Artificial intelligence and jobs: Evidence from online vacancies. *Journal of Labor Economics*, 40(S1), 293-340. <https://doi.org/10.1086/718327>
- Arntz, M., Gregory, T., & Zierahn, U. (2016). The risk of automation for jobs: OECD Social. *Employment and Migration Working Papers*, 189. OECD Publishing. <https://doi.org/10.1787/5jlz9h56dvq7-en>
- Autor, D. H. (2015). Why are there still so many jobs? *Journal of Economic Perspectives*, 29(3), 3-30. <https://doi.org/10.1257/jep.29.3.3>
- Pouliakas, K. (2021). *Artificial intelligence and job automation: an EU analysis using online job vacancy data*. Publications Office. <http://data.europa.eu/doi/10.2801/305373>
- Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023). GPTs are GPTs: An early look at the labor market impact potential of large language models. *Science*, 384(6702), 1306-1308. <https://www.science.org/doi/10.1126/science.adj0998>
- Giering, O., & Kirchner, S. (2025). Artificial intelligence and autonomy at work: Empirical insights from Germany. *Journal for Labour Market Research*, 59(20). <https://doi.org/10.1186/s12651-025-00401-5>
- Lightcast (2025). *The Speed of Skill Change: Every Job in the US is Changing. Are You?* Future-Ready Workforce Collection. <https://lightcast.io/resources/research/speed-of-skill-change>
- Mäkelä, E., & Stephany, F. (2024). Complement or substitute? How AI increases the demand for human skills (preprint). *arXiv*.
- Ministry of Digital Governance (MDG) (2025). *Digital Transformation Bible 2026–2030*. Hellenic Republic. <https://digitalstrategy.gov.gr/en/digital-transformation-bible-2026-2030>

Ministry of Digital Governance (MDG) (2022). *Revised digital skills strategic plan: Executive summary—Acceleration of the Greek digital transformation*. Hellenic Republic.

Organisation for Economic Co-operation and Development (OECD) (2024). *OECD digital economy outlook 2024 (Vol. 2): Strengthening connectivity, innovation and trust*. https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/11/oecd-digital-economy-outlook-2024-volume-2_9b2801fc/3adf705b-en.pdf

Organisation for Economic Co-operation and Development (OECD) (2025a). *Bridging the AI skills gap: Is training keeping up?* https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/04/bridging-the-ai-skills-gap_b43c7c4a/66d0702e-en.pdf

Organisation for Economic Co-operation and Development (OECD) (2025b). *OECD employment outlook 2025: Can we get through the demographic crunch?* <https://doi.org/10.1787/194a947b-en>

Organisation for Economic Co-operation and Development (OECD) (2026c). *OECD skills outlook 2025: Building the skills of the 21st century for all*. <https://doi.org/10.1787/26163cd3-en>

8.4. Οι επιπτώσεις της χρήσης Τεχνητής Νοημοσύνης στα δικαιώματα των εργαζομένων

Στον χώρο εργασίας, τα συστήματα Τεχνητής Νοημοσύνης χρησιμοποιούνται για την αυτοματοποίηση ή την υποστήριξη αποφάσεων ως προς το ανθρώπινο δυναμικό σχετικά με την πρόσληψη και την αξιολόγηση υποψηφίων, για την αυτοματοποίηση καθηκόντων που παραδοσιακά εκτελούνταν από εργαζομένους, καθώς και για την υποστήριξη διοικητικών λειτουργιών μέσω αναλυτικών εργαλείων και αλγορίθμων που βασίζονται στην ΤΝ.

Η χρήση συστημάτων ΤΝ εντείνει τους περιορισμούς στο δικαίωμα της ιδιωτικής ζωής και των προσωπικών δεδομένων των εργαζομένων, ένα δικαίωμα που, πέρα από την αυτοτελή αξία και κατοχύρωσή του, συνδέεται άρρηκτα με το δικαίωμα στην εργασία, που περιλαμβάνει την προστασία από αδικαιολόγητες επεμβάσεις στην ιδιωτικότητα.²⁰⁴ Η αυξημένη χρήση της ΤΝ στον χώρο εργασίας συνήθως προϋποθέτει ή συνεπάγεται μεγαλύτερη συλλογή και ανάλυση δεδομένων σχετικά με τους εργαζομένους και τους υποψηφίους για εργασία, για σκοπούς συναγωγής συστάσεων ή λήψης αποφάσεων ή και για τον σκοπό την εκπαίδευσης αυτών των συστημάτων. Τα δεδομένα αυτά μπορεί να περιλαμβάνουν πληροφορίες, όπως οι κινήσεις των εργαζομένων, βιομετρικά δεδομένα που μπορεί να οδηγούν και σε αναγνώριση συναισθήματος, δεδομένα υγείας (καρδιακοί παλμοί, αρτηριακή πίεση), καθώς και ψηφιακές δραστηριότητες. Τα θεμελιώδη δικαιώματα στην προστασία προσωπικών δεδομένων ενδέχεται να υπονομευτούν και από τη χρήση συστημάτων ΤΝ για την παρακολούθηση των επιδόσεων και της συμπεριφοράς των εν λόγω προσώπων.

Ένα μείζον ζήτημα προκύπτει αναφορικά με την προστασία από διακρίσεις και την αρχή της ισότητας. Όπως αναφέρεται στην αιτιολογική σκέψη 57 της AI Act, καθ' όλη τη διάρκεια της διαδικασίας πρόσληψης και κατά την αξιολόγηση, την προαγωγή ή τη διατήρηση προσώπων σε συμβατικές σχέσεις που σχετίζονται με την εργασία, τα συστήματα αυτά μπορούν να διαιωνίσουν ιστορικές νοοτροπίες διακρίσεων, για παράδειγμα εις βάρος γυναικών, ορισμένων ηλικιακών ομάδων, ατόμων με αναπηρία ή ατόμων συγκεκριμένης φυλετικής ή εθνοτικής καταγωγής ή γενετήσιου προσανατολισμού.

²⁰⁴ European Committee of Social Rights (2012). *Conclusions 2012: Statement of interpretation on Article 1§2*. Council of Europe. <https://www.coe.int/en/web/european-social-charter/statements-of-interpretation1>

Τέτοια προβλήματα μπορεί να προκύπτουν, π.χ., στο στάδιο της επιλογής/ πρόσληψης, τα οποία μπορεί να επηρεαστούν αρνητικά, για παράδειγμα σε περιπτώσεις όπου η χρήση συστημάτων μηχανικής μάθησης για την επιλογή υποψηφίων οδήγησε σε διακριτικά αποτελέσματα ή όπου συστήματα αναγνώρισης προσώπου και ανάλυσης συναισθήματος που βασίζονται στην ΤΝ είχαν ως αποτέλεσμα φυλετικές διακρίσεις.²⁰⁵

Μείζονος σημασίας είναι η διασφάλιση της αρχής της διαφάνειας και της λογοδοσίας. Το πρόβλημα ανακύπτει, ιδίως, στο πλαίσιο των προσλήψεων, του καθορισμού των αμοιβών, της επιτήρησης στον χώρο εργασίας και των διαδικασιών λήψης αποφάσεων. Για παράδειγμα, λόγω της εγγενούς αδιαφάνειας των συστημάτων ΤΝ (black box), ο καθορισμός των αμοιβών και η κατανομή εργασιών ιδίως στην εργασία μέσω πλατφορμών και στην gig work μπορεί να αφήνουν τους εργαζομένους χωρίς εξηγήσεις για τις διακυμάνσεις στις αποδοχές τους ή για τη διαθεσιμότητα θέσεων εργασίας.

²⁰⁵ Steering Committee for Human Rights (CDDH) (2026). *Human Rights and Artificial Intelligence: Handbook of the Steering Committee for Human Rights (CDDH)*. Council of Europe. <https://rm.coe.int/artificial-intelligence-handbook/48802b5bf2>

8.5. Η αλγοριθμική διοίκηση (algorithmic management) και οι εργασιακές σχέσεις

Οι τεχνολογίες Τεχνητής Νοημοσύνης μεταλλάσσουν τον χώρο εργασίας. Πρόσφατη κοινή έρευνα του Κοινού Κέντρου Ερευνών (Joint Research Centre – JRC) της Ευρωπαϊκής Επιτροπής και της Διεθνούς Οργάνωσης Εργασίας (International Labour Organization – ILO)²⁰⁶ δείχνει ότι ένα ευρύ φάσμα ψηφιακών συστημάτων, συμπεριλαμβανομένων και γενικής χρήσης, μπορεί να χρησιμοποιηθεί για τη διαχείριση και τον συντονισμό της εργασίας και τη συλλογή δεδομένων εργαζομένων, είτε αυτό γίνεται σκόπιμα είτε όχι. Σε συνδυασμό με τη δεδομενοποίηση, όπου κάθε ενέργεια, κίνηση ή κλικ παράγει δεδομένα, η ΤΝ διαμορφώνει μια νέα εργασιακή πραγματικότητα που χαρακτηρίζεται όλο και περισσότερο από την αυτοματοποίηση και την ψηφιοποίηση της διαδικασίας λήψης αποφάσεων, αλλά και τη λεγόμενη «αλγοριθμική διοίκηση». Η αλγοριθμική διοίκηση αναφέρεται κυρίως στη χρήση συστημάτων ΤΝ για τη διαχείριση και τον συντονισμό του εργατικού δυναμικού, εν ολίγοις στην εκχώρηση διοικητικών λειτουργιών (καθοδήγηση, αξιολόγηση, πειθαρχία) στον «αλγόριθμο».²⁰⁷

Πρόκειται για πρακτικές και λειτουργίες, όπως η αξιολόγηση των προσόντων ενός υποψηφίου ή η μέτρηση της απόδοσης, που υλοποιούνται εν μέρει ή εξ ολοκλήρου από «μηχανές» και έχουν ως προϋπόθεση αλλά και ως αποτέλεσμα τη συλλογή δεδομένων των εργαζομένων. Η αλγοριθμική διαχείριση της εργασίας προϋποθέτει έναν –εντεινόμενο– βαθμό ψηφιακής παρακολούθησης, ενώ, αντίστροφα, η ψηφιακή παρακολούθηση ενθαρρύνει τη χρήση αλγορίθμων για διοικητικούς σκοπούς.

²⁰⁶ Rani, U., Pesole, A., & Gonzalez Vazquez, I., (2024). *Algorithmic Management practices in regular workplaces: case studies in logistics and healthcare*. Publications Office of the European Union. <https://doi.org/10.2760/712475>

²⁰⁷ Αναφορικά με την αλγοριθμική διοίκηση, ιδιαίτερο ενδιαφέρον παρουσιάζουν τα στοιχεία και οι επισημάνσεις που περιέχονται στη μελέτη του Organisation for Economic Co-operation and Development (OECD). (2025). *Algorithmic management in the workplace: New evidence from an OECD employer survey* (OECD Artificial Intelligence Papers, No. 31). <https://doi.org/10.1787/287c13c4-en>

Θα μπορούσαμε να αναφερθούμε σε μια «ψηφιακή μορφή του τείλορισμού»²⁰⁸ που επηρεάζει τις εργασιακές και κατ' επέκταση τις κοινωνικές σχέσεις. Μελέτες δείχνουν ότι οι πρακτικές διαχείρισης που βασίζονται σε δεδομένα μπορούν να αυξήσουν την παραγωγικότητα και να ενισχύσουν την οργάνωση της εργασίας μέσω της βελτίωσης των ροών πληροφορίας και της αποδοτικότητας του συντονισμού. Τα επιχειρήματα υπέρ της χρήσης αλγορίθμων στις αποφάσεις που αφορούν τις εργασιακές σχέσεις, όπως η πρόσληψη, η αξιολόγηση, η αμοιβή, η απόλυση βασίζονται στην ιδέα ότι οι αλγόριθμοι είναι πιο αντικειμενικοί από τους ανθρώπους. Από την άλλη πλευρά, τέτοιες μορφές «αλγοριθμικής διοίκησης» μπορεί να περιορίσουν αδικαιολόγητα και υπέρμετρα την αυτονομία των εργαζομένων, να μειώσουν την ανθρώπινη επαφή και τη δυνατότητα των εργαζομένων να συζητούν την εργασία τους με τους εργοδότες ή προϊσταμένους τους ή να αμφισβητούν άμεσα και αδιαμεσολάβητα αποφάσεις που φαίνονται λανθασμένες, άδικες ή διακριτικές. Ως περαιτέρω επίπτωση θα μπορούσε να προκύψει η υπονόμευση της αίσθησης της εργασιακής ταυτότητας και της νοηματοδότησης της εργασίας με πιθανές συνέπειες κινδύνους για τη σωματική και ψυχική υγεία, αλλά και ασφάλεια στην εργασία.²⁰⁹

Σε κάθε περίπτωση, η ΤΝ δεν φαίνεται να συνιστά το μαγικό ραβδί που διά της τεχνολογικής επέμβασης απαλείφει κάθε δυσλειτουργία και επιφύλαξη για την ποιότητα και την αποτελεσματικότητα τέτοιων διαδικασιών. Οι αποφάσεις που παράγονται από συστήματα ΤΝ συνοδεύονται από το επίχρυσμα ουδετερότητας. Η αποφυγή της ανθρώπινης εμπλοκής εξαλείφει πράγματι ψυχο-κοινωνικές προκαταλήψεις; Η ΤΝ μπορεί να συνοδεύεται από την (ψευδ)αίσθηση της «αντικειμενικότητας» αλλά δεν αποκλείεται να χαρακτηρίζεται από τις προκαταλήψεις του σχεδιαστή της ή αυτού που επέλεξε τα σύνολα δεδομένων που χρησιμοποιούνται. Να ενσωματώνει στη μηχανή τις προκαταλήψεις, με συνέπεια να διαιωνίζει τις –έστω ακούσιες– άμεσες ή έμμεσες μεροληπτικές αντιλήψεις και διακρίσεις κατά ανθρώπων ή ομάδων.

Όταν τα καθήκοντα παρακολούθησης και διαχείρισης εκχωρούνται στους αλγορίθμους, δημιουργείται εξάρτηση από την αξιοπιστία των δεδομένων και των συστημάτων, μειώνοντας ενίοτε την ικανότητα να λαμβάνονται δίκαιες και ακριβείς αποφάσεις. Αποφάσεις που παράγονται υπό καθεστώς αδιαφάνειας, καθώς συχνότατα δεν είναι εφικτές η επεξήγηση και

²⁰⁸ Ως τείλορισμός (taylorism) ορίζεται το σύστημα οργάνωσης και διοίκησης της εργασίας που θεμελιώθηκε από τον Frederick W. Taylor με σκοπό την αύξηση της παραγωγικότητας μέσω της επιστημονικής ανάλυσης των εργασιακών διαδικασιών, της διάσπασης της εργασίας σε επιμέρους, μετρήσιμα καθήκοντα, της τυποποίησης της «βέλτιστης μεθόδου» εκτέλεσης, καθώς και του σαφούς διαχωρισμού μεταξύ σχεδιασμού/ελέγχου από τη διοίκηση και εκτέλεσης από τον εργαζόμενο. Βλ. Taylor, F. W. (1911), *The Principles of Scientific Management*. Harper and Brother Publishers.

²⁰⁹ Organisation for Economic Co-operation and Development (OECD). (2024). *Using AI in the workplace: Opportunities, risks and policy responses* (OECD Artificial Intelligence Papers, No. 11). <https://doi.org/10.1787/73d417f9-en>

η κατανόηση των λόγων για τους οποίους ένα μοντέλο έχει οδηγήσει σε ένα συγκεκριμένο αποτέλεσμα και ποιος συνδυασμός εισροών συνέβαλε σε αυτό.

Ένα σύστημα μπορεί να παρακολουθεί μόνο δεδομένα που μπορούν να κωδικοποιηθούν και, επομένως, η δραστηριότητα των εργαζομένων μετατρέπεται σε κώδικα, οδηγώντας σε υπεραπλούστευση και απόκρυψη «αόρατων» δραστηριοτήτων. Ο «Boss ex machina»²¹⁰ μπορεί, όμως, να αποδειχθεί άκαμπος, ανήμπορος να δείξει ενσυναίσθηση, κατανόηση και συμπόνια. Καθώς οι πρακτικές αυτές διαχέονται και σε πιο παραδοσιακούς τομείς εργασίας, τα ερωτήματα τίθενται επιτακτικά: Είναι ελέγξιμες, από ποιον και με ποια μέθοδο αυτές οι μορφές αλγοριθμικής διοίκησης; Ποιος καθορίζει αυτήν την (αναπόδραστη;) εξέλιξη; Ποια τα περιθώρια επιρροής των συλλογικών φορέων, της Πολιτείας, του νομοθέτη; Πού μετατίθεται το πεδίο της διαπραγμάτευσης;

²¹⁰ Aloisi, A., & De Stefano, V. (2025). *Εργοδότης μου είναι ο Αλγόριθμος*. Πανεπιστημιακές Εκδόσεις Κρήτης.

8.6. Το κανονιστικό πλαίσιο για τη χρήση Τεχνητής Νοημοσύνης στο πλαίσιο των εργασιακών σχέσεων

Ο Έλληνας νομοθέτης με τον Ν. 4961/2022 επιχείρησε να επιβάλει στους εργοδότες την υποχρέωση ενημέρωσης για τη χρήση Τεχνητής Νοημοσύνης στον εργασιακό τομέα, εφόσον επηρεάζει οποιαδήποτε διαδικασία λήψης αποφάσεων σχετικά με τους εργαζομένους και έχει αντίκτυπο στις συνθήκες εργασίας, στην επιλογή, στην πρόσληψη ή στην αξιολόγησή τους, καθώς και στη διασφάλιση της τήρησης της αρχής της ισότητας. Η Επιθεώρηση Εργασίας έχει ήδη εκδώσει την πρώτη σχετική απόφαση που αφορά χρήση ΤΝ στο πλαίσιο των εργασιακών σχέσεων.

Η Πράξη για την ΤΝ απαγορεύει τη διάγνωση (και την καταγραφή!) των συναισθημάτων των εργαζομένων στον χώρο εργασίας, ενώ κατατάσσει στην κατηγορία του «υψηλού κινδύνου» τα συστήματα που προορίζονται να χρησιμοποιηθούν στο πλαίσιο της απασχόλησης και διαχείρισης εργαζομένων. Πρόκειται ειδικότερα για τα συστήματα ΤΝ που προορίζονται να χρησιμοποιηθούν για την πρόσληψη ή την επιλογή φυσικών προσώπων, ιδίως για τη δημοσίευση στοχευμένων αγγελιών θέσεων εργασίας, την ανάλυση και το φιλτράρισμα των αιτήσεων εργασίας και την αξιολόγηση των υποψηφίων, ή τα συστήματα ΤΝ που προορίζονται να χρησιμοποιηθούν για τη λήψη αποφάσεων που επηρεάζουν τους όρους των εργασιακών σχέσεων, για την προώθηση ή την καταγγελία συμβατικών σχέσεων που σχετίζονται με την εργασία, για την κατανομή καθηκόντων με βάση τη συμπεριφορά του ατόμου ή τα προσωπικά γνωρίσματα ή χαρακτηριστικά, ή για την παρακολούθηση και την αξιολόγηση των επιδόσεων και της συμπεριφοράς των προσώπων που τελούν σε τέτοιες σχέσεις.

Η συμμόρφωση με τις απαιτήσεις ως προς την ανθρώπινη εποπτεία και τον έλεγχο των «στοιχείων εξόδου», η ενημέρωση των προσώπων που θίγονται όσο και των εκπροσώπων τους αναφορικά με τη χρήση συστημάτων ΤΝ στον χώρο εργασίας, η παρακολούθηση του αντικτύπου της εισαγωγής συστημάτων ΤΝ στον χώρο εργασίας στην ισότητα των φύλων και την ποικιλομορφία στο εργατικό δυναμικό, η κατάρτιση και ευαισθητοποίηση του εργατικού δυναμικού σχετικά με τις μεροληψίες των δεδομένων, τα στερεότυπα και τους κινδύνους διακρίσεων κατά τη χρήση συστημάτων ΤΝ προβλέπονται από την έννομη τάξη και εφαρμοζόμενα μπορούν να μετριάσουν τους κινδύνους για

τα δικαιώματα των εργαζομένων.²¹¹ Όπως σημειώνει η Κλειώ Σγουροπούλου, κρίσιμη παράμετρος είναι η δεοντολογική διακυβέρνηση του χώρου εργασίας κατά τρόπο ώστε να συμπεριλαμβάνονται και οι εργαζόμενοι.

Οι προκλήσεις που σχετίζονται με την ΤΝ είναι ιδιαίτερα έντονες σε νέες μορφές απασχόλησης,²¹² όπως η εργασία μέσω ψηφιακών πλατφορμών ή η λεγόμενη «gig εργασία».²¹³ Αξίζει να σημειώσουμε ότι η Οδηγία 2024/2831 –μεταξύ των άλλων– αποσκοπεί στη διασφάλιση της διαφάνειας στην αλγοριθμική διακυβέρνηση της εργασίας μέσω ψηφιακών πλατφορμών, η οποία αναδεικνύεται σε θεμελιώδη εγγύηση προστασίας των απασχολούμενων. Η Οδηγία επιβάλλει στις ψηφιακές πλατφόρμες την υποχρέωση να παρέχουν στους απασχολούμενους σαφή, επαρκή και προσβάσιμη ενημέρωση σχετικά με κρίσιμες πτυχές της συνεργασίας τους με την πλατφόρμα, η οποία συμπεριλαμβάνει τη χρήση τεχνολογικών μέσων ή συστημάτων που αξιοποιούνται για την αξιολόγηση των απασχολούμενων. Η κανονιστική κατοχύρωση της ενημέρωσης των εργαζομένων ή/και των εκπροσώπων τους επιδιώκει να αποκαταστήσει, έστω εν μέρει, την εμφανή ασυμμετρία πληροφόρησης και ισχύος που χαρακτηρίζει τη σχέση μεταξύ πλατφόρμας και απασχολούμενων. Ιδιαίτερη σημασία αποδίδεται, περαιτέρω, στον έλεγχο των αλγοριθμικών συστημάτων που χρησιμοποιούνται από τις πλατφόρμες για την κατανομή της εργασίας, τον καθορισμό των αμοιβών και την αξιολόγηση των απασχολούμενων.

Ο ενωσιακός νομοθέτης αναγνωρίζει ότι οι αλγόριθμοι δεν συνιστούν ουδέτερα τεχνολογικά εργαλεία, αλλά μηχανισμούς που μπορούν να επηρεάσουν καθοριστικά τη θέση των απασχολούμενων και, ως εκ τούτου, πρέπει να υπόκεινται σε αυστηρές εγγυήσεις ελέγχου και λογοδοσίας. Η Οδηγία απαιτεί να μη χρησιμοποιούνται τα συστήματα αυτά κατά τρόπο ώστε να συνεπάγονται άμεσες ή έμμεσες διακρίσεις οποιασδήποτε μορφής, όπως, ιδίως, λόγω φύλου, ηλικίας ή εθνικής καταγωγής. Στο πλαίσιο αυτό οι απασχολούμενοι πρέπει να έχουν τη δυνατότητα να γνωρίζουν τα βασικά κριτήρια και τις παραμέτρους βάσει των οποίων αξιολογούνται. Η απαίτηση αυτή δεν αποσκοπεί μόνο στην παροχή πληροφόρησης, αλλά και στη διαμόρφωση ενός πλαισίου ελέγχιμης και κατανοητής λήψης αποφάσεων, το οποίο καθιστά δυνατή την αμφισβήτηση τυχόν αυθαίρετων ή δυσμενών πρακτικών. Υπό την έννοια αυτήν, η διαφάνεια συνιστά αναγκαία προϋπόθεση για τη λογοδοσία των πλατφορμών και για τη διασφάλιση ελάχιστου επιπέδου δικαιοκρατικής οργάνωσης των

²¹¹ Steering Committee for Human Rights (CDDH) (2026). *Human Rights and Artificial Intelligence: Handbook of the Steering Committee for Human Rights (CDDH)* (p. 217). Council of Europe. <https://rm.coe.int/artificial-intelligence-handbook/48802b5bf2>

²¹² Ό.π.

²¹³ Η λεγόμενη gig work είναι βραχυχρόνια, βασισμένη σε επιμέρους καθήκοντα ή κατά παραγγελία εργασία, την οποία οι άνθρωποι εκτελούν για μεμονωμένες δουλειές ή «αναθέσεις» (gigs), αντί στο πλαίσιο μιας παραδοσιακής μακροχρόνιας σχέσης απασχόλησης. Σήμερα, ο όρος χρησιμοποιείται συχνά με ευρεία έννοια, ώστε να περιλαμβάνει και την εργασία που διαμεσολαβείται από πλατφόρμες, όπου εφαρμογές ή ιστότοποι συνδέουν εργαζομένους με πελάτες για υπηρεσίες, όπως μεταφορά επιβατών, διανομή, προγραμματισμός, σχεδιασμός ή διαδικτυακή ελεύθερη εργασία. Βλ. ειδικότερα Lane, M. (2020). *Regulating platform work in the digital age* (OECD Going Digital Toolkit Notes, No. 1). OECD Publishing. <https://doi.org/10.1787/181f8a7f-en>

όρων παροχής εργασίας στο ψηφιακό περιβάλλον. Επίσης, βασική κανονιστική επιλογή συνιστά η κατοχύρωση του δικαιώματος των απασχολουμένων να διατυπώνουν αντιρρήσεις και να αμφισβητούν αποφάσεις που λαμβάνονται μέσω αλγοριθμικών διαδικασιών, όταν εκτιμούν ότι οι αποφάσεις αυτές τους επηρεάζουν δυσμενώς. Η Οδηγία επιχειρεί να θέσει όρια στην τεχνολογικά διαμεσολαβημένη άσκηση εργοδοτικής εξουσίας, προωθώντας ένα πρότυπο αλγοριθμικής διακυβέρνησης σε εναρμόνιση με τις αρχές της ισότητας, της διαφάνειας και της αποτελεσματικής προστασίας των εργασιακών δικαιωμάτων.²¹⁴

Το Ευρωπαϊκό Κοινοβούλιο διατύπωσε με ευρεία πλειοψηφία τον Δεκέμβριο του 2025 πρόταση προς την Επιτροπή ώστε να διασφαλιστεί εν γένει η διαφανής, δίκαιη και ασφαλής χρήση αυτοματοποιημένων συστημάτων παρακολούθησης και λήψης αποφάσεων στον χώρο εργασίας, με έμφαση στη διασφάλιση της ανθρώπινης εποπτείας σε όλες τις αποφάσεις που λαμβάνονται ή υποστηρίζονται από τέτοια συστήματα. Οι ευρωβουλευτές θεωρούν κρίσιμη την πραγματοποίηση διαβούλευσης και την ενημέρωση των εργαζομένων για το πώς αυτά τα συστήματα επηρεάζουν τις συνθήκες εργασίας, τότε χρησιμοποιούνται για τη λήψη αυτοματοποιημένων αποφάσεων, τι είδους δεδομένα συλλέγουν ή επεξεργάζονται και πώς διασφαλίζεται η ανθρώπινη εποπτεία. Επίσης, ζητείται η κατοχύρωση του δικαιώματος να ζητούνται εξηγήσεις και επανεξέταση για αποφάσεις που ελήφθησαν ή υποστηρίχθηκαν από την αλγοριθμική διαχείριση. Η πρόταση αυτή πρέπει να αξιολογηθεί θετικά και να υποστηριχθεί λαμβάνοντας υπ' όψιν ότι «η Ευρώπη μπορεί να συνδυάσει την ανταγωνιστικότητα με την κοινωνική ευθύνη. Μπορεί να στηρίξει καινοτόμες επιχειρήσεις χωρίς να θυσιάζει τα υψηλά πρότυπα και την προστασία των εργαζομένων».²¹⁵

Όπως επισημαίνει στην έκθεσή του το Ευρωπαϊκό Κοινοβούλιο,²¹⁶ η αξιοποίηση των οφελών με ταυτόχρονη ελαχιστοποίηση των μειονεκτημάτων απαιτεί μια ανθρωποκεντρική προσέγγιση, η οποία να ενσωματώνει ισχυρά μέτρα προστασίας δεδομένων, διαβούλευση με τους εργαζομένους και συλλογικές διαπραγματεύσεις, ώστε να διασφαλίζεται η ηθική εφαρμογή αυτών των τεχνολογιών στους ευρωπαϊκούς χώρους εργασίας. Πρόκειται για ζητήματα που δεν αφορούν μόνο την αγορά εργασίας και την προστασία των δικαιωμάτων των εργαζομένων, αλλά εν τέλει την κοινωνική συνοχή και τη δυνατότητα όλων να έχουν μέρος στην ευημερία που υπόσχεται η αξιοποίηση των νέων τεχνολογιών.

²¹⁴ Για την ανάλυση της Οδηγίας βλ. Σκανδάλης, Ι. Β. (2025). Κριτική ανάλυση της Οδηγίας 2024/2831 για τη βελτίωση των συνθηκών εργασίας των απασχολούμενων σε πλατφόρμες και οι εκτιμώμενες επιπτώσεις της στο ελληνικό δίκαιο. *Δίκαιο Επιχειρήσεων & Εταιριών (ΔΕΕ)*, 4, 450-471.

²¹⁵ Βλ. τη δήλωση του Ευρωβουλευτή Andrzej Buja (ΕΛΚ, Πολωνία): Ευρωπαϊκό Κοινοβούλιο (2025, 17 Δεκεμβρίου). Το Κοινοβούλιο ζητά δικλίδες ασφαλείας για την αλγοριθμική διαχείριση στην εργασία [Δελτίο Τύπου]. https://www.europarl.europa.eu/pdfs/news/expert/2025/12/press_release/20251211IPR32175/20251211IPR32175_el.pdf

²¹⁶ De Micheli, B., Dente, G., Faioli, M., Smilari, A., Omersa, E., Patras, S., Ravaglia, M., Vancanwenbergh, S., Barslund, M., & Tobback, I. (2025). *Digitalisation, artificial intelligence and algorithmic management in the workplace: Shaping the future of work* (Study at the request of the European Added Value Unit of the Directorate for Impact Assessment and Foresight, DG EPRS, European Parliament). European Parliamentary Research Service. <https://doi.org/10.2861/0136788>

[ΠΩΣ] ΜΠΟΡΕΙ Η ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΝΑ ΕΙΝΑΙ
ΗΘΙΚΗ ΚΑΙ ΣΥΝΝΟΜΗ;
Ιούλιος 2026

Τεχνητή Νοημοσύνη και έρευνα



Μιχάλης Κρητικός, Λίλιαν Μήτρου

Συγγραφική συνεισφορά ανά υποκεφάλαιο:

9.1., 9.2. – Λίλιαν Μήτρου

9.3. – Μιχάλης Κρητικός

9.1. Η Τεχνητή Νοημοσύνη και η έρευνα

Οι πολυδιάστατες εφαρμογές και δυνατότητες της Τεχνητής Νοημοσύνης έχουν ανοίξει νέους ορίζοντες, αλλά και έχουν γεννήσει δικαιολογημένα νέες προσδοκίες αναφορικά με την έρευνα. Η ΤΝ θεωρείται μια ριζική, μετασχηματιστική τεχνολογία, που διανοίγει νέες δυνατότητες ως προς τον τρόπο που μπορεί να υποστηριχθούν η ερευνητική διαδικασία, η ενίσχυση της παραγωγικότητας των ερευνητών, η διάδοση και αξιοποίηση των αποτελεσμάτων της.

Η έρευνα συνδέεται κατά τούτο στενά με την προώθηση της καινοτομίας και της ανταγωνιστικότητας, και συνακόλουθα με την επιδίωξη της ψηφιακής κυριαρχίας στο πεδίο της ΤΝ.²¹⁷ Για τον λόγο αυτόν, άλλωστε, τα συστήματα και τα μοντέλα ΤΝ που αναπτύσσονται ειδικά και τίθενται σε λειτουργία αποκλειστικά για σκοπούς επιστημονικής έρευνας και ανάπτυξης εξαιρούνται από τις υποχρεώσεις που επιβάλλει η AI Act.²¹⁸

Δεν θα ήταν υπερβολή να σημειώσουμε ότι η ΤΝ μετασχηματίζει τον τρόπο που ασκείται η επιστήμη. Οι δυνατότητες ανάλυσης, σύνθεσης και παραγωγής γνώσης που φαίνεται να προσφέρει η ΤΝ θεωρούνταν έως πρόσφατα αδιανόητες. Αυτές οι δυνατότητες επιτρέπουν στους ερευνητές να εντοπίσουν πρότυπα (patterns) και σχέσεις που ως τώρα παρέμεναν «κρυμμένα», να προάγουν επιστημονικές ανακαλύψεις²¹⁹ και να διευρύνουν τη διεπιστημονική συνεργασία. Εκτός από την εκπαίδευση των μοντέλων (με την επιφύλαξη της ύπαρξης ποιοτικών δεδομένων), η ΤΝ επιτρέπει τη διαχείριση του τεράστιου πλήθους των πληροφοριών, συμπεριλαμβανομένου του γεωμετρικά (αν όχι εκθετικά) αυξανόμενου όγκου των δημοσιεύσεων,

²¹⁷ Βλ. τα Συμπεράσματα του Συμβουλίου της ΕΕ για τη συμπεριληπτική, βιώσιμη και ανθρωποκεντρική στρατηγική για την υιοθέτηση της ΤΝ στην επιστήμη (23/05/2025). Η Ευρωπαϊκή Στρατηγική για την ΤΝ στην Επιστήμη στοχεύει να αναπτύξει και να βελτιώσει την πρόσβαση σε εργαλεία ΤΝ και σε υπολογιστικές υποδομές, να προσελκύσει και να καλλιεργήσει talenta, και να ενισχύσει την έρευνα σε στρατηγικούς τομείς όπως η κλιματική αλλαγή, η Υγεία και οι καθαρές τεχνολογίες.

²¹⁸ Προς αποφυγήν παρανοήσεων, θα πρέπει, ωστόσο, να διευκρινιστούν τα ακόλουθα: Η εξαίρεση δεν ισχύει όταν διεξάγονται δοκιμές σε πραγματικές συνθήκες και εφόσον το σύστημα δεν έχει διατεθεί ακόμη στην αγορά. Σε κάθε περίπτωση, εφόσον και όταν τεθεί ζήτημα αξιοποίησης των προϊόντων και αποτελεσμάτων της έρευνας, θα πρέπει να είναι σχεδιασμένα και υλοποιημένα έτσι ώστε να ανταποκρίνονται στις απαιτήσεις της νομοθεσίας.

²¹⁹ Η βραβευμένη με Νόμπελ επιτυχία του AlphaFold αποτελεί χαρακτηριστικό παράδειγμα της μετασχηματιστικής δύναμης της ΤΝ.

υπερβαίνοντας τον ρόλο του υποστηρικτικού εργαλείου, καθώς εξελίσσεται σε ενεργό «συμμέτοχο» στη διαδικασία παραγωγής και ερμηνείας της επιστημονικής γνώσης.

Για αυτούς ακριβώς τους λόγους η χρήση ΤΝ θέτει ηθικά ζητήματα και πρέπει να υπόκειται σε ηθικές αρχές και κανονιστικές απαιτήσεις.²²⁰ Ωστόσο, ο αντίκτυπος της ΤΝ δεν είναι αυτονόητα θετικός, καθώς εξαρτάται σε μεγάλο βαθμό από τις συνθήκες υπό τις οποίες αναπτύσσεται και διέπεται.²²¹ Εγείρονται ζητήματα ηθικής και δεοντολογίας της έρευνας, αλλά και συμμόρφωσης με κανονιστικές απαιτήσεις. Όσο περισσότερα δεδομένα και σύνολα δεδομένων χρησιμοποιούνται τόσο αυξάνονται και περιπλέκονται οι απαιτήσεις για την προστασία της ιδιωτικότητας και του απορρήτου. Η διαδικασία επιλογής, η αντιπροσωπευτικότητα, η ακρίβεια και η εν γένει ποιότητα των δεδομένων συνιστούν σημαντικές παραμέτρους για τη θεμιτή διεξαγωγή μιας έρευνας. Εξίσου σημαντικές είναι οι απαιτήσεις σεβασμού των κανόνων πνευματικής ιδιοκτησίας κατά τη χρήση δεδομένων για την εκπαίδευση των συστημάτων, αλλά και τη συναγωγή συμπερασμάτων.

Στο πλαίσιο αυτό, επισημαίνεται ο ηθικός σχεδιασμός (ethics by design), που ενσωματώνει τις ηθικές αρχές και απαιτήσεις από τα πρώτα στάδια του σχεδιασμού ενός ερευνητικού έργου. Όπως αναφέρει η Βασιλική Μολλάκη, «με αυτόν τον τρόπο, η τεχνολογία δεν αντιμετωπίζεται ως ουδέτερο εργαλείο, αλλά ως κοινωνικό προϊόν που οφείλει να σέβεται αξίες, όπως η ισότητα, η διαφάνεια και η λογοδοσία», προσθέτοντας ότι «οι ερευνητές καλούνται να σκεφτούν εκ των προτέρων τις κοινωνικές επιπτώσεις των συστημάτων ΤΝ, να σχεδιάσουν κατάλληλες δικλίδες ασφαλείας, να προβλέψουν πιθανές μεροληψίες και να διασφαλίσουν ότι το σύστημα θα λειτουργήσει με τρόπο που θα ευνοεί την κοινωνία».²²²

²²⁰ Διευκρινίζεται ότι η παρούσα μελέτη δεν καλύπτει τα ζητήματα που αφορούν τις απαιτήσεις για την ανάπτυξη της έρευνας, αλλά τα ηθικά, δεοντολογικά και νομικά ζητήματα που τίθενται.

²²¹ Purificato, E., Bili, D., Jungnickel, R., Ruiz-Serra, V., Fabiani, J., Abendroth-Dias, K., Fernández-Llorca, D., & Gómez, E. (2025). *The role of artificial intelligence in scientific research: A science for policy, European perspective* (JRC Technical Report No. JRC143482). Publications Office of the European Union. <https://doi.org/10.2760/7217497>

²²² Μολλάκη, Β. (2026, 3 Μαρτίου). Η σημασία της ηθικής και δεοντολογικής αξιολόγησης της έρευνας. Το Βήμα. <https://www.tovima.gr/2026/03/05/the-key-issue/i-simasia-tis-ithikis-kai-deontologikis-aksiologisis-tis-ereynas>

9.2. Η ποιότητα και η ακεραιότητα της έρευνας

Το ζήτημα της χρήσης της Τεχνητής Νοημοσύνης έχει και μια άλλη διάσταση, που αφορά την ακεραιότητα και την ποιότητα της έρευνας. Η ανασκόπηση βιβλιογραφίας, η ανάλυση κειμένων, η δημιουργία περιλήψεων φαίνεται να διευκολύνουν τον ερευνητή, αν και θα πρέπει να επισημανθεί ότι εν προκειμένω η ΤΝ «μεσολαβεί» στην κατανόηση και τη συναγωγή συμπερασμάτων.

Τα αποτελέσματα είναι ορατά: Ο αριθμός των δημοσιεύσεων έχει εκτοξευτεί, χωρίς να αντιστοιχεί απαραίτητα σε ποιότητα, ενώ το φαινόμενο των hallucinations έχει ειδικότερες διαστάσεις στην έρευνα και τη δημοσίευσή της. Καταρχάς, οι κατασκευασμένες πληροφορίες ή «παραισθήσεις», που όχι σπάνια παράγει η ΤΝ, ενδέχεται, εφόσον δεν εντοπίζονται, να στρεβλώσουν την επιστημονική γνώση και κατανόηση. Περαιτέρω παρατηρείται πλέον αυτό που ορισμένοι ονομάζουν «zombie citation», «ghost reference», «phantom paper» ή ακόμη και «Frankencitation», καθώς όλο και συχνότερα διαπιστώνουμε παραπομπές σε φανταστικούς συγγραφείς ή σε κείμενα ανύπαρκτα. Η διασφάλιση της επιβεβαίωσης και της ανθρώπινης εποπτείας εγγυάται την εγκυρότητα, αλλά και την πρόληψη προκαταλήψεων και μεροληψιών ως προς τα συναγόμενα αποτελέσματα. Ο Ευρωπαϊκός Κώδικας Δεοντολογίας για την Ακεραιότητα της Έρευνας (European Code of Conduct for Research Integrity), που συντάχθηκε από την All European Academies (ALLEA), θεσπίζει ένα σύνολο αρχών για την παραγωγή αξιόπιστης έρευνας. Οι αρχές αυτές περιλαμβάνουν ηθικές διαστάσεις της έρευνας και εστιάζουν:

- α) στην αξιοπιστία ως προς τη διασφάλιση της ποιότητας της έρευνας, όπως αυτή αποτυπώνεται στον σχεδιασμό, στη μεθοδολογία, στην ανάλυση και τη χρήση των πόρων,
- β) στην εντιμότητα κατά την ανάπτυξη, διεξαγωγή, αξιολόγηση, αναφορά και κοινοποίηση της έρευνας με διαφάνεια, δικαιοσύνη, πληρότητα και αμεροληψία,
- γ) στον σεβασμό προς τους συναδέλφους, τους συμμετέχοντες στην έρευνα, τα ερευνητικά υποκείμενα, την κοινωνία, τα οικοσυστήματα, την πολιτιστική κληρονομιά και το περιβάλλον, και

δ) στη λογοδοσία για την έρευνα από τη σύλληψη της ιδέας έως τη δημοσίευση, για τη διαχείριση και την οργάνωσή της, για την εκπαίδευση, την επίβλεψη και την καθοδήγηση, καθώς και για τις ευρύτερες κοινωνικές επιπτώσεις της.²²³

Πέραν των ζητημάτων της ακαδημαϊκής ακεραιότητας²²⁴ (αντιγραφή των αποτελεσμάτων, σύνταξη εργασιών και άρθρων μέσω ΤΝ), εγείρονται ζητήματα ως προς τις επιπτώσεις στην ανθρώπινη δημιουργικότητα και στην κριτική σκέψη, που συνιστούν τον πυρήνα της ερευνητικής διαδικασίας. Τα συστήματα ΤΝ μπορούν να αναγνωρίζουν μοτίβα, να βελτιστοποιούν αποτελέσματα και να προσομοιώνουν την ανθρώπινη σκέψη και έκφραση. Πηγάζουν από την ανάλυση και την αναπαραγωγή της γλώσσας, αλλά δεν είναι απόρροια βιωματικής εμπειρίας. Προσφέρουν συσχετίσεις και λύσεις, αλλά δεν διαθέτουν εμπειρία και δομή των αιτιωδών σχέσεων. Δεν μπορούν να κατανοούν τις εμπειρίες και να ερμηνεύουν. Συνοψίζουν προϋφιστάμενες γνώσεις και πληροφορίες, αλλά υπολείπονται στην επιτέλεση λογικών σκέψεων (reasoning) και δεν κατανοούν το νόημα. Παράγεται, όμως, έτσι νέα επιστημονική γνώση; Η επιστημονική έρευνα στηρίζεται στην αμφισβήτηση, στην πρωτότυπη σκέψη, στη μεθοδικότητα και στην ερμηνευτική ικανότητα, στοιχεία που δεν μπορούν να ανατεθούν μηχανικά σε ένα σύστημα παραγωγής κειμένου.

Όπως επισημαίνεται στη μελέτη των Purificato et al., «έχει εντοπιστεί ένα νέο πρόβλημα: ο κίνδυνος της επιστημικής μετατόπισης (epistemic drift). Το φαινόμενο αυτό, το οποίο οφείλεται στη συνδημιουργία γνώσης με τις μηχανές, αναφέρεται σε μια θεμελιώδη μεταβολή του τι θεωρείται έγκυρη επιστημονική γνώση και του τρόπου με τον οποίο αυτή παράγεται. Αυτό μπορεί να εκδηλωθεί με δύο τρόπους: Είτε οι τεχνολογίες ΤΝ, ακούσια, ενισχύουν τα καθιερωμένα ερευνητικά παραδείγματα και περιορίζουν την ποικιλία των ερωτημάτων που διερευνώνται είτε οι τεχνολογίες ΤΝ καλλιεργούν μια κουλτούρα, στην οποία τα επιστημονικά συμπεράσματα αποσυνδέονται από τις ανθρώπινες πηγές και τους δημιουργούς τους, με αποτέλεσμα η παραγωγή γνώσης να αποσπάται από την ανθρώπινη εποπτεία και τον έλεγχο».²²⁵

Πώς μπορεί να υποστηριχθεί η έρευνα μέσω της ΤΝ χωρίς να απολήξουμε σε «μαρασμό», στην κοινοτοπία ή και στη μίμηση, στην προσομοίωση της σκέψης; Θα ήταν φενάκη να θεωρήσουμε ότι μπορεί να αναστραφεί η τάση και να αποκλειστεί η χρήση της ΤΝ. Είναι, ωστόσο, αναγκαίο να οριοθετηθεί

²²³ All European Academies (ALLEA) (2023). *The European Code of Conduct for Research Integrity – Revised Edition 2023*. <https://doi.org/10.26356/ECOC>

²²⁴ «Certainly, here is a possible introduction for your topic...», μια γλώσσα αρκετά οικεία σε έναν μέσο χρήστη ενός γλωσσικού μοντέλου. Με αυτόν τον τρόπο ξεκινούσε άρθρο που δημοσιεύτηκε σε ένα επιστημονικό περιοδικό της Elsevier το 2024, πυροδοτώντας τη συζήτηση για τη χρήση, κατάχρηση, ελεγχσιμότητα και τον έλεγχο της χρήσης ΤΝ.

²²⁵ Purificato, E., Bili, D., Jungnickel, R., Ruiz-Serra, V., Fabiani, J., Abendroth-Dias, K., Fernández-Llorca, D., & Gómez, E. (2025). *The role of artificial intelligence in scientific research: A science for policy, European perspective* (JRC Technical Report No. JRC143482). Publications Office of the European Union. <https://doi.org/10.2760/7217497>

και να ελεγχθεί ο ρόλος της. Η ΤΝ μπορεί να λειτουργήσει ως υποστηρικτικό εργαλείο στην οργάνωση υλικού, στην τεχνική υποβοήθηση ή στη γλωσσική επιμέλεια, δεν πρέπει, όμως, να υποκαθιστά την επιστημονική κρίση, την ερευνητική ευθύνη και την ουσιαστική διανοητική συμβολή του ερευνητή/συγγραφέα.

Μια πρώτη απαίτηση αφορά την ανάπτυξη και καλλιέργεια μιας νέας ερευνητικής παιδείας που θα αποσκοπεί στην εκπαίδευση της κριτικής και υπεύθυνης χρήσης της ΤΝ. Πανεπιστήμια και ερευνητικά κέντρα πρέπει να αναπτύξουν επεξεργασμένες και σαφείς πολιτικές κατάρτισης, καθώς και οριοθέτησης της χρήσης, ώστε να εκπαιδεύουν φοιτητές, αλλά και επιστήμονες να αποκρούουν την ευκολία του έτοιμου και να διαθέτουν τις γνώσεις και την αντίληψη του κριτικού ελέγχου του αποτελέσματος και της επαλήθευσης πηγών και παραπομπών. Η ανάγκη για ουσιαστικό έλεγχο της ποιότητας και της αυθεντικότητας της επιστημονικής εργασίας είναι ευθέως ανάλογη της ευκολίας παραγωγής φαινομενικά άρτιας επιστημονικής έρευνας και λόγου. Η επιδίωξη και επίτευξη της διαφάνειας είναι μια, επίσης, σημαντική απαίτηση: Οι ερευνητές/συγγραφείς οφείλουν να δηλώνουν με σαφήνεια εάν και σε ποιο στάδιο χρησιμοποίησαν εργαλεία ΤΝ, ώστε οι αξιολογητές και οι αποδέκτες της έρευνας, αλλά και το κοινό να γνωρίζουν τα όρια της ανθρώπινης και της μηχανικής συμβολής.

9.3. Θέτοντας τις βάσεις για μια νέα εθνική πολιτική στην έρευνα και την καινοτομία της Τεχνητής Νοημοσύνης

Εκτός από κινητήρας που τροφοδοτεί τη διεθνή οικονομική ανάπτυξη και καταλυτής για την ψηφιοποίηση του δημόσιου τομέα και την αλλαγή του τρόπου με τον οποίο ζουν, εργάζονται και αλληλεπιδρούν οι κοινωνίες, η Τεχνητή Νοημοσύνη έχει καταστεί πεδίο γεωοικονομικού ανταγωνισμού, εργαλείο στρατηγικής αυτονομίας και δείκτης τεχνολογικής κυριαρχίας. Οι χώρες που επενδύουν συστηματικά στην έρευνα για να αναπτύξουν ΤΝ σήμερα, διαμορφώνουν την οικονομική και πολιτική τους θέση για τις επόμενες δεκαετίες.

Ποια είναι θέση της χώρας μας σε αυτήν την ιστορική συγκυρία; Χωρίς να διαθέτουμε τις μεγάλες εταιρείες και επιστημονικά εργαστήρια που πρωτοπορούν, δεν υστερούμε σε επιστημονικό ταλέντο, ιδέες και ακαδημαϊκές διακρίσεις. Αν η χώρα θέλει να αποκτήσει ουσιαστική παρουσία στην έρευνα για την ανάπτυξη της ΤΝ, οφείλει πρώτα να αναγνωρίσει τις δομικές παθογένειες του ίδιου της του συστήματος έρευνας & καινοτομίας.

Πιο συγκεκριμένα, στην Ελλάδα, η έρευνα παραμένει ένα από τα πλέον υποχρηματοδοτούμενα και πλέον γραφειοκρατικοκρατούμενα πεδία άσκησης πολιτικής. Η Ελλάδα αντιμετωπίζει κρίσιμες προκλήσεις και χρόνια προβλήματα στο πεδίο αυτό, ειδικά στο πεδίο της Τεχνητής Νοημοσύνης και των ψηφιακών τεχνολογιών εν γένει: Η χρηματοδότηση της έρευνας είναι κατακερματισμένη και έχει κατακριθεί επανειλημμένα για έλλειψη διαφάνειας, γραφειοκρατικά εμπόδια, ενώ η διαρροή ταλέντων συνεχίζεται, καθώς νέοι επαγγελματίες αναζητούν ευκαιρίες στο εξωτερικό. Ταυτόχρονα, ο κατακερματισμός δεδομένων κυριαρχεί σε όλα σχεδόν τα επιστημονικά πεδία, καθώς συχνά είναι διασκορπισμένα σε διαφορετικές πλατφόρμες με ασύμβατα πρότυπα, ενώ απουσιάζει ένα κοινό πλαίσιο διαλειτουργικότητας.

Πολλές φορές ακούμε ότι οι χαμηλές πτήσεις στον χώρο της ΤΝ οφείλονται στο ότι η χώρα δεν διαθέτει το μέγεθος ή τους πόρους άλλων χωρών. Όμως το μέγεθος δεν αποτελεί καθοριστικό παράγοντα επιτυχίας. Ορισμένες από τις πιο επιτυχημένες στρατηγικές στην έρευνα για την ΤΝ προέρχονται από μικρές και μεσαίες χώρες, που κατάφεραν να αξιοποιήσουν κατά τρόπο στοχευμένο τα πλεονεκτήματά τους.

Η κυρίαρχη αφήγηση θέλει την ΤΝ να αποτελεί πεδίο αποκλειστικά των Ηνωμένων Πολιτειών και της Κίνας. Ωστόσο, χώρες όπως ο Καναδάς, το Ισραήλ, η Σιγκαπούρη, η Φινλανδία, η Ολλανδία, η Εσθονία και η Ιρλανδία απέδειξαν ότι η επιτυχία δεν εξαρτάται από το μέγεθος, αλλά από τον σχεδιασμό.

Το παράδειγμα του Καναδά είναι διεθνές σημείο αναφοράς ως προς την ικανότητα στοχευμένης συγκέντρωσης πόρων. Το 2017, ο Καναδάς υιοθέτησε την Pan-Canadian AI Strategy μέσω του Canadian Institute for Advanced Research (CIFAR). Αντί να διασπείρει πόρους σε όλα τα πανεπιστήμια, συγκέντρωσε επενδύσεις σε τρεις ισχυρούς κόμβους: Τορόντο, Μόντρεαλ και Έντμοντον. Η στρατηγική αυτή δημιούργησε κρίσιμη μάζα, διεθνή αναγνωσιμότητα και προσέλκυση κορυφαίων επιστημόνων.

Στην Ελλάδα, αντιθέτως, η χρηματοδότηση συχνά διαχέεται σε μικρά, αποσπασματικά έργα, πολλές φορές υπό την πίεση γεωγραφικών ή πολιτικών ισορροπιών. Η λογική της «ισομερούς κατανομής» υπονομεύει σε πολλές περιπτώσεις την αριστεία. Επομένως, στην έρευνα για την ανάπτυξη της ΤΝ έχει αποδειχθεί ότι η διασπορά ισοδυναμεί με απώλεια ανταγωνιστικότητας.

Μια χώρα σαν το Ισραήλ έχει καταφέρει να βρεθεί στην πρωτοπορία της καινοτομίας μέσω μιας στρατηγικής στοχευμένης εξειδίκευσης. Τι σημαίνει αυτό; Το Ισραήλ δεν επεδίωξε να καλύψει όλα τα πεδία της ΤΝ. Επένδυσε σε συγκεκριμένους τομείς, όπως η κυβερνοασφάλεια και τα αμυντικά συστήματα, αξιοποιώντας ιδρύματα, όπως το Weizmann Institute of Science. Η εστίαση αυτή το βοήθησε να δημιουργήσει ένα παγκόσμιας κλάσης οικοσύστημα. Η Ελλάδα, αντίθετα, τείνει να διατυπώνει ευρείες στρατηγικές χωρίς σαφή ιεράρχηση.

Η σταθερότητα της χρηματοδότησης είναι καθοριστικής σημασίας και αυτή είναι η περίπτωση της Φινλανδίας: Μέσω οργανισμών, όπως το Business Finland, έχει διασφαλίσει μακροπρόθεσμη και σταθερή χρηματοδότηση της έρευνας. Στην Ελλάδα, η εξάρτηση από κύκλους ευρωπαϊκής χρηματοδότησης και οι δημοσιονομικές διακυμάνσεις δημιουργούν αστάθεια. Η έρευνα ΤΝ απαιτεί ορίζοντα δεκαετίας, όχι ετήσιες αποσπασματικές ενισχύσεις.

Θα μπορούσαμε να αντλήσουμε κάποια μαθήματα από αυτές τις περιπτώσεις; Κάθε περίπτωση είναι ασφαλώς ξεχωριστή, αλλά τα συγκεκριμένα διεθνή παραδείγματα καταρρίπτουν τον μύθο περί του μεγέθους της χώρας ή της προνομιακής οικονομικής θέσης ως τα απόλυτα εκέγγυα καινοτομίας. Για την Ελλάδα, μια χώρα με ανθρώπινο κεφάλαιο μεγάλων δυνατοτήτων και αυξανόμενες τεχνολογικές φιλοδοξίες, μια συνεκτική εθνική πολιτική για την έρευνα και την καινοτομία στην ΤΝ δεν είναι απλώς επιθυμητή, είναι πλέον απαραίτητη. Χωρίς αυτήν, η Ελλάδα κινδυνεύει να μείνει πίσω στον ευρωπαϊκό και παγκόσμιο αγώνα της ΤΝ, να χάσει ευκαιρίες για οικονομική ανάπτυξη, κοινωνική καινοτομία και διεθνή επιρροή. Η Ελλάδα πρέπει να

ενεργήσει άμεσα για να παραδειγματιστεί από τις βέλτιστες διεθνείς πρακτικές στον τομέα της πολιτικής διαχείρισης της έρευνας, αναπτύσσοντας παράλληλα μια μοναδική εθνική προσέγγιση που θα αξιοποιεί τα δυνατά της σημεία και θα αντιμετωπίζει τους συγκεκριμένους περιορισμούς της.

Μια εθνική πολιτική ΤΝ της Ελλάδας θα πρέπει να στοχεύει στο να καταστήσει τη χώρα περιφερειακό ηγέτη στην αξιόπιστη, ανθρωποκεντρική ΤΝ, που θα αξιοποιεί την τεχνολογία αυτή για να προωθήσει την καινοτομία, την οικονομική ανάπτυξη και την κοινωνική ευημερία. Πώς θα γίνει αυτό; Με τη δημιουργία περισσότερων data centers; Με τη δημιουργία περισσότερων φορέων και επιτροπών για τη διακυβέρνηση αυτής της τεχνολογίας; Με την περαιτέρω εμπλοκή των περίφημων big tech στις δημόσιες προμήθειες και σε δράσεις εκπαίδευσης; Με τη συχνή παρουσία σε διεθνή συνέδρια και συνόδους; Ή υπάρχει και ένας άλλος δρόμος;

Ένα πιθανό εναλλακτικό μοντέλο θα μπορούσε να στηριχθεί κυρίως στην επιστημονική έρευνα, στην ανοιχτή επιστήμη, στη συνεργασία πανεπιστημίων, κοινωνίας και επιχειρήσεων, και στην παραγωγή δημόσιας γνώσης, το οποίο δεν θα βασίζεται αποκλειστικά σε ιδιωτικές πλατφόρμες, όπως Google, Microsoft και Amazon, αλλά θα βασίζεται στην ανοιχτή επιστήμη, στα ανοιχτά δεδομένα και στον ανοιχτό κώδικα, όπως και στη συνεργατική καινοτομία, στην τεχνολογική αυτονομία και στον κοινωνικό προσανατολισμό της ΤΝ. Η στρατηγική αυτή θα μπορούσε να επικεντρωθεί στην ανάπτυξη δημόσιων ψηφιακών υποδομών και μοντέλων ανοιχτού κώδικα, διασφαλίζοντας την ανοιχτή πρόσβαση στη γνώση, στην υιοθέτηση πλαισίων data commons για τη συλλογική διαχείριση των δεδομένων και την ενίσχυση της ερευνητικής βάσης των πανεπιστημίων και των ερευνητικών κέντρων.

Πιο συγκεκριμένα, βασική προτεραιότητα είναι η δημιουργία ενός εθνικού δικτύου ερευνητικών κέντρων ΤΝ, που θα συνδέει πανεπιστήμια, ερευνητικά ιδρύματα, δημόσιους οργανισμούς και ΜμΕ, καθώς και η δημιουργία μιας εθνικής πλατφόρμας ερευνητικών δεδομένων, που θα παρέχει ανοιχτή πρόσβαση για πανεπιστήμια και ερευνητές, και θα λειτουργεί επί τη βάση διαλειτουργικών προτύπων δεδομένων. Στόχος θα πρέπει να είναι η δημιουργία ενός συνεργατικού οικοσυστήματος γνώσης. Μια φιλόδοξη εθνική στρατηγική που θα φιλοδοξούσε να αφήσει ένα διαφορετικό στίγμα θα μπορούσε να στοχεύσει στην ανάπτυξη δημόσιων ψηφιακών υποδομών ΤΝ. Κάτι τέτοιο προϋποθέτει τη δημιουργία μιας εθνικής πλατφόρμας ανοιχτών δεδομένων, ανοιχτών μοντέλων AI (open source), δημόσιων Διεπαφών Προγραμματισμών Εφαρμογών (Application Programming Interfaces – APIs) για έρευνα και καινοτομία, και κοινών υποδομών για πανεπιστήμια και startups. Η Εσθονία, για παράδειγμα, έχει αναπτύξει μεγάλο μέρος της ψηφιακής διοίκησής της πάνω σε ανοιχτά πρότυπα και κοινή δημόσια ψηφιακή υποδομή.

Μια τέτοια στρατηγική θα μπορούσε να στηριχθεί στη δημιουργία «κοινών δεδομένων» και συνεταιρισμών δεδομένων, όπου οι πολίτες ή οι κοινότητες θα μπορούν να ελέγχουν συλλογικά τα δεδομένα τους μέσω της δημιουργίας «data trusts», στα οποία οι επιχειρήσεις και οι ερευνητές θα έχουν δίκαιη πρόσβαση. Το πλεονέκτημα μιας τέτοιας προσέγγισης είναι ότι προσφέρει μεγαλύτερη διαφάνεια και εγγυάται καλύτερη εκπροσώπηση της κοινωνίας στα δεδομένα. Ο Δήμος της Βαρκελώνης ήδη έχει θέσει σε λειτουργία διάφορα μοντέλα δημοκρατικής διακυβέρνησης δεδομένων. Προς αυτήν την κατεύθυνση θα μπορούσαν τα ελληνικά ερευνητικά ιδρύματα να αναπτύξουν μοντέλα επεξεργασίας φυσικής γλώσσας για την ελληνική, εργαλεία ανάλυσης δεδομένων και μοντέλα για δημόσιες πολιτικές. Τα μοντέλα αυτά πρέπει να είναι open source, ώστε να αξιοποιούνται από startups, δημόσιους οργανισμούς και ερευνητές.

Η εναλλακτική στρατηγική που προτείνεται δεν θα μπορούσε παρά να δώσει έμφαση στις ΜμΕ, που αποτελούν τη ραχοκοκαλιά της ελληνικής οικονομίας, μέσω της δημιουργίας προσβάσιμων δημόσιων datasets, συνεργατικών clusters, δημόσιων προμηθειών καινοτομίας και ειδικών θερμοκοιτίδων (sandboxes) για νέες τεχνολογίες. Χωρίς μια τέτοια στόχευση, η χώρα κινδυνεύει να παραμείνει καταναλωτής παρά παραγωγός καινοτομίας ΤΝ.

Μια τέτοια στρατηγική δεν θα πρέπει να βασιστεί στη λογική της προσέλκυσης υπερμεγέθων data centers (η οποία συνοδεύεται από υψηλή κατανάλωση ενέργειας, μικρή τοπική απασχόληση, χρήση μεγάλων εκτάσεων γης και εξάρτηση από ξένες εταιρείες), αλλά να προτάξει ένα υπόδειγμα κατανομημένης προώθησης της υπολογιστικής ισχύος (edge computing) και ενεργειακά αποδοτικών μοντέλων ΤΝ. Αντί για γιγαντιαία εμπορικά data centers, μια φιλόδοξη εθνική στρατηγική θα μπορούσε να επενδύσει σε εθνικά ερευνητικά υπολογιστικά κέντρα, που θα προωθούν την ανοικτή τεχνολογική καινοτομία και τις κοινωνικά ωφέλιμες εφαρμογές της ΤΝ.

Εάν, επομένως, καταφέρουμε να προτείνουμε ένα καινοτόμο πλαίσιο κανόνων και κινήτρων για την υπεύθυνη και κοινωνιοκεντρική χρήση της ΤΝ, θα έχουμε κάνει ένα σημαντικό βήμα για να αποκτήσουν οι ερευνητές και οι νεοφυείς επιχειρήσεις μια πραγματική πυξίδα που θα τους βοηθήσει να πορευτούν με ασφάλεια στα αχαρτογράφητα νερά της νέας τεχνολογικής εποχής. Ταυτόχρονα, θα έχουμε καταφέρει να βάλουμε τα θεμέλια για την οικοδόμηση ενός στιβαρού εθνικού προτύπου υπεύθυνης καινοτομίας, που θα προωθεί τη διαφάνεια και την ηθική στη χρήση αυτής της συγκεκριμένης τεχνολογίας και θα βασίζεται σε ανεξάρτητους μηχανισμούς εποπτείας για τον έλεγχο των αλγορίθμων, διασφαλίζοντας τη διαφάνεια στη χρήση ΤΝ από το κράτος και την ουσιαστική συμμετοχή των πολιτών.

Σημαντικό ρόλο θα παίξει και η ικανότητά μας, παράλληλα με την αξιοποίηση των χρηματοδοτικών ευκαιριών που μας προσφέρουν οι νέες Στρατηγικές

της Ευρωπαϊκής Ένωσης (Apply AI και AI in Science), να μπούμε σε θέση οδηγού, θεσπίζοντας μια εθνική στρατηγική που θα περιλαμβάνει τη χρήση ανοικτού λογισμικού, δημόσιες υποδομές cloud, διαλειτουργικά πρότυπα και μια πολυπρομηθευτική στρατηγική, η οποία θα υπηρετεί την κοινωνία, την οικονομία των μικρών επιχειρήσεων και τη δημοκρατία. Σε ένα παγκόσμιο περιβάλλον που χαρακτηρίζεται από υπερβολική συγκέντρωση δεδομένων, γιγαντιαία υπολογιστική ισχύ και την κυριαρχία μεγάλων τεχνολογικών εταιρειών, η Ελλάδα θα μπορούσε να προτάξει ένα εναλλακτικό μοντέλο, που θα βασίζεται σε ανοιχτά δεδομένα, συνεργατικά οικοσυστήματα, δημόσιες ψηφιακές υποδομές και σε έναν πιο ξεκάθαρα κοινωνικό προσανατολισμό της τεχνολογίας.

Ιδιαίτερης βαρύτητας για την επίτευξη των παραπάνω στόχων θα είναι και ο ρόλος ενός από τα πρώτα επτά εργοστάσια Τεχνητής Νοημοσύνης (AI Factories) της Ευρώπης, του «Pharos». Η ανεξαρτησία ενός οργανισμού, όπως ο «Pharos», είναι κρίσιμη, ώστε να μπορεί να λειτουργεί ως επιστημονικός κόμβος δημοσίου συμφέροντος και όχι ως εργαλείο είτε της κυβέρνησης είτε μεγάλων εταιρειών. Για να διασφαλιστεί αυτή η ανεξαρτησία χρειάζεται ένας συνδυασμός θεσμικών, χρηματοδοτικών και οργανωτικών μηχανισμών, που θα διασφαλίζουν διοικητική αυτονομία, επιστημονική ανεξαρτησία και διαφάνεια στη λειτουργία, και να εξασφαλιστεί ότι στη διοίκησή του θα μετέχουν πανεπιστήμια, ερευνητικά ιδρύματα, εκπρόσωποι της Κοινωνίας των Πολιτών και διεθνείς επιστήμονες.

Θα μπορούσε, μάλιστα, να διαθέτει μια διεθνή επιστημονική επιτροπή που θα αξιολογεί την έρευνα, θα καθορίζει επιστημονικές προτεραιότητες και θα διασφαλίζει ποιότητα και ανεξαρτησία. Θα πρέπει, επίσης, να λειτουργεί με δημόσιες εκθέσεις δραστηριότητας, ανοιχτές διαδικασίες χρηματοδότησης και ανεξάρτητους ελέγχους που θα ενισχύουν την εμπιστοσύνη της κοινωνίας. Ένας τέτοιος ανεξάρτητος φορέας θα μπορούσε να οδηγήσει στη βιώσιμη και δημοκρατική ανάπτυξη της ΤΝ, και να λειτουργήσει και ως κόμβος κοινωνικής καινοτομίας, υποστηρίζοντας κοινωνικές startups και επιτρέποντας τη συμμετοχή πολιτών στη διαμόρφωση τεχνολογίας μακριά από πολιτικούς παρεμβατισμούς. Ο κίνδυνος για την Ελλάδα είναι να περιοριστεί σε ρητορικές διακηρύξεις και εξαγγελίες, χωρίς να μετουσιωθούν στις απαραίτητες δομικές μεταρρυθμίσεις και στην πλαισίωση των προτεινόμενων δομών διοίκησης από ανθρώπους που έχουν διεθνή εμπειρία και θα τις στελεχώσουν βάσει διαφανών διαδικασιών και ανοιχτών προκηρύξεων.

Μια εθνική στρατηγική AI βασισμένη στους προαναφερόμενους πυλώνες μπορεί να δημιουργήσει ένα ανθεκτικό και κοινωνικά χρήσιμο οικοσύστημα ΤΝ, με τον «Pharos» ως τον κεντρικό θεσμό που συντονίζει την έρευνα, τα δεδομένα και την καινοτομία. Η ανάπτυξη μιας ισχυρής εθνικής στρατηγικής που θα προσεγγίζει την ΤΝ όχι μόνο ως οικονομική τεχνολογία, αλλά ως εργαλείο δημόσιας πολιτικής και θα εστιάζει σε ισχυρή δημόσια έρευνα και

σε εφαρμογές που θα έχουν καθαρά κοινωνική αξία μπορεί να συμβάλει ουσιαστικά στην αναχαίτιση του brain drain, δημιουργώντας ένα ελκυστικό και αξιοκρατικό περιβάλλον έρευνας και καινοτομίας. Μέσω της ενίσχυσης της δημόσιας χρηματοδότησης της έρευνας, της δημιουργίας ανεξάρτητων κέντρων αριστείας και της προώθησης ανοιχτών και διαφανών διαδικασιών αξιολόγησης, μπορούν να δημιουργηθούν ουσιαστικές ευκαιρίες για νέους επιστήμονες, αλλά και για ανθρώπους της τεχνολογίας που θα επιλέγονταν μέσω διεθνών διαγωνισμών και θα μπορούσαν να προσφέρουν τις γνώσεις και τις εμπειρίες τους στη διαχείριση τέτοιων δομών. Παράλληλα, η σύνδεση της έρευνας με κοινωνικές ανάγκες και καινοτόμες εφαρμογές μπορεί να ενισχύσει τη δημιουργία νέων θέσεων εργασίας υψηλής εξειδίκευσης και να θέσει την Ελλάδα στην πρωτοπορία της υπεύθυνης καινοτομίας.

Το ερώτημα είναι: Υπάρχει η πολιτική βούληση να ξεπεραστούν χρόνιες αδυναμίες και παθογένειες και να γίνει ουσιαστική επένδυση σε δημόσια έρευνα και υποδομές, που θα διέπεται από μια κοινωνική προσέγγιση της συγκεκριμένης τεχνολογίας; Η επιλογή είναι σαφής: Ή θα οικοδομήσουμε ένα ανεξάρτητο, εστιασμένο, εξωστρεφές και αξιόπιστο οικοσύστημα έρευνας ΤΝ ή θα παραμείνουμε θεατές μιας τεχνολογικής επανάστασης που διαμορφώνεται αλλού, εγκλωβισμένοι σε χρόνιες αβελτηρίες και μυωπικές αντιλήψεις.

[ΠΩΣ] ΜΠΟΡΕΙ Η ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΝΑ ΕΙΝΑΙ
ΗΘΙΚΗ ΚΑΙ ΣΥΝΝΟΜΗ;
Ιούλιος 2026

Τεχνητή Νοημοσύνη, δημιουργία και πνευματική ιδιοκτησία



Δημήτρης Γραμμένος, Λίλιαν Μήτρου

Συγγραφική συνεισφορά ανά υποκεφάλαιο:

10.1. – Δημήτρης Γραμμένος

10.2. – Λίλιαν Μήτρου

Η Τεχνητή Νοημοσύνη δεν έχει διεισδύσει μόνο στην εκπαίδευση ή στην έρευνα. Διευρύνει και μεταλλάσσει τις δυνατότητες της «δημιουργίας» με την παραγωγή «έργων» ή «μορφών περιεχομένου», που μέχρι πρότινος αποτελούσαν μονοπώλιο της ανθρώπινης δημιουργικότητας. Είναι η Παραγωγική ΤΝ μια νέα μορφή «εκδημοκρατισμού» και «μαζικοποίησης» της (δυνατότητας) «δημιουργίας»;

Η Παραγωγική ΤΝ φαίνεται να μεταβάλλει, σταδιακά αλλά σταθερά, την αντίληψή μας για την έννοια της δημιουργίας και του δημιουργού, της πρωτοτυπίας, της έμπνευσης, της μοναδικότητας, της διανοητικής και ψυχικής διεργασίας, που είναι απόρροια συλλογισμών, αλλά και προσωπικών βιωμάτων, συναισθημάτων, συνειδητής προσπάθειας, αποτίμησης εμπειρίας, αξιακών επιλογών και όχι μια απλή, τυχαία ανασύνθεση δεδομένων, μοτίβων, στιλ. Η σύλληψη, η δημιουργικότητα μετατοπίζονται ή εκπίπτουν στο φιλτράρισμα, στη συσχέτιση, στην επιμέλεια, στην έξυπνη «κατεύθυνση» μέσω των κατάλληλων prompts. Ο άνθρωπος της ανθρωποκεντρικής-δημιουργοκεντρικής αντίληψης θα εμπλουτιστεί ή θα αποψιλωθεί μέσω της αλληλεπίδρασης με τη μηχανή; Παρακάτω, ο Δημήτρης Γραμμένος ανιχνεύει τη σχέση μεταξύ ΤΝ και ανθρώπινης δημιουργίας και δημιουργικότητας.

Είναι προφανές ότι επαναπροσδιορίζεται η σχέση ανθρώπου, τέχνης και τεχνολογίας. Αν η Παραγωγική ΤΝ επιτρέπει να εξελιχθούμε όλοι ή, έστω, πολλοί σε «δημιουργούς» έργων, αν καθιστά τη δημιουργία πιο προσιτή και πιο συλλογική, ελλοχεύει ο φόβος της «υποκατάστασης», αλλά και της ομογενοποίησης, της ανακύκλωσης του παρελθόντος σε βάρος της δημιουργίας που καινοτομεί και δεν υποτάσσεται σε προϋπάρχουσες νόρμες και φόρμες;

10.1. Musa ex machina: Τεχνητή Νοημοσύνη και ανθρώπινη δημιουργικότητα

Οι εξελίξεις των τελευταίων ετών στον τομέα της Τεχνητής Νοημοσύνης έχουν επιτρέψει την ανάπτυξη συστημάτων ικανών να παράγουν δημιουργικά αποτελέσματα, τα οποία συχνά είναι συγκρίσιμα ή ακόμα και ανώτερα από αυτά των ανθρώπων. Πρόσφατες μελέτες οι οποίες συγκρίνουν τις δημιουργικές ικανότητες της ΤΝ με αυτές των ανθρώπων, χρησιμοποιώντας τόσο τυποποιημένες δοκιμασίες αξιολόγησης της δημιουργικότητας όσο και λιγότερο τυπικές προσεγγίσεις, καταλήγουν στο κοινό συμπέρασμα ότι συχνά δεν υπάρχει σημαντική διαφορά όσον αφορά στο τελικό αποτέλεσμα, ενώ σε αρκετές περιπτώσεις η ΤΝ υπερτερεί. Προς το παρόν, τουλάχιστον, οι άνθρωποι τείνουν να διατηρούν το προβάδισμα όσον αφορά τις υψηλότερες επιδόσεις, την πρωτοτυπία, την ποικιλομορφία και την αξιοπιστία των αποτελεσμάτων.

Ο συχνός χαρακτηρισμός της Παραγωγικής ΤΝ (Generative AI ή GenAI) ως «δημιουργικής» ώθησε τον Marc Runco, γνωστικό ψυχολόγο και κορυφαίο ερευνητή στον τομέα της δημιουργικότητας, να αναθεωρήσει τον ευρέως χρησιμοποιούμενο τυπικό ορισμό που είχε εισαγάγει το 2012, σύμφωνα με τον οποίο η δημιουργικότητα προϋποθέτει δύο βασικά χαρακτηριστικά: *πρωτοτυπία* και *αποτελεσματικότητα*. Εφόσον και τα δύο χαρακτηριστικά αξιολογούνται αποκλειστικά βάσει του εκάστοτε δημιουργικού αποτελέσματος, η ΤΝ είναι ικανή να επιτυγχάνει συγκρίσιμες επιδόσεις. Ωστόσο, εξακολουθεί να στερείται ορισμένων άλλων χαρακτηριστικών που αποτελούν συστατικά στοιχεία της ανθρώπινης δημιουργικότητας, όπως, π.χ., εσωτερικά κίνητρα, σκοπιμότητα, αυθεντικότητα, ανακάλυψη προβλημάτων. Επομένως, σύμφωνα με τον Marc Runco, η ΤΝ μπορεί να παράγει μόνο «τεχνητή δημιουργικότητα». Προκειμένου, λοιπόν, αυτή να διακριθεί από την ανθρώπινη, πρότεine την επικαιροποίηση του τυπικού ορισμού ώστε να συμπεριλαμβάνει επιπρόσθετα την *αυθεντικότητα* και τη *σκοπιμότητα*. Επιπλέον αυτής της διάστασης, αρκετοί επιστήμονες τονίζουν ότι, παρόλο που τα τελικά προϊόντα της τεχνητής και της ανθρώπινης δημιουργικότητας ενδέχεται να είναι συγκρίσιμα ως προς την πρωτοτυπία και την αποτελεσματικότητα, οι αντίστοιχες εσωτερικές διεργασίες είναι θεμελιωδώς διαφορετικές, όπως επίσης και η εμπειρία (ικανοποίηση, απογοήτευση, χαρά κλπ.) που αντλείται (ή όχι, στην περίπτωση της ΤΝ) μέσα από αυτές.

Πέραν, όμως, των θεωρητικών προβληματισμών σχετικά με τον ρόλο και τη φύση της τεχνητής δημιουργικότητας, δεν είναι λίγοι αυτοί που οραματίζονται την ανάδυση μιας νέας μορφής υβριδικής νοημοσύνης, η οποία συχνά αναφέρεται ως «συν-δημιουργικότητα» (co-creativity). Σε αυτό το πλαίσιο, η ΤΝ νοείται ως ένας «συνεργάτης» που, αφενός, θα επιτρέψει στους επιστήμονες να μελετήσουν και να κατανοήσουν σε βάθος την ανθρώπινη δημιουργικότητα και, αφετέρου, θα βοηθήσει τους δημιουργούς να ξεπεράσουν τους περιορισμούς του ανθρώπινου νου και να κατακτήσουν, απροσπέλαστα προς το παρόν, δημιουργικά επιτεύγματα.

Μέρα με τη μέρα καθίσταται ολοένα και πιο σαφές ότι η ΤΝ, συνεργαζόμενη με ανθρώπους σε ποικίλους βαθμούς, αναπόφευκτα θα διαδραματίσει κεντρικό ρόλο στις περισσότερες (αν όχι σε όλες) τις δημιουργικές ανάγκες κάθε τομέα ανθρώπινης δραστηριότητας. Πρόσφατα, προτάθηκε η χρήση της λέξης «κυβερνο-δημιουργικότητα» (cyber-creativity) ως γενικού όρου που καλύπτει ολόκληρο το φάσμα των πιθανών σεναρίων συνεργασίας, από την πλήρη ανθρώπινη έως την πλήρη μηχανική αυτονομία. Το 2025, μια ανάλυση 1,5 εκατομμυρίου συνομιλιών με το ChatGPT αποκάλυψε ότι το 13% του συνόλου των μηνυμάτων που σχετίζονταν με την εργασία αφορούσε τη «δημιουργική σκέψη».

Επιπλέον, υπάρχουν βάσιμες ενδείξεις ότι η συνέργεια ανθρώπου-ΤΝ σε δημιουργικές εργασίες έχει συνήθως θετικό αποτέλεσμα, καθώς και ότι οι άνθρωποι που συνεργάζονται με την ΤΝ για δημιουργικές εργασίες υπερτερούν σημαντικά έναντι εκείνων που εργάζονται χωρίς τη βοήθειά της. Ωστόσο, προς το παρόν υπάρχουν ασαφή συμπεράσματα σχετικά με το αν η συν-δημιουργικότητα ανθρώπου-ΤΝ:

- α) δεν υπερτερεί ουσιαστικά έναντι των αντίστοιχων μεμονωμένων βέλτιστων ανθρώπινων ή υπολογιστικών επιδόσεων,
- β) ευνοεί άτομα με χαμηλότερα επίπεδα δημιουργικού δυναμικού και εξειδίκευσης,
- γ) ευνοεί άτομα με υψηλότερο επίπεδο δεξιοτήτων, ενώ δεν έχει καμία ή ακόμη και αρνητική επίδραση σε άτομα με χαμηλότερο επίπεδο δεξιοτήτων.

Το 1988, ο Mihaly Csikszentmihalyi παρουσίασε το «Συστημικό Μοντέλο της Δημιουργικότητας» (Systems Model of Creativity), το οποίο υποστηρίζει ότι η δημιουργικότητα είναι το αποτέλεσμα της αλληλεπίδρασης τριών μερών: του *ατόμου*, του *γνωστικού/πολιτισμικού τομέα* και του *κοινωνικού πεδίου*. Σύμφωνα με αυτό το μοντέλο, το άτομο προτείνει μια νέα ιδέα σε έναν τομέα, το πεδίο την αξιολογεί και, εφόσον την αποδεχθεί, η ιδέα αναγνωρίζεται ως «δημιουργική» και ενσωματώνεται στον τομέα, αποτελώντας πλέον μέρος της συλλογικής γνώσης.

Το 2019, οι Daniel T. Gruner και Mihaly Csikszentmihalyi επικαιροποίησαν το μοντέλο και το μετονόμασαν σε «Δημιουργικότητα 4.0» (Creativity 4.0). Σε αυτήν την εκδοχή, συμπεριέλαβαν την ΤΝ ως αυτόνομο συστατικό που αλληλεπιδρά με το άτομο και τον τομέα, αλλά όχι με το πεδίο. Το 2023, οι Paul Atkinson και Richie Barker υποστήριξαν ότι η εκδοχή δεν λαμβάνει υπ' όψιν την ΤΝ που είναι ενσωματωμένη σε καθημερινά συστήματα, ούτε το πώς αυτή επηρεάζει τις επικοινωνιακές πρακτικές. Έτσι, πρότειναν με τη σειρά τους μια νέα εκδοχή, η οποία ενσωματώνει τον ρόλο της ΤΝ ως φίλτρου ιδεών και θεματοφύλακα του τομέα, καθώς και ως υποστηρικτή της δημιουργικής πρακτικής, με αποτέλεσμα να ταυτίζεται σε μεγάλο βαθμό με το άτομο.

Τα τελευταία 25 χρόνια έχουν προταθεί διάφορα εννοιολογικά και πρακτικά πλαίσια για τη διερεύνηση του ρόλου της ΤΝ στη δημιουργική διαδικασία.²²⁶ Στα πρώιμα πλαίσια, ο άνθρωπος είναι ξεκάθαρα εκείνος που δρα δημιουργικά και αποτελεί συνήθως και τον αποκλειστικό κριτή του τελικού αποτελέσματος, ενώ η ΤΝ περισσότερο διευκολύνει παρά κατασκευάζει τα τελικά δημιουργικά αποτελέσματα, υιοθετώντας έναν υποστηρικτικό, μη αυτόνομο ρόλο. Το 1998, η Margaret Boden προσδιόρισε τρεις τομείς για την παραγωγή ιδεών με τη βοήθεια της ΤΝ: δημιουργία νέων συνδυασμών από οικείες ιδέες, εξερεύνηση των δυνατοτήτων εννοιολογικών χώρων και πραγματοποίηση μετασχηματισμών που επιτρέπουν τη γέννηση ιδεών που προηγουμένως ήταν αδύνατες. Η αντίληψη της ΤΝ ως εργαλείου αντικατοπτρίζεται, επίσης, στο έργο της Kumiyo Nakakoji, η οποία την οραματίστηκε να λειτουργεί μεταφορικά ως: αλτήρες (για την απόκτηση γνώσεων και ανάπτυξη δεξιοτήτων), παπούτσια τρεξίματος (για την επιτάχυνση διαδικασιών), πέδιλα του σκι (για την πρόσβαση σε νέες δυνατότητες). Αντίθετα, ο Todd Lubart το 2005 απέδωσε στην ΤΝ ρόλους εμπνευσμένους από ανθρώπινες ιδιότητες: νταντά (διαχείριση της εργασίας), φίλος δι' αλληλογραφίας (επικοινωνία με συνεργάτες), προπονητής (τεχνικές ενίσχυσης της δημιουργικότητας), συνάδελφος (συνεργασία στην παραγωγή ιδεών). Το 2012, η Mary Lou Maher επινόησε ένα πλαίσιο με παρόμοιους ρόλους για τα υπολογιστικά συστήματα (υποστήριξη, ενίσχυση, παραγωγή), αλλά επιπλέον εξέτασε την κλιμάκωση της αλληλεπίδρασης πέρα από την τυπική περίπτωση όπου μόνο ένας άνθρωπος χρησιμοποιεί μόνο ένα υπολογιστικό σύστημα.

Μεταγενέστερες προσπάθειες προώθησαν μια πιο ισορροπημένη προσέγγιση, όπου άνθρωποι και ΤΝ αξιοποιούν αμοιβαία τα δημιουργικά τους πλεονεκτήματα για να επιτύχουν περισσότερα. Βασισμένοι στο πλαίσιο του

²²⁶ Εκτενείς και σχετικές βιβλιογραφικές αναφορές περιλαμβάνονται στα τρία άρθρα των Grammenos & Lubart (2025· 2026· in press).

Wiggins για δημιουργικά συστήματα,²²⁷ οι Anna Kantosalo και Hannu Toivonen εισήγαγαν το 2016 την έννοια της «εναλλασσόμενης συν-δημιουργικότητας» (alternating co-creativity), κατά την οποία ένας άνθρωπος και ένας υπολογιστικός πράκτορας εναλλάσσονται στην παραγωγή και τροποποίηση ενός δημιουργικού αντικειμένου ή έννοιας. Πρότειναν, επίσης, τη «συν-δημιουργικότητα με καταμερισμό καθηκόντων» (task-divided co-creativity) ως μια κατάσταση όπου κάθε εταίρος αναλαμβάνει αποκλειστικά συγκεκριμένους ρόλους. Επιπλέον, εντόπισαν και πρότειναν λύσεις σε τέσσερις βασικές προκλήσεις για την επίτευξη αμοιβαία επωφελούς συν-δημιουργίας ανθρώπου-TN, δίνοντας προτεραιότητα στις ανθρώπινες ανάγκες. Οι προκλήσεις αυτές περιλαμβάνουν τις: «καθολική ασυμφωνία» (universal mismatch), όπου ο άνθρωπος ή ο υπολογιστικός πράκτορας παράγει μια έννοια που βρίσκεται εκτός του «σύμπαντος» του επόμενου που πρέπει να δράσει, «εννοιολογική ασυμφωνία» (conceptual mismatch), όπου ο υπολογιστικός πράκτορας αδυνατεί να αναγνωρίσει την έννοια που παρείχε ο άνθρωπος ως έγκυρη ή το αντίστροφο, «καλλιτεχνική διαφωνία» (artistic disagreement), όπου ο άνθρωπος και ο υπολογιστικός πράκτορας διαφωνούν ως προς την αισθητική αξία μιας έννοιας που παρήγαγε ο άλλος, «παραγωγική αδυναμία» (generative impotence), όπου ο άνθρωπος ή ο υπολογιστικός πράκτορας αδυνατεί να συνεχίσει τη δημιουργική αναζήτηση με αφετηρία την έννοια που παρείχε ο άλλος.

Πιο πρόσφατες εργασίες εστιάζουν κυρίως στην Παραγωγική TN και την ενσωμάτωσή της στη δημιουργική διαδικασία, προσφέροντας πρακτική καθοδήγηση και κριτικές επισημάνσεις. Για παράδειγμα, το 2023, οι Tojin Eapen et al. παρέθεσαν μια σειρά από παραδείγματα που δείχνουν πώς η Παραγωγική TN μπορεί να: συσχετίζει απομακρυσμένες έννοιες, αμφισβητεί προκαταλήψεις, βοηθά στην αξιολόγηση και ανάλυση ιδεών, υποστηρίζει τη βελτίωση και τον συνδυασμό ιδεών, διευκολύνει τη συνεργασία μεταξύ των χρηστών.

Το 2024, σε μια μελέτη σχετικά με τη συστηματική χρήση της Παραγωγικής TN σε σύντομους κύκλους σχεδιασμού (design sprints), οι Grange et al. πρότειναν:

²²⁷ Το πλαίσιο προσφέρει έναν τυπικό τρόπο για την περιγραφή, ανάλυση και σύγκριση ανθρώπινων και υπολογιστικών δημιουργικών συστημάτων. Θεωρεί ότι υπάρχει ένα σύμπαν με όλες τις δυνατές έννοιες, μέσα στο οποίο βρίσκεται ένας μικρότερος εννοιολογικός χώρος με τις έννοιες που είναι σχετικές με έναν δημιουργικό τομέα. Το σύστημα λειτουργεί βάσει τριών συνόλων κανόνων, που αντίστοιχα ορίζουν τι ανήκει στον χώρο, τις τεχνικές αναζήτησης νέων ιδεών και πώς κρίνεται το πόσο καλή είναι μια ιδέα. Το πλαίσιο περιγράφει δύο προσεγγίσεις της δημιουργικότητας: τη «διερευνητική» (exploratory), η οποία χρησιμοποιεί προϋπάρχουσες τεχνικές για την εύρεση νέων ιδεών που ταιριάζουν στους κανόνες, και τη «μετασχηματιστική» (transformational), κατά την οποία ο δημιουργός αλλάζει τους κανόνες προκειμένου να προσεγγίσει ιδέες που προηγουμένως ήταν απρόσιτες. Βλ. Wiggins, G. A. (2006). A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems*, 19(7), 449-458. <https://doi.org/10.1016/j.knosys.2006.04.009>

- α) δημιουργία ενός συνεχούς κύκλου συνεργασίας με την Παραγωγική ΤΝ που παράγει προστιθέμενη αξία,
- β) οικοδόμηση εμπιστοσύνης προς την Παραγωγική ΤΝ,
- γ) διαμόρφωση δομημένων διαδικασιών συλλογής και ποιοτικής διαλογής δεδομένων,
- δ) καλλιέργεια μιας νοοτροπίας δεξιotechνίας (craftsmanship mindset).

Αντιθέτως, οι Marrone et al. (2024) υιοθέτησαν μια πιο κριτική οπτική απέναντι στην Παραγωγική ΤΝ, υποστηρίζοντας ότι:

- α) δεν μπορεί αυτόνομα να επιδείξει ουσιαστικά επίπεδα δημιουργικότητας,
- β) μπορεί να συνεργαστεί με ανθρώπους μόνο σε συγκεκριμένες επιμέρους εργασίες της δημιουργικής διαδικασίας,
- γ) μπορεί να συμβάλει στην επιτάχυνση της δημιουργικής διαδικασίας των ανθρώπων (αφήνοντας τις σύνθετες εργασίες σε αυτούς),
- δ) μπορεί να παρέχει γρήγορη και ακριβή ανατροφοδότηση σχετικά με την ανθρώπινη δημιουργικότητα.

Πέρα όμως από τα ενδεχόμενα εντυπωσιακά δημιουργικά οφέλη, η υπερβολική και άκριτη χρήση της ΤΝ ενδέχεται να έχει σημαντικό αρνητικό αντίκτυπο στην ανθρώπινη δημιουργικότητα. Για παράδειγμα, η «γνωστική ανάθεση» (cognitive offloading) δημιουργικών εργασιών στην ΤΝ μπορεί σταδιακά να οδηγήσει σε απώλεια θεμελιωδών δημιουργικών δεξιοτήτων. Επίσης, εάν η διαδικασία του ιδεασμού (ideation) βασίζεται αποκλειστικά σε δεδομένα εκπαίδευσης της ΤΝ αντί σε προσωπικές εμπειρίες και σκέψεις, αναπόφευκτα θα μειωθεί δραστικά και η αυθεντικότητα των δημιουργικών αποτελεσμάτων.

Επιπλέον, η εξάρτηση από ένα κοινόχρηστο σύνολο εργαλείων ΤΝ και παραγόμενης από αυτά γνώσης θα έχει σημαντική αρνητική επίδραση στην «ιδεοποικιλότητα», δηλαδή στην ποικιλομορφία και διαφοροποίηση των ιδεών. Έχει ήδη αποδειχθεί πειραματικά ότι τα δημιουργικά αποτελέσματα των εργαλείων ΤΝ παρουσιάζουν σημαντικά μεγαλύτερη ομοιότητα μεταξύ τους σε σύγκριση με αυτά των ανθρώπων, με αποτέλεσμα τη δημιουργία ενός περιορισμένου συνόλου ιδεών, το οποίο κατά συνέπεια περιορίζει και τις δυνατότητες αποκλίνουσας σκέψης.

Μάλιστα, υπάρχουν επαναλαμβανόμενες ενδείξεις ότι τα αποτελέσματα των συστημάτων Παραγωγικής ΤΝ είναι εξαιρετικά ομοιογενή, ακόμα και

όταν αυτά συνεργάζονται με ανθρώπους. Σε αυτό το πλαίσιο, έχει επισημανθεί ο ενδεχόμενος κίνδυνος να οδηγήσει η υπερβολική εξάρτηση από περιεχόμενο που παράγεται από ΤΝ σε μια γενικευμένη «κατάρρευση της γνώσης» (knowledge collapse), δηλαδή σε μια προοδευτική συρρίκνωση του συνόλου της ανθρώπινης λειτουργικής γνώσης, όπου σπάνιες ή λιγότερο διαδεδομένες γνώσεις σταδιακά εξαφανίζονται για πάντα.

Τα προαναφερθέντα φαινόμενα σύγκλισης των ιδεών μπορούν να αντισταθμιστούν μέσω της ενίσχυσης της ανθρώπινης «δημιουργικής αυτενέργειας» (creative agency), δηλαδή της ικανότητας και της συνειδητής επιλογής του ατόμου να δρα αυτόνομα, να αναλαμβάνει πρωτοβουλία και έλεγχο και να φέρει την ευθύνη για τη δημιουργική διαδικασία και τα αποτελέσματά της. Ωστόσο, ελλοχεύει ο κίνδυνος οι εντυπωσιακές δυνατότητες της ΤΝ να αναστέλλουν τη δημιουργική αυτενέργεια, υπονομεύοντας το κίνητρο και την προθυμία των ατόμων να εμπλακούν σε δημιουργικές προκλήσεις με υψηλές γνωστικές απαιτήσεις. Επιπρόσθετα, είναι πολύ πιθανό να καταστέλλουν τη δημιουργική σκέψη και τη δημιουργική αυτοπεποίθηση των αρχάριων δημιουργών. Οι Ryan Watkins και Eran Barak-Medina (2024) έχουν προτείνει ένα πλαίσιο που περιλαμβάνει τρεις αντιληπτικούς παράγοντες, οι οποίοι διαμορφώνουν την επίδραση της ΤΝ στη δημιουργική αυτενέργεια:

α) το κατά πόσον η ΤΝ εκλαμβάνεται ότι λειτουργεί ανταγωνιστικά ή συμπληρωματικά με ανθρώπινες δεξιότητες,

β) την αντιλαμβανόμενη αποτελεσματικότητα και απόδοση της ΤΝ,

γ) το κατά πόσον τα συστήματα ΤΝ επιτελούν μια λειτουργία υψηλού ή χαμηλού ρίσκου.

Διαφορετικοί συνδυασμοί επιμέρους χαρακτηριστικών αυτών των παραγόντων ενδέχεται να επηρεάσουν τη δημιουργική αυτενέργεια θετικά, αρνητικά ή και ουδέτερα, ενώ ο βαθμός γραμματισμού στην ΤΝ λειτουργεί ως ρυθμιστικός παράγοντας ως προς το πρόσημο της επίδρασης αυτής.

Ένας ακόμα παράγοντας που συμβάλλει ουσιαστικά στη δημιουργική αυτενέργεια είναι οι «δημιουργικές πεποιθήσεις για τον εαυτό» (creative self-beliefs) ενός ατόμου, οι οποίες περιλαμβάνουν τη δημιουργική αυτοπεποίθηση (creative confidence), τη δημιουργική κεντρικότητα (creative centrality), τη δημιουργική ανάληψη ρίσκου (creative risk-taking) και τη δημιουργική αυτορρύθμιση (creative self-regulation), καθώς λειτουργούν ως γέφυρα μεταξύ του «δημιουργικού δυναμικού» (creative potential) (τι θα μπορούσε εν δυνάμει να δημιουργήσει το άτομο) και της «δημιουργικής δράσης» (creative action) (εάν και τι δημιουργεί τελικά). Ωστόσο, οι Faiella et al., στην πρόσφατη έρευνά τους (2025), καταλήγουν στο

συμπέρασμα ότι ακόμη και ισχυρές «δημιουργικές πεποιθήσεις για τον εαυτό» δεν εγγυώνται ότι το άτομο θα διαθέτει και το αντίστοιχο επίπεδο δημιουργικής αυτοπεποίθησης κατά τη συνεργασία του με την ΤΝ.

Οι «δημιουργικές πεποιθήσεις για τον εαυτό» αποτελούν επιπλέον βασικό συστατικό της «δημιουργικής μεταγνώσης» (creative metacognition), δηλαδή της ικανότητας ενός ατόμου να έχει επίγνωση της δημιουργικής του σκέψης και να μπορεί να την προσαρμόζει ανάλογα με τον εκάστοτε στόχο ή πρόβλημα. Σχετικά ερευνητικά ευρήματα υποδεικνύουν ότι η αποτελεσματικότητα υποστήριξης της δημιουργικότητας από την Παραγωγική ΤΝ διαμεσολαβείται από το επίπεδο των μεταγνωστικών στρατηγικών που εφαρμόζονται. Σύμφωνα, δε, με τον καθηγητή του MIT, Jackson G. Lu: «Η μεταγνώση –το να σκέφτεσαι σχετικά με τον τρόπο που σκέφτεσαι– είναι ο συνδετικός κρίκος που λείπει ανάμεσα στην απλοϊκή και στην ουσιαστική χρήση της ΤΝ. Επιτρέπει στους ανθρώπους να θέτουν καλύτερα ερωτήματα, να αναγνωρίζουν κενά γνώσης και να αντλούν πραγματική αξία από τα εργαλεία ΤΝ».

Το 2023, μια πολυπληθής ομάδα διακεκριμένων ερευνητών της δημιουργικότητας δημοσίευσε ένα μανιφέστο,²²⁸ που διερευνά ποικίλα μελλοντικά σενάρια συνεργασίας ανάμεσα σε ανθρώπους και μηχανές. Μεταξύ αυτών, ως βέλτιστο θεωρείται η «επαυξημένη δημιουργικότητα» (augmented creativity), στο πλαίσιο της οποίας επιτυγχάνεται μια υβριδική σύζευξη που δεν θα ήταν εφικτή ούτε από τους ανθρώπους ούτε από την ΤΝ μεμονωμένα. Σε αυτό το σενάριο, το τελικό αποτέλεσμα αποτελεί προϊόν συνεργατικής προσπάθειας, στην οποία και τα δύο μέρη συμμετέχουν λίγο-πολύ ισότιμα, με σαφή αναγνώριση της συνεισφοράς του καθενός.

Για την επίτευξη αυτού του θετικού αποτελέσματος και την ενίσχυση της ηθικής και υπεύθυνης χρήσης της ΤΝ, το μανιφέστο προτείνει τέσσερις θεμελιώδεις «νόμους», οι οποίοι ορίζουν ότι η ΤΝ οφείλει:

α) να μη διαπράττει λογοκλοπή σε βάρος ανθρώπινων έργων και να γνωστοποιεί τις πηγές εκπαίδευσής της τόσο στους χρήστες όσο και στους αρχικούς δημιουργούς,

β) να μην παράγει επιβλαβές περιεχόμενο ή «κακόβουλη δημιουργικότητα» (malevolent creativity),

γ) να υπηρετεί ή να συνεργάζεται με τους ανθρώπους και όχι να τους ανταγωνίζεται,

²²⁸ Vinchon, F., Lubart, T. I., Bartolotta, S., Gironnay, V., Botella, M., Bourgeois-Bougrine, S., Burkhardt, J., Bonnardel, N., Corazza, G. E., Glăveanu, V., Hanchett Hanson, M., Ivcevic, Z., Karwowski, M., Kaufman, J. C., Okada, T., Reiter-Palmon, R., & Gaggioli, A. (2023). Artificial Intelligence & Creativity: A Manifesto for Collaboration. *The Journal of Creative Behavior*, 57(4), 472-484. <https://doi.org/10.1002/jocb.597>

δ) να δηλώνει ρητά πότε το περιεχόμενο έχει παραχθεί τεχνητά.

Για την υποστήριξη της εφαρμογής των νόμων αυτών, προτείνεται ένα σύστημα πιστοποίησης με τρεις ετικέτες ποιότητας, οι οποίες απονέμονται για:

- α) ηθική ανάπτυξη της ΤΝ,
- β) κατάλληλες συνθήκες συνεργασίας ανθρώπου-ΤΝ και
- γ) έργα που έχουν δημιουργηθεί με τη συνδρομή ΤΝ.

Οι τέσσερις «νόμοι» που διατυπώνονται στο μανιφέστο αντικατοπτρίζονται, επίσης, στην πρόσφατη Πράξη για την Τεχνητή Νοημοσύνη,²²⁹ μέσω των απαιτήσεων του άρθρου 53 για τα πνευματικά δικαιώματα και τη γνωστοποίηση των δεδομένων εκπαίδευσης (Πρώτος Νόμος), των απαγορεύσεων και υποχρεώσεων διαχείρισης κινδύνου των άρθρων 5 και 9 (Δεύτερος Νόμος), της πρόβλεψης του άρθρου 14 για ανθρώπινη εποπτεία (Τρίτος Νόμος) και των υποχρεώσεων διαφάνειας και σήμανσης του συνθετικού περιεχομένου του άρθρου 50 (Τέταρτος Νόμος).

Το 2025, μια, επίσης, μεγάλη ομάδα διακεκριμένων ερευνητών συνέταξε ένα άρθρο για την κυβερνο-δημιουργικότητα,²³⁰ με στόχο τη μεγιστοποίηση των οφελών και την ταυτόχρονη ελαχιστοποίηση των κινδύνων της εφαρμογής της ΤΝ στη δημιουργική διαδικασία. Το άρθρο παρουσιάζει δύο αντίθετα πιθανά μελλοντικά σενάρια. Ένα ουτοπικό όραμα, στο οποίο η ΤΝ λειτουργεί ενισχυτικά και όχι υποκαθιστώντας την ανθρώπινη δημιουργικότητα, και μια δυστοπική εκδοχή, η οποία προβλέπει την κατάλυση κάθε μορφής ανθρώπινης δημιουργικής αυτενέργειας λόγω της ανόδου της ΤΝ. Προκειμένου η επιστημονική κοινότητα να βοηθήσει την ανθρωπότητα να κινηθεί προς το ουτοπικό σενάριο, αποφεύγοντας παράλληλα το δυστοπικό, οι ερευνητές καταλήγουν σε έναν δεκάλογο ερευνητικών προκλήσεων, ο οποίος καλύπτει τις ακόλουθες διαστάσεις:

- α) θεωρητικό πλαίσιο της κυβερνο-δημιουργικότητας,
- β) κοινωνικο-πολιτισμικές προοπτικές,

²²⁹ European Parliament & Council of the European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. Official Journal of the European Union. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>

²³⁰ Corazza, G. E., Agnoli, S., Jorge Artigau, A. J., Beghetto, R. A., Bonnardel, N., Coletto, I., Faiella, A., Gerardini, K., Gilhooly, K., Glăveanu, V. P., Hanson, M. H., Kapoor, H., Kaufman, J. C., Kenett, Y. N., Kharkhurin, A. V., Luchini, S., Mangion, M., Mirabile, M., Obialo, F. K., ... Lubart, T. (2025). Cyber-Creativity: A Decalogue of Research Challenges. *Journal of Intelligence*, 13(8), 103. <https://doi.org/10.3390/jintelligence13080103>

- γ) κυβερνο-δημιουργική διαδικασία,
- δ) δημιουργικός δρων (creative agent),
- ε) συν-δημιουργική ομάδα,
- στ) κυβερνο-δημιουργικά προϊόντα,
- ζ) πεδία της κυβερνο-δημιουργικότητας,
- η) εκπαίδευση στην κυβερνο-δημιουργικότητα,
- θ) ηθικές πτυχές,
- ι) σκοτεινή πλευρά της κυβερνο-δημιουργικότητας.

Στην ιστορική διαδρομή της ανθρωπότητας, η δημιουργικότητα υπήρξε αυτονόητα ανθρώπινη και άρρηκτα συνδεδεμένη με την εξέλιξη του είδους και του πολιτισμού. Ωστόσο, στο άμεσο μέλλον, ο αναδυόμενος χώρος της κυβερνο-δημιουργικότητας αναμένεται να επαναπροσδιορίσει τόσο την έννοια όσο και την ίδια τη φύση του «δημιουργού».

Προκειμένου να διασφαλιστεί η ανθρώπινη δημιουργική αυτενέργεια, καθώς και η ουσιαστική παρουσία σε ένα ολοένα και πιο μηχανοκρατούμενο περιβάλλον, ο ρόλος της εκπαίδευσης καθίσταται καθοριστικός. Όχι όμως αποκλειστικά υπό τη μορφή του απαραίτητου ψηφιακού γραμματισμού γύρω από τις δυνατότητες και τους κινδύνους της εκάστοτε τεχνολογίας, αλλά κυρίως μέσω της συστηματικής καλλιέργειας θεμελιωδών ανθρώπινων γνωρισμάτων, όπως η φαντασία, η αυθεντικότητα, η πρόθεση, η ηθική κρίση, η ενσυναίσθηση, η ανάληψη ευθύνης, η κριτική και η δημιουργική σκέψη.

Σε έναν κόσμο όπου οι δημιουργικές ικανότητες των μηχανών εξελίσσονται ραγδαία, το εναπομένον ανταγωνιστικό πλεονέκτημα των ανθρώπων έγκειται στο να γίνουν «περισσότερο άνθρωποι», επενδύοντας σε στοιχεία που η τεχνολογία ενδεχομένως μπορεί να μιμηθεί, αλλά σίγουρα όχι να βιώσει.

Βιβλιογραφία

- Atkinson, D. P., & Barker, D. R. (2023). AI and the social construction of creativity. *Convergence: The International Journal of Research into New Media Technologies*, 29(4), 1054-1069. <https://doi.org/10.1177/13548565231187730>
- Boden, M. A. (1998). Creativity and artificial intelligence. *Artificial Intelligence*, 103(1-2), 347–356. [https://doi.org/10.1016/S0004-3702\(98\)00055-1](https://doi.org/10.1016/S0004-3702(98)00055-1)
- Chatterji, A., Cunningham, T., Deming, D., Hitzig, Z., Ong, C., Shan, C. Y., & Wadman, K. (2025). *How People Use ChatGPT* [NBER Working Paper 34255]. National Bureau of Economic Research. <https://doi.org/10.3386/w34255>
- Corazza, G. E., Agnoli, S., Jorge Artigau, A. J., Beghetto, R. A., Bonnardel, N., Coletto, I., Faiella, A., Gerardini, K., Gilhooly, K., Glăveanu, V. P., Hanson, M. H., Kapoor, H., Kaufman, J. C., Kenett, Y. N., Kharkhurin, A. V., Luchini, S., Mangion, M., Mirabile, M., Obialo, F. K., ... Lubart, T. (2025). Cyber-Creativity: A Decalogue of Research Challenges. *Journal of Intelligence*, 13(8), 103. <https://doi.org/10.3390/jintelligence13080103>
- Eapen, T. T., Finkenstadt, D. J., Folk, J., & Venkataswamy, L. (2023, July). How Generative AI Can Augment Human Creativity. *Harvard Business Review*. <https://hbr.org/2023/07/how-generative-ai-can-augment-human-creativity>
- European Parliament, & Council of the European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. Official Journal of the European Union. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- Faiella, A., Zielińska, A., Karwowski, M., & Corazza, G. E. (2025). Am I Still Creative? The Effect of Artificial Intelligence on Creative Self-Beliefs. *The Journal of Creative Behavior*, 59(2). <https://doi.org/10.1002/jocb.70011>

- Grammenos, D., & Lubart, T. (2025). Peeping at creAltivity through a keyhole: Creative self-perceptions, potential, and enhancement of GenAI chatbots. *Artificial Intelligence Review*, 58(10), 293. <https://doi.org/10.1007/s10462-025-11288-6>
- Grammenos, D., & Lubart, T. (2026). Are Eight Chatbots Better Than One? Boosting Chatbot Creative Outcomes via Exposure to Self- and Peer-Generated Examples. Preprint. <https://doi.org/10.21203/rs.3.rs-8718021/v1>
- Grammenos, D., & Lubart, T. (In Press). Using a componential approach for developing AI-enhanced creativity training interventions. In R. A. Beghetto (Ed.), *Oxford Handbook of Human Creativity x Generative AI in Education*. Oxford University Press.
- Grange, C., Demazure, T., Ringeval, M., Bourdeau, S., & Martineau, C. (2024). The Human-GenAI value loop in human-centered innovation: Beyond the magical narrative. *arXiv*. <https://doi.org/10.48550/arXiv.2407.17495>
- Gruner, D. T., & Csikszentmihalyi, M. (2019). Engineering Creativity in an Age of Artificial Intelligence. In I. Lebudă & V. P. Glăveanu (Eds.), *The Palgrave Handbook of Social Creativity Research* (pp. 447-462). Springer International Publishing. https://doi.org/10.1007/978-3-319-95498-1_27
- Jackson, A.-L. (2024, January 24). *Teaching creativity and critical-thinking skills is imperative as AI transforms education*. Fast Company. <https://www.fastcompany.com/91014583/ai-education-impact-creativity-critical-thinking-skills-davos-2024>
- Kantosalo, A., & Toivonen, H. (2016). Modes for creative human-computer collaboration: Alternating and task-divided co-creativity. In *Proceedings of the International Conference on Innovative Computing and Cloud Computing* (pp. 77-84).
- Maher, M. L. (2012). Computational and collective creativity: Who's being creative? In *Proceedings of the Third International Conference on Computational Creativity* (pp. 67-71).
- Marrone, R., Crompton, D., & Medeiros, K. (2024). How Does Narrow AI Impact Human Creativity? *Creativity Research Journal*, 38(1), 150-160. <https://doi.org/10.1080/10400419.2024.2378264>
- MIT Sloan Office of Communications (2025). *Does generative AI actually enhance creativity in the workplace?* MIT Sloan School of Management.

<https://mitsloan.mit.edu/press/does-generative-ai-actually-enhance-creativity-workplace>

Nakakoji, K. (2006). Meanings of tools, support, and uses for creative design processes. In *Proceedings of the International Design Research Symposium '06* (pp. 156-165).

Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal*, 24(1), 92-96. <https://doi.org/10.1080/10400419.2012.650092>

Vinchon, F., Lubart, T. I., Bartolotta, S., Gironnay, V., Botella, M., Bourgeois-Bougrine, S., Burkhardt, J., Bonnardel, N., Corazza, G. E., Glăveanu, V., Hanchett Hanson, M., Ivcevic, Z., Karwowski, M., Kaufman, J. C., Okada, T., Reiter-Palmon, R., & Gaggioli, A. (2023). Artificial Intelligence & Creativity: A Manifesto for Collaboration. *The Journal of Creative Behavior*, 57(4), 472-484. <https://doi.org/10.1002/jocb.597>

Watkins, R., & Barak-Medina, E. (2024). AI's Influence on Human Creative Agency. *Creativity Research Journal*, 38(1), 137-149. <https://doi.org/10.1080/10400419.2024.2437264>

10.2. Τεχνητή Νοημοσύνη και Ζητήματα πνευματικής ιδιοκτησίας

Μια άλλη διάσταση που έχει απασχολήσει πολύ τη δημόσια και νομική συζήτηση είναι η χρήση έργων που προστατεύονται από δικαιώματα πνευματικής ιδιοκτησίας ως «δεδομένων εκπαίδευσης» της Παραγωγικής Τεχνητής Νοημοσύνης. Μια εκπαίδευση που επιτρέπει περαιτέρω τη «δημιουργία» νέου ψηφιακού περιεχομένου με ίδια ή παρόμοια μορφολογικά χαρακτηριστικά. Αναπόφευκτα τίθεται το ερώτημα της επάρκειας των ρυθμίσεων για την αντιμετώπιση αυτών των νέων ζητημάτων.

Ακόμη και εάν η «παραδοσιακή» νομική μας προσέγγιση εξακολουθεί να αρνείται (και ορθά) την ιδιότητα του δημιουργού σε ένα –μη ανθρώπινο– σύστημα ΤΝ ή να μην αναγνωρίζει ως πρωτότυπη δημιουργία το έργο που παρήχθη ως αποτέλεσμα prompts, τα ζητήματα της χρήσης των έργων για την εκπαίδευση των συστημάτων απασχολούν με ιδιαίτερη ένταση: Η ποιότητα των «απαντήσεων», των «στοιχείων εξόδου», εξαρτάται από τη διαθεσιμότητα πληροφορίας για την εκπαίδευση των μοντέλων. Η ανάπτυξη, η δοκιμή και η επικύρωση μοντέλων ΤΝ βασίζονται σε μεγάλο βαθμό στην πρόσβαση σε μεγάλα σύνολα δεδομένων, γεγονός που έχει οδηγήσει σε εκρηκτική αύξηση της ζήτησης για δεδομένα εκπαίδευσης.

Μια ευρέως χρησιμοποιούμενη μέθοδος συλλογής τέτοιων δεδομένων είναι η «συλλογή δεδομένων μέσω εξόρυξης» (data scraping), η οποία αναφέρεται στην αυτοματοποιημένη εξαγωγή πληροφοριών από ιστότοπους, βάσεις δεδομένων ή πλατφόρμες κοινωνικής δικτύωσης τρίτων. Οι δραστηριότητες scraping μπορούν να εγείρουν ζητήματα που αφορούν δικαιώματα διανοητικής ιδιοκτησίας, στα οποία συγκαταλέγονται δικαιώματα πνευματικής ιδιοκτησίας, τα δικαιώματα επί βάσεων δεδομένων, τα εμπορικά σήματα, τα εμπορικά απόρρητα.²³¹

Υπάρχει μια υποχρέωση «συνεισφοράς» στην προαγωγή της γενικής γνώσης; Είναι θεμιτή χρήση (fair use) η χρήση των έργων χωρίς την άδεια των δημιουργών τους; Η συζήτηση και οι αντίστοιχες δικαστικές διαμάχες είναι ενεργές, ιδίως στις ΗΠΑ.

²³¹ Organisation for Economic Co-operation and Development (OECD). (2025). *Intellectual property issues in artificial intelligence trained on scraped data* (OECD Artificial Intelligence Papers, No. 33). <https://doi.org/10.1787/d5241a23-en>

Η πρακτική του scraping επηρεάζει άμεσα τους δημιουργούς και τους κατόχους έργων που προστατεύονται από δικαιώματα διανοητικής ιδιοκτησίας, ιδίως όταν πραγματοποιείται χωρίς τη συναίνεση ή την καταβολή αμοιβής στους δικαιούχους. Το μείζον ζήτημα έγκειται στην αναζήτηση μιας (νέας) ισορροπίας μεταξύ, αφενός, του αιτήματος για πρόσβαση σε προστατευόμενα έργα, χάριν της ανάπτυξης της Παραγωγικής ΤΝ και της καινοτομίας εν γένει εντός της ΕΕ, και, αφετέρου, των δικαιούχων πνευματικών δικαιωμάτων για έλεγχο της πρόσβασης, αλλά και ανταποδοτικότητα της πρόσβασης στα έργα αυτά. Σε τελευταία ανάλυση, τίθεται ζήτημα μιας δίκαιης κατανομής του οφέλους που προκύπτει από τον εμπλουτισμό των μοντέλων.

Είναι μόνο ζήτημα επιβεβαίωσης ή τροποποίησης των ισχυουσών ρυθμίσεων; Όπως αναφέρεται και σε μελέτη του Ευρωπαϊκού Κοινοβουλίου, η ΤΝ (AI), και ιδίως η Παραγωγική ΤΝ (Generative AI – GenAI), έχει μειώσει περαιτέρω το κόστος παραγωγής και έχει καταστήσει δυνατή μια έκρηξη περιεχομένου. Και όπως έχει ήδη καταδείξει η γενικότερη εμπειρία, το ζήτημα δεν είναι μόνο νομικό, αλλά αφορά και το οικονομικό μοντέλο της Παραγωγικής ΤΝ και τα οφέλη που (προσδοκάται να) επιφέρει για την οικονομία και την κοινωνία.

Η «ευρωπαϊκή λύση» φαίνεται να εστιάζει στους όρους διαφάνειας και συμμόρφωσης. Η προσέγγιση αυτή αποτυπώνεται ήδη στην AI Act με την υποχρέωση υιοθέτησης πολιτικής συμμόρφωσης με τους κανόνες της πνευματικής ιδιοκτησίας, αλλά και την υποχρέωση διαφάνειας και σήμανσης του περιεχομένου και των έργων που υποστηρίζονται ή παράγονται με τη χρήση ΤΝ. Αντίστοιχες απαιτήσεις περιέχονται και στον Κώδικα Ορθής Πρακτικής για την ΤΝ Γενικού Σκοπού. Αξιοσημείωτη εξέλιξη αποτελεί το πρόσφατο (10 Μαρτίου 2026) ψήφισμα του Ευρωπαϊκού Κοινοβουλίου, που εστιάζει στην πληρέστερη διαφάνεια ως προς τα προστατευόμενα έργα που χρησιμοποιούνται στην εκπαίδευση, στην ενίσχυση της δυνατότητας opt-out των δικαιούχων, δηλαδή της δυνατότητας να επιλέγουν την εξαίρεση των έργων τους από τη χρήση τους για εκπαίδευση και την ανάπτυξη μηχανισμών αμοιβής ή αδειοδότησης.

Το ψήφισμα του ΕΚ συνοψίζει ακριβώς την πολλαπλή –και πάντως δυσχερώς επιτεύξιμη– στόχευση της ΕΕ: «Η Ευρωπαϊκή Ένωση βρίσκεται αντιμέτωπη με τη στρατηγική πρόκληση της υστέρησης έναντι άλλων διεθνών δυνάμεων στον τομέα της Τεχνητής Νοημοσύνης (ΤΝ). Στο πλαίσιο αυτό, η Ευρωπαϊκή Επιτροπή παρουσίασε, κατά τη Σύνοδο Δράσης για την ΤΝ στο Παρίσι τον Φεβρουάριο του 2025, το σχέδιο δράσης "AI Continent", με σκοπό την ανάδειξη της Ευρώπης σε παγκόσμιο ηγέτη στον συγκεκριμένο τομέα. Το ψήφισμα υπογραμμίζει ότι η επίτευξη του στόχου αυτού προϋποθέτει τη στήριξη και προώθηση των τεχνολογιών και υπηρεσιών Παραγωγικής ΤΝ, αντί της επιβολής αδικαιολόγητων περιορισμών που θα μπορούσαν να ανακόψουν την ανάπτυξή τους. Η ενίσχυση της καινοτομίας στον τομέα της ΤΝ θεωρείται

απαραίτητη για τη διασφάλιση της τεχνολογικής κυριαρχίας της Ευρώπης, την ενδυνάμωση της ανταγωνιστικότητάς της, την προστασία της πολιτιστικής και γλωσσικής πολυμορφίας της και τη διατήρηση της ικανότητάς της να καινοτομεί, πάντοτε με σεβασμό στις θεμελιώδεις ευρωπαϊκές αξίες. Παράλληλα, τονίζεται ότι η τεχνολογική πρόοδος πρέπει να συμβάλλει στη βιώσιμη οικονομική ανάπτυξη, στην παραγωγικότητα και την ανταγωνιστικότητα, ενώ οι τεχνολογίες ΤΝ οφείλουν να είναι προσβάσιμες σε ευρύ φάσμα χρηστών και φορέων σε ολόκληρη την Ένωση. Τέλος, επισημαίνεται ότι η ανάπτυξη προηγμένων συστημάτων Παραγωγικής ΤΝ εξαρτάται σε μεγάλο βαθμό από τη διαθεσιμότητα υψηλής ποιότητας δεδομένων εκπαίδευσης. Για τον λόγο αυτόν, ο δημιουργικός τομέας αναγνωρίζεται ως κρίσιμη πηγή ποιοτικού περιεχομένου και κρίνεται αναγκαία η δίκαιη αμοιβή των δημιουργών, ώστε να διασφαλίζεται τόσο η βιωσιμότητα της πολιτιστικής παραγωγής όσο και η περαιτέρω εξέλιξη των τεχνολογιών ΤΝ» (Ψήφισμα – Σημείο Β).²³²

Το διακύβευμα δεν είναι μόνο ένα ζήτημα πνευματικών δικαιωμάτων, αλλά ένα ακόμη πεδίο άσκησης ψηφιακής κυριαρχίας. Είναι ζήτημα δίκαιης κατανομής των πόρων και των ωφελειών της τεχνολογίας.

²³² Για το σχετικό κείμενο βλ. Ευρωπαϊκό Κοινοβούλιο (2026, 10 Μαρτίου). Προστασία των δικαιωμάτων των δημιουργών της ΕΕ στην εποχή της τεχνητής νοημοσύνης [Δελτίο Τύπου]. <https://www.europarl.europa.eu/news/el/press-room/20260306IPR37511/prostasia-ton-dikaionaton-ton-dimiourgon-tis-ee-stin-epochi-tis-tn>

[ΠΩΣ] ΜΠΟΡΕΙ Η ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΝΑ ΕΙΝΑΙ
ΗΘΙΚΗ ΚΑΙ ΣΥΝΝΟΜΗ;
Ιούλιος 2026

Ο ρόλος και η ευθύνη της Πολιτείας



Αντώνης Κουρουτάκης, Λίλιαν Μήτρου

Συγγραφική συνεισφορά ανά υποκεφάλαιο:

11.1., 11.2., 11.3. – Λίλιαν Μήτρου

11.4. – Αντώνης Κουρουτάκης

11.1. Ποιες είναι οι υποχρεώσεις του κράτους στην εποχή της Τεχνητής Νοημοσύνης;

Η τεχνολογική ανάπτυξη που λαμβάνει χώρα υπό την επίδραση της Τεχνητής Νοημοσύνης είναι άνευ προηγουμένου στην ανθρώπινη ιστορία λόγω της ταχύτητας, του εύρους και του βάθους της. Η ΤΝ, τόσο ως απότοκο όσο και ως κορύφωση μιας ευρύτερης τεχνολογικής διαδικασίας ψηφιοποίησης που έχει μετατρέψει πρωτοφανή όγκο πληροφοριών σε ψηφιακή μορφή, προκαλεί βαθιές έως ανατρεπτικές αλλαγές στις κοινωνίες μας και επιτάσσει τον αναστοχασμό για τον ρόλο του κράτους και τις ευθύνες του έναντι των πολιτών.

Ο νομοθέτης δεν παράγει τεχνολογία. Παραμένει, όμως, υπεύθυνος για την ένταξη των νέων τεχνολογιών στην κοινωνία και για τη διαχείριση της σύνθετης έντασης μεταξύ των οικονομικών και κοινωνικών πλεονεκτημάτων και των κινδύνων που συνδέονται με αυτές. Το κράτος της ψηφιακής εποχής καλείται να αναζητήσει νέες σταθμίσεις και νέες ισορροπίες, αλλά και να αντιμετωπίσει νέες κατηγορίες απειλών και κινδύνων, που –ουσιαστικά– συνδέονται μεταξύ τους.²³³ Πρωτίστως καλείται:

α) να προστατεύσει τους πολίτες έναντι κυβερνοαπειλών και ασύμμετρων απειλών που μπορεί να έχουν συνέπειες τόσο στη δυνατότητα απόλαυσης των δικαιωμάτων τους, αλλά και στην ίδια την ύπαρξή τους,

β) να προστατεύσει τους πολίτες από την έκθεση στην παραπληροφόρηση, που μπορεί να εξελιχθεί σε πρόβλημα δημοκρατίας και κυριαρχίας, αλλά και τη χειραγώγηση με τη στοχευμένη ενημέρωση, που μπορεί να οδηγήσει στη «στένωση» του ορίζοντα και κατ' αποτέλεσμα των απόψεων και επιλογών μας,

γ) να αντιμετωπίσει τον ραγδαίο ρυθμό εξέλιξης του οικοσυστήματος των πλατφορμών και τη δυναμική φύση των εταιρειών αυτών, την ικανότητά τους να υπερβαίνουν τις προκλήσεις που θέτουν τα ζητήματα δικαιοδοσίας, καθώς και τον όγκο και την κλίμακα της πληροφορίας που διακινείται μέσω

²³³ Περισσότερα βλ. σε Μήτρου, Λ. (2024), Σύνταγμα και Νέες Τεχνολογίες, Στο Χ. Ανθόπουλος, Λ. Παπαδοπούλου, Ν. Παπαχρήστος και Κ. Χρυσόγονος (Επιμ.), *Η θεωρία της συνταγματικής πράξης* (σ. 107 επ.). Σάκκουλας.

των υπηρεσιών τους και παράλληλα, να αποκαταστήσει, να αντιμετωπίσει ή έστω να μετριάσει την ασυμμετρία ισχύος που προκύπτει και από το γεγονός ότι οι πλατφόρμες καταλήγουν να εκτελούν ημιδημόσιες λειτουργίες σε ένα παγκοσμιοποιημένο ψηφιακό περιβάλλον,

δ) να διασφαλίσει την προστασία των θεμελιωδών δικαιωμάτων, του κράτους δικαίου και της δημοκρατίας στο εξελισσόμενο οικοσύστημα της αλγοριθμικής διακυβέρνησης μέσω ισχυρών συστημάτων AI, γεγονός που θα μπορούσε να οδηγήσει σε ένα νέο μοντέλο διακυβέρνησης από αλγορίθμους, την αλγοκρατία (algorocracy).

Ο κατάλογος της ευρωπαϊκής διακήρυξης σχετικά με τα ψηφιακά δικαιώματα και τις ψηφιακές αρχές για την ψηφιακή δεκαετία²³⁴ είναι ίσως μια ενδεικτική απαρίθμηση των προκλήσεων που καλείται να αντιμετωπίσει ο νομοθέτης: από τη διακηρυκτικού χαρακτήρα απαίτηση για ψηφιακό μετασχηματισμό με επίκεντρο τους ανθρώπους, μια επιλογή που αντανάκλαται στην ανθρωποκεντρική προσέγγιση της AI Act, στην έμφαση στην αλληλεγγύη και την ένταξη (μέσω και της εκπαίδευσης και κατάρτισης σε ψηφιακές δεξιότητες), αλλά και στη διασφάλιση δίκαιων και πρόσφορων συνθηκών εργασίας.

Η τεχνολογία ή ακριβέστερα οι συνέπειές της επιβάλλουν στον νομοθέτη να λάβει μια σειρά από σύνθετες αποφάσεις: εάν και πότε να παρέμβει, με ποιο είδος κανονιστικής παρέμβασης, ποιοι είναι οι εμπλεκόμενοι, πόσο πρέπει να διαρκέσει η κανονιστική παρέμβαση. Το τελευταίο διάστημα έχει τεθεί προς συζήτηση εάν θα πρέπει να συμπεριληφθεί στο Σύνταγμα νέα διάταξη που θα (επιχειρήσει να) ρυθμίσει τη χρήση της ΤΝ. Χρειάζεται μια νέα συνταγματική διάταξη και, εάν ναι, με ποιο περιεχόμενο; Ή μήπως δεν είναι μόνο ζήτημα μιας συνταγματικής ρύθμισης;

²³⁴ European Parliament, Council of the European Union, & European Commission. (2023). *European declaration on digital rights and principles for the digital decade* (2023/C 23/01; PUB/2023/89). Official Journal of the European Union. [https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32023C0123\(01\)](https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32023C0123(01))

11.2. Τεχνητή Νοημοσύνη και Συνταγματική Αναθεώρηση²³⁵

Η Συνταγματική Αναθεώρηση του 2001, εντάσσοντας στο Σύνταγμα ρυθμίσεις όπως η ρητή κατοχύρωση της προστασίας προσωπικών δεδομένων (άρθρο 9Α) και κυρίως του δικαιώματος συμμετοχής στην Κοινωνία της Πληροφορίας (άρθρο 5Α), κωδικοποιούσε και ενίσχυε ορισμένα δικαιώματα, αλλά ταυτόχρονα συγκροτούσε την απαρχή μιας νέας «πληροφοριακής συνταγματικής τάξης», δημιουργούσε μια βάση ρύθμισης της Κοινωνίας της Πληροφορίας, καθώς οι Τεχνολογίες Πληροφορικής και Επικοινωνιών επιδρούσαν ήδη καθοριστικά στην κοινωνία, στην οικονομία και το κράτος.

Αν και εν πολλοίς αντιμετωπίστηκε ως «συμβολική», η σημαντικότερη ίσως αλλαγή συνίστατο στη δεύτερη παράγραφο του άρθρου 5Α, δηλαδή το δικαίωμα συμμετοχής στην Κοινωνία της Πληροφορίας. Με τη διατύπωση της ρύθμισης να οδηγεί στην κατοχύρωση ενός «συμμετοχικού δικαιώματος», το οποίο αναφέρεται σε όλες τις εκφάνσεις της ζωής και στο σύνολο των δυνατοτήτων και ευκαιριών του καθενός, ώστε να μπορεί να λαμβάνει μέρος σε όλες τις διεργασίες της Κοινωνίας της Πληροφορίας (κοινωνικές, οικονομικές, πολιτικές, πολιτιστικές, εργασιακές, ψυχαγωγικές) και να συνδιαμορφώνει την κοινωνία αυτήν, όντας ενεργό μέλος της και όχι απλώς «χρήστης» (user) των τεχνολογικών μέσων. Πέραν μιας ειδικότερης έκφρασης του δικαιώματος συμμετοχής «στην κοινωνική, οικονομική και πολιτική ζωή της χώρας» που πλέον εξελισσόταν όλο και περισσότερο σε ψηφιακά περιβάλλοντα, η διάταξη ενέχει μία ειδικότερη εκδοχή της αρχής της ισότητας, το κράτος υποχρεούται σε διευκόλυνση της πρόσβασης στις πληροφορίες, καθώς και της παραγωγής, ανταλλαγής και διάδοσής τους, χωρίς αποκλεισμούς. Η συσχέτιση με την παραγωγή, ανταλλαγή και διάδοση πληροφορίας είναι ιδιαίτερης σημασίας, καθώς οι πληροφορίες, τα δεδομένα και η «διακυβέρνησή» τους είναι κομβικής σπουδαιότητας για την κοινωνία στην εποχή της ΤΝ.

Ωστόσο, πλέον υπό την επίδραση της ραγδαίας εξέλιξης και κυρίως διείσδυσης του νέου κύματος τεχνολογίας, της ΤΝ, σε κάθε πτυχή της ζωής,

²³⁵ Εκτενέστερη εκδοχή του κειμένου αυτού έχει δημοσιευτεί ως: Μήτρου, Λ. (2026). Χρειαζόμαστε ένα Σύνταγμα για την Τεχνητή Νοημοσύνη. Στο Ε. Βενιζέλος (Επιμ.), *Πενήντα χρόνια από το Σύνταγμα του 1975* (σ. 675 επ.). Σάκκουλας.

προκύπτουν νέες απαιτήσεις από το κράτος αναφορικά με την εγγύηση της ελευθερίας, των δικαιωμάτων και της ισότιμης συμμετοχής, καθώς αμφισβητείται η ρυθμιστική επάρκεια των κανόνων. Αφορούν αυτές οι νέες προκλήσεις και το συνταγματικό κείμενο; Είναι επαρκές ώστε να δημιουργήσει τις προϋποθέσεις για τη θεσμική υποδοχή του μέλλοντος;

Στο πλαίσιο της συζήτησης για τη Συνταγματική Αναθεώρηση ετέθη μεταξύ άλλων ως στόχος η αναδιατύπωση διατάξεων ή η προσθήκη νέων, ώστε να λαμβάνονται υπ' όψιν οι εξελίξεις και η προσαρμογή στα δεδομένα της εποχής της ΤΝ.²³⁶ Τα προκαταρκτικά ερωτήματα είναι: Θα πρέπει να αναθεωρούμε το Σύνταγμα με κάθε –έστω μείζονος χαρακτήρα– τεχνολογική αλλαγή ή να επιδιώξουμε ένα Σύνταγμα που χαρακτηρίζεται από τεχνολογική ουδετερότητα;²³⁷ Οι χρήσεις της Τεχνητής Νοημοσύνης είναι ρυθμιστέες και ρυθμίσιμες, και ποιος ο ρόλος του Συντάγματος; Επιτρέπει το Σύνταγμα ως έχει τη ρύθμιση και αντιμετώπιση των επιπτώσεων των τεχνολογικών αλλαγών, έτσι ώστε να προσδιορίσει και να περιορίσει την άσκηση εξουσιών, να εδραιώσει και να κατοχυρώσει δικαιώματα, και να επαναπροσδιορίσει τη θέση του ατόμου σε έναν ραγδαία μεταβαλλόμενο κοινωνικο-οικονομικό μετασχηματισμό; Εάν θεωρήσουμε ότι το Σύνταγμα δεν μπορεί να χαρακτηριστεί στο σύνολό του ως ένα Σύνταγμα με προσανατολισμό την Κοινωνία της Πληροφορίας, αυτό να σημαίνει ότι παρεμποδίζεται θεσμικά η μετάβαση αυτή; Επαρκεί το πλαίσιο και πλέγμα που έχει διαμορφώσει κυρίως η ενωσιακή νομοθεσία κατά την τελευταία πενταετία; Ή χρειαζόμαστε και ένα νέο, ψηφιακό, Σύνταγμα; Έχει όμως περιθώρια (και νόημα;) η αποτύπωση θεμελιακών έστω επιλογών στο εθνικό Σύνταγμα, σε ένα περιβάλλον που εξ ορισμού υπερβαίνει τα εθνικά σύνορα και την κανονιστική εμβέλεια του εθνικού νομοθέτη;

Η πρόταση που έχει διατυπωθεί από τον Ευριπίδη Στυλιανίδη²³⁸ εδράζεται στις ακόλουθες επιλογές:

α) στη «δέσμευση» ότι η εφαρμογή της Τεχνητής Νοημοσύνης πρέπει να συμβάλλει στην προστασία των ανθρωπίνων δικαιωμάτων, της δημοκρατίας και του κράτους δικαίου,

²³⁶ Στυλιανίδης, Ε. (2025). Πρόταση για την Τεχνητή Νοημοσύνη εν όψει της Αναθεώρησης του Ελληνικού Συντάγματος. Στο Ε. Στυλιανίδης (Επιμ.), *Τεχνητή Νοημοσύνη-Ανθρώπινα δικαιώματα, Δημοκρατία και Κράτος Δικαίου* (σ. 633 επ.). Νομική Βιβλιοθήκη.

²³⁷ Πρέπει να είναι τεχνολογικά ουδέτερο (technology neutral) ή τεχνολογικά αγνωστικιστικό (technology agnostic), δηλαδή να είναι διατυπωμένο κατά τρόπο ώστε να μην αντιστοιχεί σε μια συγκεκριμένη τεχνολογική φάση, αλλά να μπορεί να αναπτύσσει τη ρυθμιστική ενέργειά του ανεξάρτητα από τη ρυθμιζόμενη τεχνολογία.

²³⁸ Στο κείμενο της 3ης Ιουνίου 2026 «Πρόταση της Κοινοβουλευτικής Ομάδας της ΝΔ για την αναθεώρηση του Συντάγματος», εισηγητής της οποίας είναι ο Ε. Στυλιανίδης, αναφέρεται: «Άρθρο 5 ή άρθρο 5Α ή νέο άρθρο 5Β - Η τεχνητή νοημοσύνη οφείλει να υπηρετεί την ελευθερία του ατόμου και την ευημερία της κοινωνίας ώστε να μετριάζονται οι κίνδυνοι και να αξιοποιούνται τα πλεονεκτήματα που προσφέρει».

β) στη ρητή διατύπωση της υποχρέωσης της Πολιτείας να μετριάξει τους κινδύνους και να αξιοποιεί τα πλεονεκτήματα που προσφέρει η Τεχνητή Νοημοσύνη για την προώθηση μιας καλύτερης ζωής για όλους,

γ) στη διατύπωση των βασικών αρχών που πρέπει να διέπουν την εφαρμογή της Τεχνητής Νοημοσύνης, δηλαδή:

- 1) Διαφάνεια, συμπεριλαμβανομένης της προσβασιμότητας και της επεξηγηματικότητας
- 2) Δικαιοσύνη και ακεραιότητα, συμπεριλαμβανομένης της μη διάκρισης
- 3) Ανθρώπινη ευθύνη για τις αποφάσεις, συμπεριλαμβανομένης της ευθύνης και της διαθεσιμότητας ενδίκων μέσων
- 4) Ασφάλεια και προστασία
- 5) Προστασία της ιδιωτικής ζωής και των προσωπικών δεδομένων

δ) στην πρόβλεψη Ανεξάρτητης Αρχής για τη διασφάλιση της εφαρμογής των ως άνω προτεινόμενων προβλέψεων.

Αν και είναι σημαντικό να γίνουν κατανοητές και να «μεταφραστούν» οι θεμελιώδεις συνταγματικές αρχές και ελευθερίες υπό το πρίσμα των μεταβαλλόμενων τεχνολογικών (και όχι μόνο) συνθήκων, θα πρέπει να αναρωτηθούμε για την προστιθέμενη αξία μιας νέας διάταξης. Η Ευρωπαϊκή Ένωση, χρησιμοποιώντας μάλιστα το κανονιστικό εργαλείο των Κανονισμών, των Acts, καθώς και τον έλεγχο των μεγάλων πλατφορμών, επιχειρεί κατά την τελευταία πενταετία την αντιμετώπιση του παγκόσμιου ιδιωτικού ψηφιακού οικοσυστήματος και των επιπτώσεών του – ακόμη και στη δημοκρατία.

Η υπό διαμόρφωση νέα γενιά δικαιωμάτων αναφέρεται σε τελευταία ανάλυση στα θεμέλια, στις κατευθυντήριες αρχές και τα πρωταρχικά δικαιώματα της σύγχρονης ευρωπαϊκής και εθνικής συνταγματικής δικαιοταξίας, όπως η δημοκρατία και το κράτος δικαίου, η αξία του ανθρώπου, η ισότητα, η ελευθερία ανάπτυξης της προσωπικότητας και συμμετοχής. Οι αρχές και απαιτήσεις για διαφάνεια, επεξηγησιμότητα, λογοδοσία, ανθρώπινη παρέμβαση, δικαιοσύνη και ισότητα, προστασία δεδομένων και επιδίωξη της κοινωνικής ευημερίας έχουν ήδη αποτυπωθεί στις Κατευθυντήριες Γραμμές Δεοντολογίας για αξιόπιστη Τεχνητή Νοημοσύνη που διατύπωσε η Ομάδα Εμπειρογνομόνων Υψηλού Επιπέδου για την Τεχνητή Νοημοσύνη. Ο αντίστοιχος Κανονισμός, δηλαδή η Πράξη για την Τεχνητή Νοημοσύνη, ορίζει στο πρώτο άρθρο ως σκοπό την προώθηση της υιοθέτησης ανθρωποκεντρικής και αξιόπιστης ΤΝ, διασφαλίζοντας παράλληλα υψηλό επίπεδο προστασίας της Υγείας, της Ασφάλειας, των θεμελιωδών δικαιωμάτων που

κατοχυρώνονται στον Χάρτη, συμπεριλαμβανομένης της δημοκρατίας, του κράτους δικαίου και της περιβαλλοντικής προστασίας από τις επιζήμιες συνέπειες των συστημάτων ΤΝ στην Ένωση, και στηρίζοντας την καινοτομία. Αναφορικά με την πρόταση για Ανεξάρτητη Αρχή, η Πράξη για την Τεχνητή Νοημοσύνη περιέχει ήδη σχετική πρόβλεψη, ενώ στην ελληνική έννομη τάξη λειτουργούν ήδη Ανεξάρτητες Αρχές, όπως η Αρχή Προστασίας Δεδομένων Προσωπικού Χαρακτήρα, επιφορτισμένες με την εποπτεία της προστασίας δικαιωμάτων και την τήρηση εγγυήσεων.

Ένα νέο συνταγματικό άρθρο θα επιβεβαίωνε, έστω και συμβολικά, την εδραίωση και επιβεβαίωση των βασικών συνταγματικών επιλογών με διαφορετικές, ενισχυμένες διατυπώσεις, προσαρμοσμένες στις νέες τεχνολογίες και τους κινδύνους που συνεπιφέρουν; Οι θετικές υποχρεώσεις του κράτους αφορούν την εγγύηση των ατομικών δικαιωμάτων και ελευθεριών, και την υποστήριξη της ακώλυτης άσκησής τους στο περιβάλλον της ΤΝ, η οποία καταλαμβάνει πλέον κάθε πτυχή της ανθρώπινης δραστηριότητας και για αυτό θα ήταν ανακόλουθο να αντιμετωπιστεί ως ένα φαινόμενο που απλώς συμπληρώνει τα προϋφιστάμενα. Δύο κρίσιμες παράμετροι για την απόφαση επί της αναγκαιότητας και του περιεχομένου μιας συνταγματικής ρύθμισης είναι α) η έλλειψη προβλεψιμότητας αναφορικά με τις τεχνολογικές εξελίξεις, το εύρος και το βάθος των επιπτώσεών τους, και β) ο διεθνής, πολυκεντρικός και εξελικτικός χαρακτήρας του κυβερνοχώρου και της Τεχνητής Νοημοσύνης.

Η συνταγματική «απάντηση» στις απαιτήσεις μιας ψηφιακής κοινωνίας της διακινδύνευσης δεν μπορεί παρά να αναζητηθεί σε επίπεδο που υπερβαίνει το μεμονωμένο κράτος.²³⁹ Η θέση αυτή ενισχύεται από τη διαπίστωση ότι το μείζον διακύβευμα είναι ο δημόσιος έλεγχος των παγκόσμιας εμβέλειας και επιρροής ιδιωτών φορέων, οι δραστηριότητες και πρακτικές των οποίων επηρεάζουν άμεσα την επικοινωνία και τον δημόσιο λόγο, δημόσιες και ιδιωτικές σχέσεις, και ενέχουν σοβαρούς κινδύνους για τα δικαιώματα και τη δημοκρατία. Συνεπώς, εφόσον τεθεί ζήτημα στη συζήτηση για τη Συνταγματική Αναθεώρηση, η προτεινόμενη αλλαγή θα μπορούσε να περιοριστεί σε μια λιτή αναδιατύπωση του άρθρου 5Α παρ. 2²⁴⁰ ή στην προσθήκη μιας –πάντα λιτής– τρίτης παραγράφου κατά τρόπο ώστε να προσαρμοστεί στις ειδικότερες απαιτήσεις της ΤΝ.

²³⁹ Βενιζέλος, Ε. (2020). Σύνταγμα και νέες τεχνολογίες. *Δίκαιο Επιχειρήσεων & Εταιριών (ΔΕΕ)*, 4, 542-548.

²⁴⁰ Η διατύπωση ως έχει σήμερα είναι «καθένας έχει δικαίωμα συμμετοχής στην Κοινωνία της Πληροφορίας. Η διευκόλυνση της πρόσβασης στις πληροφορίες που διακινούνται ηλεκτρονικά, καθώς και της παραγωγής, ανταλλαγής και διάδοσής τους αποτελεί υποχρέωση του κράτους, τηρουμένων πάντοτε των εγγυήσεων των άρθρων 9, 9Α και 19».

11.3. Ψηφιακός συνταγματισμός ή/και ένα νέο «κοινωνικό συμβόλαιο»;

Ανεξάρτητα από τη συζήτηση στο πλαίσιο της εξαγγελθείσας Αναθεώρησης, πρέπει να συζητήσουμε εν γένει την ανάγκη για έναν νέο ψηφιακό συνταγματισμό, προκειμένου να «μεταφραστούν» οι θεμελιώδεις συνταγματικές αρχές και ελευθερίες υπό το πρίσμα των μεταβαλλόμενων τεχνολογικών (και όχι μόνο) συνθηκών,²⁴¹ έτσι ώστε να προστατεύονται τα δικαιώματα αλλά και να περιορίζεται η ανάπτυξη ιδιωτικών εξουσιών εκτός δημοσίου ελέγχου.²⁴² Όπως παρατηρεί ο Πέτρος Τερζής, ο ψηφιακός συνταγματισμός απαντά σε έναν φόβο, τον φόβο ότι οι συνταγματικές αρχές και προστασίες καταρρέουν γρήγορα υπό τις αδιόρατες δυνάμεις νέων ψηφιακών τεχνολογιών και των ιδιωτικών οντοτήτων που ασκούν τεχνολογική ισχύ. Και διαισθητικά, το εγχείρημα φαίνεται να ευσταθεί.²⁴³

Όπως και σε άλλες περιπτώσεις, πολλοί επικαλούνται την έννοια και την ανάγκη ενός ψηφιακού συνταγματισμού, και μάλιστα με διάφορες παραλλαγές και διαστάσεις, όπως ο υβριδικός συνταγματισμός, ο αλγοριθμικός συνταγματισμός, ο πληροφοριακός συνταγματισμός, αλλά δεν υπάρχει συμφωνία αναφορικά με το τι εννοούμε με αυτό. Ωστόσο, παρατηρείται σύγκλιση απόψεων ως προς τη διαπίστωση σημαντικής ανισορροπίας εξουσίας στον ψηφιακό κόσμο, καθώς και την επιτακτική ανάγκη χρήσης κανονιστικών εργαλείων για τον έλεγχο αυτής της δύναμης. Θα μπορούσε να υποστηριχθεί κανείς ότι η έννοια αυτή εκφράζει περισσότερο μια αγωνία να συμφωνήσουμε στη σκοπιμότητα καθιέρωσης και διασφάλισης της ύπαρξης ενός κανονιστικού πλαισίου για την προστασία των θεμελιωδών δικαιωμάτων και την εξισορρόπηση των εξουσιών στο ψηφιακό περιβάλλον.²⁴⁴

Η προσέγγιση αυτή εδράζεται σε ορισμένους βασικούς άξονες-πυλώνες που έχουν ήδη επισημανθεί στο πλαίσιο της μελέτης. Πρόκειται καταρχάς για την επιταγή περιορισμού και ανα-εξισορρόπησης εξουσιών μεταξύ

²⁴¹ Celeste, E. (2022). *Digital Constitutionalism: The Role of Internet Bills of Rights* (p. 76 ff.). Routledge.

²⁴² De Gregorio, G. (2021). The rise of digital constitutionalism in the European Union. *International Journal of Constitutional Law*, 19(1) 2021, 41-70, 57, 65. <https://doi.org/10.1093/icon/moab001>

²⁴³ Terzis, P. (2024). Against digital constitutionalism. *European Law Open*, 3(2), 336-352. <https://doi.org/10.1017/elo.2024.15>

²⁴⁴ Celeste, E. (2025). Conceptual approaches to digital constitutionalism: A counter-critique. In *Digital constitutionalism* (p. 24 ff.). Nomos.

κράτους, ιδιωτικών φορέων εξουσίας και πολιτών.²⁴⁵ Ως προς τον δεύτερο άξονα που αναφέρεται στο κράτος δικαίου, επισημαίνεται ότι –πέραν των προφανών διαστάσεων (δέσμευση από σαφείς κανόνες, εγγυήσεις έννομης προστασίας κλπ.)– απαιτούνται πολιτικές και μέτρα που θα διασφαλίζουν την ανθρώπινη παρέμβαση, τη λογοδοσία και τον έλεγχο από δημοκρατικά νομιμοποιημένους φορείς εξουσίας. Στενά συνδεδεμένη με τον δεύτερο άξονα είναι η διασφάλιση μιας πολυδιάστατης διαφάνειας που λειτουργεί τόσο ως μέσο προστασίας των δικαιωμάτων όσο και ως μέσο επιβολής συμμόρφωσης σε όσους χρησιμοποιούν τέτοια συστήματα.

Η Πράξη για την Τεχνητή Νοημοσύνη αντιπροσωπεύει ένα ορόσημο της μετάβασης της ΕΕ από τον «ψηφιακό φιλελευθερισμό» και την αποχή από την επιβολή κανόνων σε μια μορφή «ψηφιακού συνταγματισμού», καθώς επιχειρεί την προστασία των θεμελιωδών δικαιωμάτων, αλλά και τη ρύθμιση της δραστηριότητας κυρίως των ιδιωτικών ψηφιακών «παικτών». Η επίτευξη αυτού του στόχου εξαρτάται χωρίς αμφιβολία από ένα ευέλικτο, αλλά και στιβαρό σύστημα εποπτείας και επιβολής των κανόνων δικαίου.²⁴⁶ Βέβαια, η διαδικασία της νομοθέτησης, ιδίως σε ένα ρευστό και μη προδιαγνώσιμο περιβάλλον, όπως αυτό της ΤΝ, είναι μια διαρκής διαδικασία.

Η ικανότητα του νομοθέτη να επιτελέσει τον ρυθμιστικό του ρόλο στην Κοινωνία της ΤΝ εξαρτάται, επίσης, από την ικανότητά του να «συνδιαλλαγεί» με τους άλλους νομοθέτες σε ένα παγκοσμιοποιημένο περιβάλλον, αλλά και με τα άλλα ρυθμιστικά «συστήματα» και να δημιουργήσει νέες μορφές «διαλόγου» και σύνθεσης, αφενός, μεταξύ των περισσότερων κρατικών, οικονομικών και κοινωνικών θεσμών και, αφετέρου, μεταξύ των περισσότερων κανονιστικών τεχνικών, αλλά και των άλλων –ιδιωτών– παραγόντων, που έχουν έναν εκ των πραγμάτων κομβικό ρόλο στο οικοσύστημα της πληροφορίας και της ΤΝ.

Η εποχή της ΤΝ μεταμορφώνει πτυχές της ζωής μας με τρόπους που θέτουν σε δοκιμασία τις αξίες, τους κανόνες που διέπουν την κοινωνία, την οικονομία, το κράτος, τον τρόπο ζωής, εργασίας και σκέψης, αλλά δεν είναι ακόμη σαφές αν αυτή η εποχή θα είναι μια εποχή στην οποία η προστασία δικαιωμάτων και ελευθεριών, οι συνθήκες ισότητας, το δημοκρατικό κράτος δικαίου και η ευημερία θα είναι εγγυημένα. Υποστηρίζεται ότι σε όλη τη διάρκεια της Ιστορίας, η τεχνολογική πρόοδος δεν μεταφραζόταν αυτόματα σε ευημερία για όλους.²⁴⁷ Αντίθετα, η κατανομή των οφελών

²⁴⁵ Με ορισμένους να επισημαίνουν ότι η κρίσιμη παράμετρος και επιδίωξη πρέπει να είναι η απόδοση ισχύος στους χρήστες έναντι των παρόχων υπηρεσιών, που είναι ταυτόχρονα φορείς ιδιωτικής, αλλά και οιοινεί δημόσιας εξουσίας.

²⁴⁶ Pollicino, O., & Paolucci, F. (2025). Digital Constitutionalism. In L. Floridi & M. Taddeo (Eds.), *A Companion to Digital Ethics* (p. 313 ff.). Wiley.

²⁴⁷ Acemoglu, D., & Johnson, S. (2023). *Power and progress: Our thousand-year struggle over technology and prosperity*. PublicAffairs.

από την τεχνολογία εξαρτάται σε μεγάλο βαθμό από τις πολιτικές επιλογές, τις θεσμικές ρυθμίσεις και την ισορροπία δυνάμεων μεταξύ διαφορετικών παραγόντων ισχύος και κοινωνικών ομάδων. Αυτή η ιστορική οπτική υποδηλώνει ότι, χωρίς συνειδητή και ολιστική παρέμβαση, τα οφέλη και τα κέρδη που προκύπτουν από την ΤΝ θα είναι άνισα κατανεμημένα. Το αίτημα της κοινωνικής συνοχής και ανάπτυξης σε δημοκρατικές χώρες συνδέεται με την ικανότητα των κοινωνιών να προσαρμόζονται, βρίσκοντας νέες ισορροπίες και αντισταθμίσεις.

Ίσως αυτό που πρέπει να επανασυζητηθεί είναι η αναγκαιότητα ενός (νέου) «κοινωνικού συμβολαίου»²⁴⁸ στην εποχή της ΤΝ. Ήδη στις αρχές του 21ου αιώνα, όταν καθίστατο σαφής η εξάρχουσα σχέση της πληροφορίας και η μετάβαση σε μια Κοινωνία της Πληροφορίας, διατυπωνόταν η σκέψη ότι «το «σύμφωνο αλληλεγγύης» της βιομηχανικής και μεταβιομηχανικής κοινωνίας, βασισμένο στις αρχές της διανεμητικής δικαιοσύνης, που παραδοσιακά υποστήριξε το κοινωνικό κράτος, μοιάζει απαρχαιωμένο. Η κοινωνία των δικτύων υπονομεύει αυτό το «σύμφωνο αλληλεγγύης» μέσω διαδικασιών και αλλαγών, όπως η εξατομίκευση της εργασίας και των αξιών, η απώλεια του κοινωνικού κεφαλαίου και η ελάττωση της εμπιστοσύνης μεταξύ των πολιτών, αλλά και μεταξύ πολιτών και κράτους».²⁴⁹ Στις Τεχνολογίες Πληροφορικής και Επικοινωνίας είχαν εναποτεθεί αντίστοιχα και πολλές προσδοκίες για μετασχηματισμό του κοινωνικού κράτους σε «πληροφοριακό κοινωνικό κράτος»,²⁵⁰ ενώ σήμερα ο ενωσιακός νομοθέτης επιχειρεί να καλλιεργήσει ρυθμιστικά τον «αλτρουισμό των δεδομένων».²⁵¹ Η ανάγκη δημόσιας συζήτησης για τους όρους ενός ψηφιακού «κοινωνικού συμβολαίου»²⁵² σε εθνικό και υπερεθνικό επίπεδο καθίσταται επιτακτική, καθώς

²⁴⁸ Ήδη από την εποχή του Διαφωτισμού η έννοια του «κοινωνικού συμβολαίου» έχει λειτουργήσει ως το θεμέλιο της οργανωμένης κοινωνίας. Από το όραμα του Thomas Hobbes για μια κυρίαρχη εξουσία που προστατεύει τους πολίτες από το χάος, το πλαίσιο του John Locke για τα φυσικά δικαιώματα και τη διακυβέρνηση μέσω συναίνεσης ή το ιδανικό της γενικής βούλησης του Jean-Jacques Rousseau, το «κοινωνικό συμβόλαιο» υπήρξε πάντοτε μια ζωντανή διαπραγμάτευση ανάμεσα στα άτομα, στους θεσμούς και στην εξουσία.

²⁴⁹ Μήτρου, Λ. (2006). Δικαίωμα συμμετοχής στην Κοινωνία της Πληροφορίας. Στο Θ. Παπαχρίστου, Τ. Βιδάλης, Λ. Μήτρου, Α. Τάκης (Επιμ.), *Το δικαίωμα συμμετοχής στην Κοινωνία της Πληροφορίας* (σ. 48 επ.). Σάκκουλα.

²⁵⁰ Στη θεωρία γίνεται ήδη λόγος για «πληροφοριακό κοινωνικό κράτος» (informational welfare state), αν και πρέπει να παρατηρήσει κανείς ότι η υπό μελέτη περίπτωση (Φινλανδία) δεν αντιστοιχεί στον μέσο όρο, ούτε όσον αφορά τον βαθμό εξάπλωσης της τεχνολογίας ούτε όσον αφορά τον βαθμό ανάπτυξης του κοινωνικού κράτους. Βλ. Castells, M., Pekka, H. (2002). *The Information Society and the Welfare State. The Finnish Model* (p. 76 ff.). Oxford 2002.

²⁵¹ Μέσω των σχετικών ρυθμίσεων της Πράξης για τη Διακυβέρνηση των Δεδομένων (Data Governance Act) ή και των ρυθμίσεων και σχεδιασμών για τα European Data Spaces.

²⁵² De Gregorio, G. (2022). *Digital Constitutionalism in Europe: Reframing Rights and Powers in the Algorithmic Society* (p. 25). Cambridge University Press. Η σκέψη του «κοινωνικού συμβολαίου» έχει επικριθεί, καθώς διαστρεβλώνει τις απαρχές της συνταγματικής κοινωνίας, ανιχνεύοντας την καταγωγή της σε μια προ-πολιτική και «φυσική» τάξη. Όπως γράφει ο P. Terzis, «ο ψηφιακός συνταγματισμός και η συνταγματική τάξη, σύμφωνα με το επιχείρημα, δεν αναδύονται από κατάσταση φύσης, αλλά από τη σύγκλιση ορισμένων πολιτικών και κοινωνικών δυνάμεων πάνω σε μια σειρά βασικών πολιτικών σκοπών και μια ικανότητα επιβεβαίωσής τους». Βλ. Terzis, P. (2024). Against digital constitutionalism. *European Law Open*, 3(2), 336-352. <https://doi.org/10.1017/el0.2024.15>

διαπιστώνουμε, εκτός των άλλων, τη μετάβαση σε νέες μορφές οργάνωσης της οικονομίας και της εργασίας, όπως η «πλατφορμοποίηση».²⁵³

Η ΕΕ ανταποκρίνεται ήδη υπό μια έννοια σε αυτήν την πρόκληση οριοθετώντας τι επιτρέπεται να κάνει και τι όχι η ΤΝ ή, ακριβέστερα, για ποιους σκοπούς δεν επιτρέπεται να τη χρησιμοποιούν οι άνθρωποι.²⁵⁴ Ωστόσο, τίθενται και άλλα μείζονος σημασίας ζητήματα: Πώς διασφαλίζεται ότι η μετάβαση στην εποχή της ΤΝ είναι συμπεριληπτική και δεν διευρύνει υπάρχοντα χάσματα ή δεν αφήνει πίσω και εκτός αυτούς που δεν μπορούν να ακολουθήσουν; Η εγγύηση αποτελεσματικής και δημοκρατικής διακυβέρνησης που υπόκειται σε λογοδοσία, η καταπολέμηση της κυριαρχίας και της συγκέντρωσης οικονομικής και πολιτικής ισχύος σε μη δημοκρατικά νομιμοποιημένους και ελέγξιμους φορείς και η διασφάλιση ότι η κοινωνία θα είναι ανθεκτική και δίκαιη (μέσω και της διασφάλισης ενός κοινωνικού πλέγματος ασφάλειας) εντάσσονται στη βασική προβληματική μιας τέτοιας συζήτησης. Το ερώτημα σε τελευταία ανάλυση δεν είναι εάν η ΤΝ θα μεταμορφώσει την κοινωνία, αλλά ποιος θα οδηγήσει αυτήν τη μεταμόρφωση, προς ποια κατεύθυνση και στόχο, και επί τη βάση ποιων επιλογών και αξιών.²⁵⁵

Τα ζητήματα του ψηφιακού συνταγματισμού και του νέου «κοινωνικού συμβολαίου» εξετάζει ο Αντώνης Κουρουτάκης στη συμβολή του στην παρούσα μελέτη.

²⁵³ Βλ. σχετικά την πρόσφατη μελέτη του Ευρωπαϊκού Κοινοβουλίου που συντάχθηκε από τους González Vázquez, I., Fernández Macías, E., Wright, S., & Villani, D. (2025). *Digital monitoring, algorithmic management and the platformisation of work in Europe* (JRC Research Report No. JRC143072). Publications Office of the European Union. <https://doi.org/10.2760/9406086> Στο ίδιο πλαίσιο ιδιαίτερα σημαντικές οι επισημάνσεις των, Aloisi, A., & De Stefano, V. (2025). *Ο εργοδότης μου είναι ένας αλγόριθμος: Τεχνητή νοημοσύνη, πλατφόρμες εργασίας και εργατικό δυναμικό* (σ. 39 επ.). Πανεπιστημιακές Εκδόσεις Κρήτης

²⁵⁴ Προσδιορίζοντας στο άρθρο 5 της AI Act τους μη αποδεκτούς κινδύνους και άρα τις απαγορευμένες χρήσεις.

²⁵⁵ Βλ. και τους προβληματισμούς που εκφράζει ο Rahwan, I. (2018). Society-in-the-loop: programming the algorithmic social contract. *Ethics and information technology*, 20(1), 5-14. <https://doi.org/10.1007/s10676-017-9430-8>

11.4. Η ρύθμιση ως έκφραση της ψηφιακής κυριαρχίας: Η Τεχνητή Νοημοσύνη, το «κοινωνικό συμβόλαιο» και η εμφάνιση του ψηφιακού συνταγματισμού

11.4.1. Εισαγωγή

Η ταχεία διάδοση των συστημάτων Τεχνητής Νοημοσύνης στη δημόσια διοίκηση, στις αγορές, στην εκπαίδευση, στην υγειονομική περίθαλψη, στην αστυνόμευση και την καθημερινή ζωή έχει προκαλέσει μια ολοκληρωμένη ρυθμιστική αντίδραση στην Ευρωπαϊκή Ένωση. Μέσω της Πράξης για την Τεχνητή Νοημοσύνη, η ΕΕ έχει θεσπίσει ένα καινοτόμο κανονιστικό πλαίσιο που επικεντρώνεται σε μια αξιολόγηση βάσει κινδύνου. Ενώ τα συστήματα ΤΝ που θεωρούνται ότι ενέχουν απαράδεκτο κίνδυνο απαγορεύονται, τα άλλα συστήματα κατηγοριοποιούνται ανάλογα με το επίπεδο κινδύνου τους, υψηλό, περιορισμένο ή ελάχιστο, και πρέπει να συμμορφώνονται με τις αντίστοιχες νομικές υποχρεώσεις.²⁵⁶ Επιπλέον, η Πράξη για την Τεχνητή Νοημοσύνη εισάγει μια συγκεκριμένη νομική διάκριση μεταξύ των εταιρειών που αναπτύσσουν την τεχνολογία («πάροχοι», σύμφωνα με το άρθρο 3 παρ. 3) και των οντοτήτων που τη χρησιμοποιούν σε επαγγελματικό περιβάλλον («χρήστες», σύμφωνα με το άρθρο 3 παρ. 4). Ο Κανονισμός για την ΤΝ περιγράφεται συχνά ως νόμος περί της ευθύνης για ελαττωματικά προϊόντα,²⁵⁷ πλαίσιο διαχείρισης κινδύνων²⁵⁸ ή εργαλείο διακυβέρνησης της καινοτομίας.²⁵⁹ Αυτές οι περιγραφές είναι χρήσιμες από θεωρητικής σκοπιάς, αλλά ατελείς από κανονιστικής. Παραβλέπουν το γεγονός ότι ο Κανονισμός για την ΤΝ δεν αφορά μόνο τη ρύθμιση της αγοράς ή την τεχνολογική ασφάλεια, αλλά και τον τρόπο με τον οποίο ασκείται η εξουσία επί των ατόμων σε μια αυτοματοποιημένη κοινωνία.

²⁵⁶ Βλ. για παράδειγμα τα Κεφάλαια 3 και 4.

²⁵⁷ Almada, M., & Petit, N. (2025). The EU AI Act: Between the rock of product safety and the hard place of fundamental rights. *Common Market Law Review*, 62(1), 85-120. <https://doi.org/10.54648/COLA2025004>

²⁵⁸ Schuett, J. (2024). Risk management in the Artificial Intelligence Act. *European Journal of Risk Regulation*, 15(2), 367-385. <https://doi.org/10.1017/err.2023.1>

²⁵⁹ European Commission. (2023, September 13). *State of the Union 2023 address by President von der Leyen* (Speech/23/4426). https://commission.europa.eu/strategy-and-policy/state-union/state-union-2023_en

Γιατί αυτή η μορφή εξουσίας δεν ρυθμίζεται επαρκώς από το παραδοσιακό Δημόσιο Δίκαιο; Η δυσκολία υπαγωγής της ΤΝ στο παραδοσιακό Δημόσιο Δίκαιο έγκειται στο γεγονός ότι η ΤΝ δεν αποτελεί απλώς μια ακόμη τεχνολογία που απαιτεί ρύθμιση της αγοράς. Αντιθέτως, συνιστά ένα σύνολο τεχνολογιών που αναδιαρθρώνουν την επιστημολογική και αποφασιστική υποδομή, μέσω της οποίας ασκείται εξουσία επί των ατόμων και των πληθυσμών. Ως ακρογωνιαίος λίθος της επανάστασης της τεχνολογίας των πληροφοριών, η ΤΝ πολλαπλασιάζει την ανθρώπινη νοημοσύνη, συγκρίσιμη σε μετασχηματιστικό εύρος με τη Γεωργική Επανάσταση, που πολλαπλασίασε την παραγωγή τροφίμων, και τη Βιομηχανική Επανάσταση που πολλαπλασίασε τη φυσική δύναμη. Αυτή η κλίμακα μετασχηματισμού μοιάζει με αυτό που ο Michel Foucault περιέγραψε ως μηχανισμό κυβερνητικότητας:²⁶⁰ ένα σύστημα μέσω του οποίου η εξουσία ασκείται όχι με άμεση εντολή, αλλά μέσω της γνώσης, της ταξινόμησης και της συνεχούς αξιολόγησης των πληθυσμών. Για τον λόγο αυτόν, η απάντηση της ΕΕ δεν μπορεί να γίνει κατανοητή απλώς ως τεχνικός κανονισμός. Πρέπει να ερμηνευθεί σε σχέση με ευρύτερα ζητήματα κυριαρχίας, «κοινωνικού συμβολαίου» και συνταγματισμού.

Η παρούσα έκθεση υποστηρίζει ότι η Πράξη για την Τεχνητή Νοημοσύνη υπερβαίνει την απλή τεχνική εποπτεία και αντιπροσωπεύει μια προσπάθεια αντιμετώπισης της νέας μορφής εξουσίας που εισάγει η ΤΝ. Δεν δημιουργεί νέα δικαιώματα, αλλά εξειδικεύει την εφαρμογή των υφιστάμενων σε *ex ante system design* (εκ των προτέρων σχεδιασμό του συστήματος). Αποτελεί έναν ζωτικό μηχανισμό για την επιβολή της κυριαρχίας της ΕΕ, καθορίζοντας τα συγκεκριμένα νομικά, κανονιστικά και ηθικά όρια, εντός των οποίων μπορούν να λειτουργούν νόμιμα τα συστήματα αυτά. Με τον τρόπο αυτόν, η Πράξη για την Τεχνητή Νοημοσύνη αποσκοπεί στη διατήρηση της θεμελιώδους σχέσης μεταξύ των ατόμων, των δημόσιων και ιδιωτικών Αρχών και της αυτοματοποιημένης λήψης αποφάσεων με τη βοήθεια της ΤΝ. Τελικά, μεταφράζοντας τις συνταγματικές αρχές σε εκ των προτέρων υποχρεώσεις σχεδιασμού και διακυβέρνησης για τα τεχνικά συστήματα, το πλαίσιο αυτό πρωτοπορεί σε μια μορφή συνταγματισμού, προσαρμοσμένου στα ψηφιακά δεδομένα, διασφαλίζοντας ότι ο κανόνας δικαίου είναι ενσωματωμένος στην ίδια την αρχιτεκτονική της ψηφιακής εποχής.

11.4.2. Ποια νέα μορφή εξουσίας εισάγει η Τεχνητή Νοημοσύνη στην κοινωνία;

Τα συστήματα Τεχνητής Νοημοσύνης, τα οποία χρησιμοποιούνται εξίσου από ιδιωτικούς και δημόσιους φορείς, επηρεάζουν όλο και περισσότερο τις αποφάσεις που διαμορφώνουν την πρόσβαση στην ελευθερία του λόγου,

²⁶⁰ Foucault, M. (2007). *Security, territory, population: Lectures at the Collège de France, 1977-1978* (M. Senellart, Ed.; G. Burchell, Trans.). Palgrave Macmillan.

στις αγορές και τις υπηρεσίες. Τόσο στον δημόσιο όσο και στον ιδιωτικό τομέα, η ΤΝ χρησιμοποιείται όλο και περισσότερο για την ταξινόμηση των ατόμων, την πρόβλεψη της συμπεριφοράς τους και την κατανομή των πόρων. Είτε αυτά τα συστήματα αξιολογούν τον κίνδυνο, καθορίζουν την επιλεξιμότητα ή φιλτράρουν πληροφορίες, λειτουργούν ως οι κύριοι κινητήριοι μοχλοί που καθοδηγούν ή αντικαθιστούν την ανθρώπινη κρίση. Όπως έχουν παρατηρήσει οι Karen Yeung²⁶¹ και Mireille Hildebrandt,²⁶² αυτή η μετατόπιση σηματοδοτεί την άνοδο της αλγοριθμικής διακυβέρνησης: μια νέα μορφή εξουσίας που ασκείται μέσω αυτοματοποιημένων αρχιτεκτονικών λήψης αποφάσεων. Σε αντίθεση με την παραδοσιακή διακυβέρνηση, η οποία βασίζεται στην ανθρώπινη διακριτική ευχέρεια και σε διαφανείς θεσμικές διαδικασίες, η αλγοριθμική διακυβέρνηση λειτουργεί μέσω της «αόρατης» λογικής του κώδικα, αλλάζοντας ριζικά τον τρόπο με τον οποίο ασκείται η εξουσία επί των ατόμων και των πληθυσμών.

Η ΤΝ αντιπροσωπεύει μια ριζική εξέλιξη της ορθολογικής-νομικής εξουσίας του Max Weber, μετατοπίζοντας τη βάση της διοικητικής νομιμότητας από κανόνες που ερμηνεύονται από τον άνθρωπο σε αδιαφανείς και πολύπλοκες στατιστικές πράξεις.²⁶³ Αυτή η μετάβαση ενισχύει την προειδοποίηση της Hannah Arendt για μια «κυβέρνηση από κανέναν», όπου η μεταβίβαση εξουσίας σε αλγορίθμους εξασφαλίζει ότι η ευθύνη δεν διαχέεται απλώς σε μια ιεραρχία, αλλά διαλύεται εντελώς μέσα σε ένα σύστημα «μαύρου κουτιού».²⁶⁴

Αυτή η εξουσία είναι διακριτή από πολλές απόψεις. Η εξουσία που εισάγουν τα συστήματα ΤΝ δεν είναι απλώς μια ταχύτερη εκδοχή της παραδοσιακής γραφειοκρατίας, αλλά είναι ποιοτικά διαφορετική ως προς την εκτέλεση και τον αντίκτυπό της. Ο πρώτος πυλώνας αυτής της αλλαγής είναι η μεταβίβαση της κρίσης σε στατιστικά συστήματα. Ενώ τα νομικά και διοικητικά πλαίσια βασίζονταν παραδοσιακά στην ανθρώπινη λογική, στη διακριτική ευχέρεια και την υποχρέωση αιτιολόγησης, τα συστήματα ΤΝ λειτουργούν μέσω στατιστικών συμπερασμάτων και αναγνώρισης προτύπων, σχηματίζοντας μια «αλγοριθμική κυβερνητικότητα».²⁶⁵ Όπως δείχνει η Danielle Keats Citron, η αυτοματοποιημένη λήψη αποφάσεων μπορεί να παρακάμψει τις παραδοσιακές διαδικαστικές εγγυήσεις και να υπονομεύσει τις προστασίες

²⁶¹ Yeung, K. (2017). 'Hypnudge': Big data as a mode of regulation by design. *Information, Communication & Society*, 20(1), 118-136. <https://doi.org/10.1080/1369118X.2016.1186713>

²⁶² Hildebrandt, M. (2016). *Smart technologies and the end(s) of law: Novel entanglements of law and technology*. Edward Elgar Publishing.

²⁶³ Weber, M. (1978). *Economy and society: An outline of interpretive sociology* (G. Roth & C. Wittich, Eds.). University of California Press. (Original work published 1922).

²⁶⁴ Arendt, H. (1998). *The human condition* (2nd ed.). University of Chicago Press. (Original work published 1958).

²⁶⁵ Rouvroy, A. (2016). The end(s) of critique: Data-behaviourism vs. due-process. In M. Hildebrandt & K. De Vries (Eds.), *Privacy, due process and the computational turn: The philosophy of law meets the philosophy of technology*. Routledge.

της δίκαιης διαδικασίας.²⁶⁶ Όταν οι αποφάσεις βασίζονται σε μοντέλα και όχι σε λογικούς συλλογισμούς, η παραδοσιακή ικανότητα και το δικαίωμα ενός πολίτη να αμφισβητήσει μια απόφαση αρχίζει να διαλύεται.

Επιπλέον, η ΤΝ ασκεί μια δομική και αόρατη εξουσία. Σε αντίθεση με ένα ορατό δικαστήριο ή μια κυβερνητική υπηρεσία, η ΤΝ, η οποία στη συντριπτική πλειονότητα των περιπτώσεων αναπτύσσεται, σχεδιάζεται και πωλείται από κερδοσκοπικές εταιρείες, είναι συχνά ενσωματωμένη στο παρασκήνιο της ψηφιακής ζωής, καθιστώντας δύσκολη την αναγνώρισή της, πόσω μάλλον την αμφισβήτησή της. Ο Frank Pasquale²⁶⁷ περιγράφει αυτό το φαινόμενο ως «κοινωνία μαύρου κουτιού», όπου ιδιόκτητοι αλγόριθμοι καθορίζουν τις ευκαιρίες της ζωής πίσω από ένα πέπλο εταιρικής μυστικότητας. Ομοίως, η Julie E. Cohen²⁶⁸ σημειώνει ότι αυτή η υποδομή εξουσίας δημιουργεί μια νέα «νομική γεωγραφία», όπου η εξουσία ασκείται μέσω του τεχνικού σχεδιασμού των πλατφορμών και όχι μέσω των παραδοσιακών θεσμικών διαδικασιών.

Τα συστήματα ΤΝ δεν επεξεργάζονται απλώς την πραγματικότητα, αλλά την κατηγοριοποιούν. Προκαθορίζοντας τους δανειολήπτες ως «υψηλού κινδύνου» ή τους πολίτες ως «δυσνητικά ριζοσπάστες», η ΤΝ περιορίζει την αυτενέργεια των ανθρώπων αξιωματούχων. Με αυτόν τον τρόπο, η ΤΝ ασκεί γνωσιακή εξουσία, καθώς η εσωτερική της λογική υπαγορεύει την «αλήθεια» και το «αποτέλεσμα», που οι άνθρωποι διαχειριστές είναι στη συνέχεια υποχρεωμένοι να εκτελέσουν.²⁶⁹

Ωστόσο, αυτή η εξουσία εκτείνεται από τις δομές κράτους στην αγορά, όπου εξελίσσεται σε αυτό που η Shoshana Zuboff²⁷⁰ αποκαλεί «εργαλειακή εξουσία». Σε αυτό το πλαίσιο, η ΤΝ δεν κατηγοριοποιεί απλώς το άτομο, αλλά χρησιμοποιεί αυτές τις κατηγοριοποιήσεις για να «ρυθμίζει» και να «κατευθύνει» την ανθρώπινη συμπεριφορά. Εδώ ο στόχος μετατοπίζεται από τη διοικητική αποτελεσματικότητα στην αποκόμιση κέρδους μέσω της συνεχούς παρακολούθησης και της αυτοματοποιημένης χειραγώγησης του περιβάλλοντος του καταναλωτή.

Τέλος, η κλίμακα και η αυτοματοποίηση αυτών των συστημάτων μετατρέπουν τη διακυβέρνηση από μια περιστασιακή ανθρώπινη παρέμβαση σε μια συνεχή, ευρεία μεσολαβούμενη από μηχανές αξιολόγηση. Όπως έχει παρατηρήσει

²⁶⁶ Citron, D. K. (2008). Technological Due Process. *Washington University Law Review*, 85(6), 1249-1313.

²⁶⁷ Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.

²⁶⁸ Cohen, J. E. (2019). *Between Truth and Power: The Legal Constructions of Information Capitalism*. Oxford University Press.

²⁶⁹ Hildebrandt, M. (2015). *Smart technologies and the end(s) of law: Novel entanglements of law and technology*. Edward Elgar Publishing.

²⁷⁰ Zuboff, S. (2019). *The Age of Surveillance Capitalism*. PublicAffairs.

η Karen Yeung,²⁷¹ αυτό δημιουργεί ένα περιβάλλον «υπερ-ώθησης», όπου η αυτοματοποιημένη εξουσία λειτουργεί συνεχώς και σε παγκόσμια κλίμακα. Πρόκειται για μια μορφή εξουσίας που βρίσκεται σε έναν μεταβατικό χώρο: Δεν είναι ούτε καθαρά ιδιωτική ούτε καθαρά δημόσια. Είναι μια αυτοματοποιημένη εξουσία που υπαγορεύει τα θεμελιώδη δικαιώματα και τις ευκαιρίες στη ζωή, απαιτώντας μια ρυθμιστική απάντηση που υπερβαίνει την τεχνική ασφάλεια για να αντιμετωπίσει τα ίδια τα θεμέλια του «κοινωνικού μας συμβολαίου».

Η ανησυχία αυτή δεν είναι μόνο θεωρητική. Το Ενωσιακό Δίκαιο έχει ήδη αναγνωρίσει την ιδιαίτερη απειλή της αυτοματοποιημένης λήψης αποφάσεων. Το άρθρο 22 του Γενικού Κανονισμού για την Προστασία Δεδομένων (GDPR) καθιερώνει το δικαίωμα του ατόμου να μην υπόκειται σε απόφαση που λαμβάνεται αποκλειστικά βάσει αυτοματοποιημένης επεξεργασίας και παράγει έννομα αποτελέσματα ή το επηρεάζει σημαντικά. Η διάταξη αυτή αντικατοπτρίζει την ιδέα ότι η νομιμότητα των αποφάσεων δεν εξαρτάται μόνο από την ακρίβειά τους, αλλά και από τη δυνατότητα ανθρώπινης κρίσης, αιτιολόγησης και αμφισβήτησης. Ωστόσο, ο GDPR περιορίζεται στο επίπεδο της επεξεργασίας δεδομένων και δεν ρυθμίζει τη συνολική αρχιτεκτονική του συστήματος που παράγει την απόφαση. Η Πράξη για την Τεχνητή Νοημοσύνη έρχεται να καλύψει αυτό το κενό, μεταφέροντας τη λογική του άρθρου 22 από το αποτέλεσμα της απόφασης στον σχεδιασμό του συστήματος.

11.4.3. Χρησιμοποιεί η Ευρωπαϊκή Ένωση τη ρύθμιση για να επιβάλει την κυριαρχία της στον τρόπο με τον οποίο τα συστήματα Τεχνητής Νοημοσύνης λαμβάνουν αποφάσεις;

Στον τρέχοντα πολιτικό διάλογο, η Ευρωπαϊκή Ένωση επικαλείται συχνά την έννοια της «ψηφιακής κυριαρχίας» ως μέρος της ευρύτερης στρατηγικής ατζέντας της.²⁷² Στο πλαίσιο της ΤΝ, η έννοια αυτή γίνεται πιο ακριβής και πιο βαθιά. Η ψηφιακή κυριαρχία για την ΕΕ αναφέρεται στην ικανότητα να διαμορφώνει και να ελέγχει τις συνθήκες υπό τις οποίες λειτουργούν οι ψηφιακές τεχνολογίες εντός των συνόρων της και, όλο και περισσότερο, εκτός αυτών, μια μορφή κανονιστικής εξουσίας που δεν βασίζεται στον εδαφικό έλεγχο, αλλά στην κανονιστική αρχή.²⁷³

Η ΕΕ δεν είναι παγκόσμιος ηγέτης στην ανάπτυξη της ΤΝ, στις υποδομές ή στην εταιρική κυριαρχία, όπως είναι οι Ηνωμένες Πολιτείες ή η Κίνα. Ωστόσο, μέσω ρυθμιστικών μέτρων, αναδεικνύεται σε παγκόσμιο ηγέτη στον καθορισμό του τρόπου με τον οποίο η ΤΝ μπορεί να χρησιμοποιηθεί νόμιμα.

²⁷¹ Yeung, K. (2017). 'Hypernudge': Big data as a mode of regulation by design. *Information, Communication & Society*, 20(1), 118-136. <https://doi.org/10.1080/1369118X.2016.1186713>

²⁷² von der Leyen, U. (2025, September 10). *State of the Union address 2025*. European Commission. https://commission.europa.eu/strategy-and-policy/state-union/state-union-2025_en

²⁷³ Bradford, A. (2020). *The Brussels Effect: How the European Union Rules the World*. Oxford University Press.

Κεντρικό ρόλο σε αυτήν την προσπάθεια διαδραματίζει η Πράξη της ΕΕ για την Τεχνητή Νοημοσύνη, η οποία κατηγοριοποιεί τα συστήματα ΤΝ ανάλογα με τον κίνδυνο και επιβάλλει αντίστοιχες υποχρεώσεις.

Σύμφωνα με την Πράξη για την Τεχνητή Νοημοσύνη, η ΕΕ διεκδικεί κυριαρχία επί των κανονιστικών όρων για την ανάπτυξη της ΤΝ με διάφορους τρόπους: Καθορίζει ποιοι τύποι χρήσεων της ΤΝ είναι απαράδεκτοι, θεσπίζει απαιτήσεις εποπτείας βάσει κινδύνου, επιβάλλει ενσωματωμένα μέτρα προστασίας πριν από την ανάπτυξη και καθορίζει πρότυπα διαφάνειας, ικνηλασιμότητας και ανθρώπινης εποπτείας. Για παράδειγμα, η Πράξη απαγορεύει τις πρακτικές ΤΝ που χαρακτηρίζει ως «απαράδεκτες», όπως εκείνες που εκμεταλλεύονται ευπάθειες ή επιτρέπουν την κοινωνική βαθμολόγηση, ενώ απαιτεί αυστηρούς ελέγχους και τεκμηρίωση για συστήματα «υψηλού κινδύνου».

Αυτή η διεκδίκηση ρυθμιστικής εξουσίας υπερβαίνει την κυριαρχία επί του εδάφους ή των επιχειρήσεων. Αυτό που διακυβεύεται είναι η κυριαρχία επί των κανονιστικών όρων, υπό τους οποίους τα αυτοματοποιημένα συστήματα μπορούν να ασκούν εξουσία. Δηλαδή, η ΕΕ καθορίζει όχι μόνο ποιος μπορεί να αναπτύξει ΤΝ, αλλά και με ποιους τρόπους τα συστήματα ΤΝ μπορούν να συλλογίζονται και να ενεργούν έναντι των ανθρώπων χωρίς να παραβιάζουν τα δικαιώματα ή τα πρότυπα ασφάλειας. Η ρυθμιστική επιρροή της ΕΕ συχνά εκτείνεται πέρα από τα επίσημα σύνορά της μέσω ενός φαινομένου που ονομάζεται συνήθως «Brussels Effect». Ο όρος αυτός, που επινοήθηκε από την Anu Bradford,²⁷⁴ περιγράφει τον τρόπο με τον οποίο οι κανόνες της ΕΕ, λόγω του μεγέθους και της ελκυστικότητας της αγοράς της ΕΕ και της αυστηρότητας των προτύπων της, τείνουν να γίνονται *de facto* παγκόσμια πρότυπα, επειδή οι επιχειρήσεις τα υιοθετούν σε παγκόσμιο επίπεδο για να απλοποιήσουν τη συμμόρφωση και να εξασφαλίσουν την πρόσβαση στην αγορά.

Στο πλαίσιο της ΤΝ, αυτό σημαίνει ότι η ΕΕ δεν ρυθμίζει απλώς την εφαρμογή της ΤΝ σε εθνικό και ενωσιακό επίπεδο, αλλά προσπαθεί να καθορίσει παγκόσμιες νόρμες για την αλγοριθμική λήψη αποφάσεων. Εάν οι πολυεθνικοί κολοσσοί και οι διεθνείς δικαιοδοσίες υιοθετήσουν το ευρωπαϊκό πλαίσιο για την ΤΝ με σκοπό την πρόσβαση στην ενιαία αγορά ή την εναρμόνιση των λειτουργικών τους εξόδων, τότε τα ρυθμιστικά πρότυπα της ΕΕ θα καταστούν *de facto* ο παγκόσμιος κανόνας διακυβέρνησης της τεχνολογίας αυτής. Υπό αυτό το πρίσμα, η κυριαρχία εδώ γίνεται κυριαρχία επί των προτύπων αποδεκτής μηχανικής συλλογιστικής σχετικά με τους ανθρώπους. Αυτή η μορφή ρυθμιστικής φιλοδοξίας, που διαμορφώνει όχι μόνο τις αγορές, αλλά και τα κανονιστικά πλαίσια για τον τρόπο με τον οποίο επιτρέπονται τα συστήματα ΤΝ σε παγκόσμιο επίπεδο, είναι σχετικά

²⁷⁴ Ό.π.

καινοτόμος και αντικατοπτρίζει μια μετατόπιση στον τρόπο με τον οποίο μπορεί να ασκηθεί η γεωπολιτική επιρροή σε μια εποχή ταχέως εξελισσόμενων τεχνολογιών ΤΝ.

11.4.4. Είναι η Πράξη για την Τεχνητή Νοημοσύνη μια επαναδιαπραγμάτευση του «κοινωνικού συμβολαίου» στην εποχή της αυτοματοποιημένης λήψης αποφάσεων;

Το «κοινωνικό συμβόλαιο» είναι μια θεμελιώδης έννοια στην πολιτική φιλοσοφία, που περιγράφει μια σιωπηρή συμφωνία μεταξύ των ατόμων και της δημόσιας Αρχής (του κράτους), με την οποία τα άτομα συναινούν στη διακυβέρνηση σε αντάλλαγμα για την προστασία των θεμελιωδών δικαιωμάτων τους και του κράτους δικαίου. Κλασικοί θεωρητικοί του «κοινωνικού συμβολαίου», όπως ο Thomas Hobbes, ο John Locke και ο Jean-Jacques Rousseau, διατύπωσαν ο καθένας τη δική του εκδοχή αυτής της ιδέας.²⁷⁵ Παρά τις διαφορές τους, η δομική λογική του «κοινωνικού συμβολαίου» και για τους τρεις είναι ότι τα άτομα συμφωνούν να παραχωρήσουν ορισμένες φυσικές ελευθερίες προκειμένου να αποκτήσουν την ασφάλεια και τη δημόσια τάξη που παρέχει μια κυρίαρχη ή πολιτική κοινότητα. Αυτό το πλαίσιο χρησιμεύει για να δικαιολογήσει τη νομιμότητα της πολιτικής εξουσίας και να θεμελιώσει τις υποχρεώσεις της κυβέρνησης στη νομιμότητα, στη λογοδοσία και την προστασία των δικαιωμάτων.

Σύμφωνα με τις αρχές του κράτους δικαίου που θεμελιώνουν το σύγχρονο «κοινωνικό συμβόλαιο», είναι αυτονόητο ότι οι αποφάσεις που επηρεάζουν τη ζωή των ανθρώπων πρέπει να είναι διαφανείς, αμφισβητήσιμες και να αποδίδονται σε έναν υπεύθυνο φορέα, και ότι οι άνθρωποι που επηρεάζονται έχουν το δικαίωμα να αμφισβητήσουν ουσιαστικά τέτοιες αποφάσεις. Στην εποχή της ΤΝ, ωστόσο, αυτή η βάση διαταράσσεται, καθώς τα άτομα αντιμετωπίζουν όλο και περισσότερο αποφάσεις που δεν προέρχονται από ανθρώπινες αρχές, αλλά από αυτοματοποιημένα συστήματα που ενδέχεται να λειτουργούν εκτός των καθιερωμένων νομικών και διαδικαστικών πλαισίων. Αυτό εγείρει κρίσιμα ερωτήματα σχετικά με τον τρόπο με τον οποίο εφαρμόζονται βασικές αρχές, όπως η νόμιμη διαδικασία, η λογοδοσία και η εκπροσώπηση όταν η λήψη αποφάσεων ανατίθεται σε αλγορίθμους.

Η ενσωμάτωση της ΤΝ σε ευαίσθητους τομείς, από την απασχόληση και τη χορήγηση δανείων έως την επιβολή του νόμου και δυνητικά όλες τις δημόσιες υπηρεσίες, φέρνει αυτές τις ανησυχίες στο προσκήνιο. Για παράδειγμα, η αυτοματοποιημένη λήψη αποφάσεων που επηρεάζει τα νομικά

²⁷⁵ Hobbes, T. (1996). *Leviathan* (R. Tuck, Ed.). Cambridge University Press (Original work published 1651). Locke, J. (1988). *Two treatises of government* (P. Laslett, Ed.). Cambridge University Press (Original work published 1689). Rousseau, J.-J. (1997). *The social contract* (V. Gourevitch, Ed. & Trans.). Cambridge University Press (Original work published 1762).

δικαιώματα ή την πρόσβαση σε βασικές υπηρεσίες μπορεί να υπονομεύσει τις παραδοσιακές έννοιες της δίκαιης διαδικασίας, εάν τα επηρεαζόμενα άτομα δεν διαθέτουν ουσιαστικές πληροφορίες σχετικά με τον τρόπο λήψης των αποφάσεων ή τον τρόπο αμφισβήτησής τους. Μια βασική πρόκληση είναι ότι τα συστήματα ΤΝ συχνά λειτουργούν αδιαφανώς, με περιορισμένη πρόσβαση όσον αφορά την εσωτερική τους λογική, καθιστώντας δύσκολο για τους ανθρώπους να κατανοήσουν, να αμφισβητήσουν ή να ζητήσουν αποζημίωση για αποφάσεις που τους επηρεάζουν.

Αυτή η εξέλιξη εγείρει διάφορα θεμελιώδη ερωτήματα που είναι κεντρικά για την αναδιαμόρφωση του «κοινωνικού συμβολαίου» σε μια αυτοματοποιημένη εποχή: Τι σημαίνει δίκαιη διαδικασία όταν οι αποφάσεις λαμβάνονται από μηχανές και όχι από ανθρώπους; Πώς μπορούν τα άτομα να αμφισβητήσουν τα αποτελέσματα που παράγονται από αδιαφανή μοντέλα; Ποιος φέρει την ευθύνη για τις βλάβες που προκαλούνται από αυτοματοποιημένα συστήματα; Και πώς μπορούν να διατηρηθούν η ανθρώπινη δράση και η λογοδοσία όταν τα τεχνολογικά συστήματα είναι βαθιά ενσωματωμένα στις δημόσιες και ιδιωτικές διαδικασίες λήψης αποφάσεων;

Η Πράξη της ΕΕ για την Τεχνητή Νοημοσύνη μπορεί να ερμηνευθεί ως μια προσπάθεια να απαντηθούν αυτά τα ερωτήματα, επαναβεβαιώνοντας τις βασικές νομικές, κανονιστικές και ηθικές αρχές στη ρυθμιστική αντιμετώπιση της ΤΝ. Η Πράξη ενσωματώνει υποχρεώσεις που αποσκοπούν στη διατήρηση της ανθρώπινης εξουσίας και των ανθρωπίνων δικαιωμάτων εν όψει της αυτοματοποίησης, συμπεριλαμβανομένων απαιτήσεων ανθρώπινης εποπτείας, που διασφαλίζουν ότι τα συστήματα μπορούν να παρακολουθούνται και να παρακάμπτονται από ανθρώπους όπου είναι απαραίτητο, καθώς και υποχρεώσεων διαφάνειας, τεκμηρίωσης και εξήγησης, που καθιστούν τη συμπεριφορά και τις αποφάσεις του συστήματος πιο κατανοητές και υπεύθυνες.

Παράλληλα με αυτά τα μέτρα, η Πράξη για την Τεχνητή Νοημοσύνη απαιτεί επίσης αξιολογήσεις κινδύνου πριν από την ανάπτυξη και δομές λογοδοσίας σε όλη την αλυσίδα αξίας των συστημάτων ΤΝ [συλλογή δεδομένων, εκπαίδευση του μοντέλου (training), εφαρμογή του σε προϊόντα (deployment), χρήση από τον τελικό χρήστη], προκειμένου να διασφαλίζεται ότι οι πιθανές βλάβες εντοπίζονται και μετριαζονται εκ των προτέρων. Για τα συστήματα ΤΝ «υψηλού κινδύνου», τα οποία είναι πιο πιθανό να επηρεάσουν τα θεμελιώδη δικαιώματα, η Πράξη επιβάλλει λεπτομερή τεκμηρίωση, καταγραφή και μηχανισμούς εποπτείας, ενώ παράλληλα εξουσιοδοτεί τους χρήστες και τις Αρχές να παρεμβαίνουν όταν τα συστήματα δυσλειτουργούν ή λειτουργούν με απρόβλεπτο τρόπο.

Επιπλέον, οι οργανώσεις της Κοινωνίας των Πολιτών και οι υπερασπιστές των δικαιωμάτων έχουν τονίσει τη σημασία των Εκτιμήσεων Αντικτύπου στα Θεμελιώδη Δικαιώματα (FRIAs). Οι FRIAs αποτελούν νομική υποχρέωση

βάσει του άρθρου 27 της Πράξης για την Τεχνητή Νοημοσύνη για ορισμένους χρήστες συστημάτων ΤΝ «υψηλού κινδύνου». Παράλληλα, οι οργανώσεις της Κοινωνίας των Πολιτών και οι υπερασπιστές των δικαιωμάτων έχουν αναπτύξει οδηγίες και εργαλεία για να διασφαλίσουν ότι αυτές οι αξιολογήσεις λειτουργούν ως ουσιαστικές διασφαλίσεις²⁷⁶. Αυτές οι αξιολογήσεις έχουν σχεδιαστεί για να εκτιμούν τις πιθανές επιπτώσεις της ανάπτυξης συστημάτων ΤΝ στα θεμελιώδη δικαιώματα και να επιτρέπουν στα άτομα και στους φορείς δημοσίου συμφέροντος να εντοπίζουν, να αμφισβητούν και να διορθώνουν επιβλαβή ή διακριτικά αποτελέσματα.

Συνολικά, αυτά τα στοιχεία δεν υποδηλώνουν επαναδιαπραγμάτευση του «κοινωνικού συμβολαίου», αλλά αποτελούν μια μεταβολή του μηχανισμού εφαρμογής του στην εποχή της αυτοματοποιημένης λήψης αποφάσεων.

Η βασική πρόταση είναι ότι κανένα αυτοματοποιημένο σύστημα δεν μπορεί να ασκεί σημαντική εξουσία επί των ατόμων χωρίς να υπόκειται σε διασφαλίσεις ισοδύναμες με αυτές που επιβάλλονται στις παραδοσιακές δημόσιες Αρχές. Υπό αυτήν την έννοια, η Πράξη για την Τεχνητή Νοημοσύνη δεν επαναπροσδιορίζει τη σχέση μεταξύ ατόμων και Αρχών, αλλά αποτρέπει την αποδυνάμωση αυτής της σχέσης λόγω της τεχνολογικής εξουσιοδότησης. Η Πράξη ενσωματώνει εγγυήσεις για το κράτος δικαίου, τη διαφάνεια, την αμφισβητήσιμη φύση και την ανθρώπινη εποπτεία στη νομική αρχιτεκτονική που διέπει τη λήψη αλγοριθμικών αποφάσεων, διατηρώντας έτσι τις βασικές αρχές του «κοινωνικού συμβολαίου» υπό τις νέες τεχνολογικές συνθήκες.

Εάν η αυτοματοποιημένη λήψη αποφάσεων επιτρεπόταν να λειτουργεί χωρίς εγγυήσεις ισοδύναμες με εκείνες που επιβάλλονται στις δημόσιες Αρχές, ο κίνδυνος δεν θα ήταν απλώς ένα κανονιστικό κενό. Αντίθετα, θα δημιουργούσε μια κατάσταση στην οποία η εξουσία επί των ατόμων θα ασκούταν εκτός των συνταγματικών και νομικών πλαισίων που παραδοσιακά δικαιολογούν την εξουσία στις δημοκρατικές κοινωνίες. Οι εν λόγω κίνδυνοι προσεγγίζονται μέσα από τους άξονες της διακυβέρνησης, της διαφάνειας και της λογοδοσίας. Τα συστήματα ΤΝ μπορούν να επηρεάσουν βαθιά τη ζωή των ανθρώπων σε πρωτοφανή κλίμακα (για παράδειγμα στη Δικαιοσύνη, στην Υγεία ή την απασχόληση), ενώ λειτουργούν με τρόπους που είναι τεχνικά περίπλοκοι και κοινωνικά αδιαφανείς, υπονομεύοντας τις καθιερωμένες νόρμες νομιμότητας και μη διακριτικής μεταχείρισης.²⁷⁷

²⁷⁶ European Center for Not-for-Profit Law, & Danish Institute for Human Rights (2025). *A guide to fundamental rights impact assessments under the EU Artificial Intelligence (AI) Act*. Danish Institute for Human Rights. <https://www.humanrights.dk/publications/guide-fundamental-rights-impact-assessments-under-eu-artificial-intelligence-ai-act> European Union (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. Official Journal of the European Union, L. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>

²⁷⁷ O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing Group.

Η δημόσια εξουσία περιορίζεται από νόρμες που καθιστούν τις αποφάσεις γνωστές, αναθεωρήσιμες και υποκείμενες σε αμφισβήτηση. Ωστόσο, τα αλγοριθμικά συστήματα λειτουργούν συχνά ως «μαύρα κουτιά», με εσωτερική λογική που είναι είτε ιδιοσταγής είτε τόσο περίπλοκη, που ακόμη και οι ειδικοί δυσκολεύονται να την ερμηνεύσουν. Το έργο του Frank Pasquale σχετικά με την «κοινωνία του μαύρου κουτιού»²⁷⁸ υπογραμμίζει τον τρόπο με τον οποίο αδιαφανείς αλγόριθμοι μπορούν να ελέγχουν τα αποτελέσματα με τρόπους που διαφεύγουν του δημοκρατικού ελέγχου, ενδεχομένως διαιωνίζοντας συστημικά μειονεκτήματα χωρίς σαφή νομική λογοδοσία.

Ένας δεύτερος κίνδυνος είναι η διάχυση της ευθύνης. Στα παραδοσιακά διοικητικά συστήματα, συγκεκριμένα πρόσωπα ή όργανα (όπως αξιωματούχοι ή οργανισμοί) είναι υπεύθυνοι για τις αποφάσεις και μπορούν να κληθούν να λογοδοτήσουν εάν παραβιάσουν νομικά ή ηθικά πρότυπα. Αντίθετα, τα αυτοματοποιημένα συστήματα περιλαμβάνουν ένα δίκτυο σχεδιαστών, προγραμματιστών, υπευθύνων ανάπτυξης και προμηθευτών δεδομένων, γεγονός που καθιστά δύσκολο τον εντοπισμό της ευθύνης όταν οι αποφάσεις προκαλούν βλάβη. Υποστηρίζεται ότι η λογοδοσία στην ΤΝ δεν απαιτεί απαραίτητα νέα πλαίσια, αλλά προσαρμογή και βελτίωση με τρόπο ώστε να είναι ρητά δομημένη, σαφής και να εφαρμόζεται προσεκτικά στη μοναδική κοινωνικο-τεχνική φύση των συστημάτων ΤΝ.²⁷⁹

Η διάβρωση θεμελιωδών εγγυήσεων του κράτους δικαίου προκύπτει από την έλλειψη διαφάνειας και λογοδοσίας. Η νόμιμη διακυβέρνηση και το κράτος δικαίου απαιτούν τα άτομα να κατανοούν τις αποφάσεις που τα επηρεάζουν και να έχουν την ευκαιρία να τις αμφισβητήσουν. Χωρίς επαρκή εξήγηση και αιτιολόγηση των αποφάσεων της ΤΝ, η αντίληψη της νομιμότητας υποβαθμίζεται και η εμπιστοσύνη διαβρώνεται.²⁸⁰ Η διαφάνεια σε αυτό το πλαίσιο δεν είναι απλώς τεχνική, αλλά και κανονιστική, καθώς επηρεάζει τον τρόπο με τον οποίο οι άνθρωποι αντιλαμβάνονται τη δικαιοσύνη και τη νομιμότητα όταν τα συστήματα ΤΝ χρησιμοποιούνται σε δημόσιες λειτουργίες.

Ένα σχετικό ζήτημα είναι ότι πολλά συστήματα ΤΝ με σημαντικό κοινωνικό αντίκτυπο αναπτύσσονται και χρησιμοποιούνται από ιδιωτικούς φορείς, εκτός των παραδοσιακών μηχανισμών του Δημοσίου Δικαίου. Χωρίς διασφαλίσεις ισοδύναμες με τους περιορισμούς του δημόσιου τομέα, οι ιδιωτικοί φορείς λήψης αποφάσεων θα μπορούσαν να ασκήσουν σημαντική επιρροή στην πρόσβαση των πολιτών σε υπηρεσίες, ευκαιρίες ή δικαιώματα, χωρίς να

²⁷⁸ Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.

²⁷⁹ Novelli, C., Taddeo, M., & Floridi, L. (2024). Accountability in artificial intelligence: what it is and how it works. *AI & Society*, 39(4), 1871-1882. <https://doi.org/10.1007/s00146-023-01635-y>

²⁸⁰ Kouroutakis, A. (2025). Rule of law in the AI era: addressing accountability, and the digital divide. *Discover Artificial Intelligence*, 4, 115. <https://doi.org/10.1007/s44163-024-00191-8>

δεσμεύονται από αρχές όπως η αναλογικότητα, η μη διάκριση ή η διαδικαστική δικαιοσύνη.²⁸¹ Αυτό δημιουργεί μια μορφή ιδιωτικής διακυβέρνησης με δημόσιες επιπτώσεις, αλλά χωρίς δημοκρατική νομιμότητα ή δομές λογοδοσίας.

Σε ένα βαθύτερο κοινωνικο-νομικό επίπεδο, η απουσία ρύθμισης των αυτοματοποιημένων συστημάτων μπορεί να υπονομεύσει την ανθρώπινη δράση.²⁸² Η λήψη αποφάσεων με τη βοήθεια της ΤΝ βασίζεται συχνά σε στατιστικά συμπεράσματα και συσχετίσεις, αντί για τη συλλογιστική και την αιτιώδη συνάφεια που χαρακτηρίζουν την ανθρώπινη κρίση. Χωρίς επαρκείς διασφαλίσεις, τα άτομα κινδυνεύουν να αντιμετωπίζονται ως προφίλ δεδομένων και όχι ως αυτόνομοι πολίτες με αξιοπρέπεια και δικαιώματα, αποδυναμώνοντας έτσι το ηθικό υπόβαθρο της δημοκρατικής νομιμότητας και της ατομικής αυτονομίας.

Το πιο σημαντικό είναι ότι αυτή η αλλαγή θα μπορούσε να συμβεί χωρίς ρητή πολιτική επιλογή. Αντί να αποτελεί συνειδητή επαναδιαπραγμάτευση του νοήματος της διακυβέρνησης, θα μπορούσε να συμβεί σταδιακά, καθώς η κοινωνία υιοθετεί όλο και περισσότερα αυτοματοποιημένα συστήματα. Αυτό θα ισοδυναμούσε με μια *de facto* νέα κοινωνική τάξη, στην οποία σημαντικές αποφάσεις για τη ζωή των ανθρώπων λαμβάνονται από συστήματα που δεν δεσμεύονται από συνταγματικές αρχές διαφάνειας, λογοδοσίας και προστασίας των δικαιωμάτων. Αυτός είναι ο λόγος για τον οποίο η ενσωμάτωση εγγυήσεων, όπως η ανθρώπινη εποπτεία, η εξήγηση και η λογοδοσία, στη διακυβέρνηση της ΤΝ είναι απαραίτητη για τη διατήρηση της ακεραιότητας των δημοκρατικών συστημάτων στην εποχή της ΤΝ.

11.4.5. Αποτελεί αυτό μια μορφή ψηφιακού συνταγματισμού;

Η παραπάνω ανάλυση ανέδειξε δύο θέματα. Πρώτον, ότι η Πράξη για την Τεχνητή Νοημοσύνη επιδιώκει να αποτρέψει τα αυτοματοποιημένα συστήματα ΤΝ από το να ξεφύγουν από τις εγγυήσεις που επιβάλλει παραδοσιακά το «κοινωνικό συμβόλαιο» στην άσκηση εξουσίας. Δεύτερον, ότι η ΕΕ διεκδικεί μια μορφή ψηφιακής κυριαρχίας, καθορίζοντας τις κανονιστικές προϋποθέσεις υπό τις οποίες τα συστήματα ΤΝ μπορούν να λειτουργούν νόμιμα. Όταν συνδυάζονται αυτά τα δύο συμπεράσματα, το αποτέλεσμα δεν είναι απλώς η ρύθμιση της τεχνολογίας, αλλά ένα συνταγματικό σχέδιο προσαρμοσμένο στο ψηφιακό περιβάλλον.

Ο ψηφιακός συνταγματισμός, στην ουσία του, έχει ως στόχο να διασφαλίσει ότι οι συνταγματικές αρχές, το κράτος δικαίου, οι περιορισμοί της εξουσίας,

²⁸¹ Bloch-Wehba, H. (2019). Global Platform Governance: Private Power in the Shadow of the State. *SMU Law Review*, 72(1), 27-80.

²⁸² Kouroutakis, A. (2025). Rule of law in the AI era: addressing accountability, and the digital divide. *Discover Artificial Intelligence*, 4, 115. <https://doi.org/10.1007/s44163-024-00191-8>

η λογοδοσία, η προστασία των θεμελιωδών δικαιωμάτων και οι εγγυήσεις του κράτους δικαίου εν γένει συνεχίζουν να περιορίζουν την άσκηση της εξουσίας, όταν αυτή η εξουσία ασκείται μέσω ψηφιακών συστημάτων και όχι μέσω παραδοσιακών δημόσιων θεσμών.²⁸³ Αυτό είναι ακριβώς το πρόβλημα που εντοπίζει η ανάλυση του «κοινωνικού συμβολαίου»: Η ΤΝ κινδυνεύει να δημιουργήσει έναν χώρο όπου οι αποφάσεις που επηρεάζουν τα άτομα λαμβάνονται εκτός του νομικού και ηθικού πλαισίου που ιστορικά δικαιολογεί τη διακυβέρνηση. Η Πράξη για την Τεχνητή Νοημοσύνη ανταποκρίνεται, επιμένοντας ότι αυτό το πλαίσιο συνεχίζει να ισχύει ακόμη και όταν οι αποφάσεις λαμβάνονται από μηχανές.

Ταυτόχρονα, η ψηφιακή κυριαρχία δείχνει ότι η ΕΕ δεν ρυθμίζει μόνο την ΤΝ εντός της επικράτειάς της, αλλά καθορίζει και τα πρότυπα για τον τρόπο με τον οποίο μπορεί να ασκείται η αλγοριθμική εξουσία σε γενικότερο επίπεδο. Στο πλαίσιο αυτό, η κυριαρχία δεν εκδηλώνεται ως έλεγχος των υποδομών ή των επιχειρήσεων, αλλά ως έλεγχος της κανονιστικής αρχιτεκτονικής που διέπει την αυτοματοποιημένη λήψη αποφάσεων. Η προσέγγιση αυτή αποτελεί επιτομή του ψηφιακού συνταγματισμού: της θέσης, δηλαδή, ότι οι συνταγματικές αξίες οφείλουν να καθορίζουν τον σχεδιασμό και την ανάπτυξη των τεχνολογιών που δομούν την κοινωνική συμβίωση.

Αυτό που καθιστά την Πράξη για την Τεχνητή Νοημοσύνη εμβληματικό δείγμα αυτού του συνταγματισμού είναι, πρωτίστως, η ρυθμιστική τεχνική την οποία επιστρατεύει. Αντί να βασίζεται αποκλειστικά στα δικαστήρια, στους κανόνες ευθύνης ή στα εκ των υστέρων ένδικα μέσα, η Πράξη μεταφράζει τις συνταγματικές αρχές σε τεχνικές και οργανωτικές απαιτήσεις και προδιαγραφές. Υποχρεώσεις όπως η εκτίμηση κινδύνου, η τεκμηρίωση, η ικνηλασιμότητα, η ανθρώπινη εποπτεία και η διαφάνεια δεν είναι απλώς μηχανισμοί συμμόρφωσης, αλλά λειτουργούν ως αρχές όπως η αναλογικότητα, η λογοδοσία και οι εγγυήσεις του κράτους δικαίου, στο πλαίσιο του σχεδιασμού και της διακυβέρνησης των συστημάτων ΤΝ. Οι συνταγματικές εγγυήσεις ενσωματώνονται στην αρχιτεκτονική του συστήματος πριν από την ανάπτυξη.

Σε αυτό το μοντέλο, η συμμόρφωση δεν επαφίεται στην εκ των υστέρων εφαρμογή των νόμων (adjudication), αλλά επιβάλλεται προληπτικά μέσω του κώδικα, ο οποίος καθορίζει τα όρια του εφικτού πριν καν εκδηλωθεί μια πράξη. Το Δίκαιο δεν παύει να λειτουργεί ως εξωτερικός κριτής, αλλά επιπλέον μετατρέπεται σε δομικό υλικό της ψηφιακής πραγματικότητας. Αυτήν την εξέλιξη του Δικαίου από τη νομική διακυβέρνηση στην

²⁸³ Suzor, N. P. (2019). *Lawless: The secret rules that govern our digital lives*. Cambridge University Press· Padovani, C., & Santaniello, M. (2018). Digital constitutionalism: Fundamental rights and power limitation in the Internet eco-system. *International Communication Gazette*, 80(4), 295-301. <https://doi.org/10.1177/1748048518757114>· De Gregorio, G. (2022). *Digital constitutionalism in Europe: Reframing rights and powers in the algorithmic society*. Cambridge University Press· Celeste, E. (2023). *Digital constitutionalism: The role of internet bills of rights*. Routledge.

τεχνολογική διαχείριση περιγράφει ο Roger Brownsword ως «Law 3.0».²⁸⁴ Αυτό το μοντέλο της τεχνικής διαχείρισης απαντάται και σε άλλους τομείς, όπως η αυτοκινητοβιομηχανία και η αεροναυπηγική, όπου οι νομικοί κανόνες ασφαλείας ενσωματώνονται απευθείας στον τεχνικό σχεδιασμό.

Υπό αυτό το πρίσμα, η Πράξη για την Τεχνητή Νοημοσύνη μπορεί να θεωρηθεί ως το κανονιστικό επίπεδο που γεφυρώνει τον Χάρτη Θεμελιωδών Δικαιωμάτων της ΕΕ και τον GDPR. Ο Χάρτης θεσπίζει αφηρημένα δικαιώματα [μεταξύ άλλων, το δικαίωμα στην ανθρώπινη αξιοπρέπεια (άρθρο 1), στην προστασία δεδομένων προσωπικού χαρακτήρα (άρθρο 8), στη μη διάκριση (άρθρο 21) και στην αποτελεσματική δικαστική προστασία (άρθρο 47)], ο GDPR τα μεταφράζει σε υποχρεώσεις σχετικά με την επεξεργασία προσωπικών δεδομένων²⁸⁵ και η Πράξη για την Τεχνητή Νοημοσύνη επεκτείνει αυτήν τη μετάφραση στον ίδιο τον μηχανισμό λήψης αποφάσεων. Το αποτέλεσμα είναι μια μορφή πολυεπίπεδου συνταγματισμού, όπου τα θεμελιώδη δικαιώματα δεν προστατεύονται μόνο μέσω δικαστικής κρίσης, αλλά και μέσω του σχεδιασμού τεχνολογικών συστημάτων.

Με αυτόν τον τρόπο, η Πράξη για την Τεχνητή Νοημοσύνη δεν περιορίζεται στην οριζόντια επέκταση των θεμελιωδών δικαιωμάτων στους ιδιωτικούς φορείς, αν και το κάνει και αυτό. Πιο σημαντικά, διασφαλίζει ότι τα αλγοριθμικά συστήματα δεν μπορούν να γίνουν ζώνες όπου οι συνταγματικοί περιορισμοί παύουν να ισχύουν, ανεξάρτητα από το αν αναπτύσσονται από δημόσιες Αρχές ή ιδιωτικούς φορείς. Διατηρεί τη συνέχεια του «κοινωνικού συμβολαίου» και εκφράζει την ψηφιακή κυριαρχία της ΕΕ μέσω της συνταγματικοποίησης της διακυβέρνησης της ΤΝ.

Για τους λόγους αυτούς, η Πράξη για την Τεχνητή Νοημοσύνη μπορεί να θεωρηθεί ως μια συγκεκριμένη εκδήλωση του ψηφιακού συνταγματισμού: ένα νομικό πλαίσιο που προστατεύει τη συνταγματική τάξη, ενσωματώνοντας τις αρχές της απευθείας στις τεχνικές και οργανωτικές προϋποθέσεις υπό τις οποίες επιτρέπεται η αυτοματοποιημένη λήψη αποφάσεων.

²⁸⁴ Brownsword, R. (2022). *Rethinking law, regulation, and technology*. Edward Elgar Publishing. <https://doi.org/10.4337/9781800886476>

²⁸⁵ Mitrou, L. (2018). Data protection, artificial intelligence and cognitive services: Is the General Data Protection Regulation (GDPR) “artificial intelligence-proof”? *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3386914>

[ΠΩΣ] ΜΠΟΡΕΙ Η ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΝΑ ΕΙΝΑΙ
ΗΘΙΚΗ ΚΑΙ ΣΥΝΝΟΜΗ;
Ιούλιος 2026

Η μεγάλη εικόνα



Γιάννης Ασσαέλ, Δημήτρης-Φραγκίσκος Λέκκας,
Λίλιαν Μήτρου

Συγγραφική συνεισφορά ανά υποκεφάλαιο:

12.1. – Γιάννης Ασσαέλ

12.2., 12.4., 12.5., 12.6. – Λίλιαν Μήτρου

12.3. – Δημήτρης-Φραγκίσκος Λέκκας

Τα ηθικά και τα νομικά ζητήματα αφορούν εν τέλει την ίδια την απόφαση για τη χρήση της Τεχνητής Νοημοσύνης, την κατεύθυνση και την έκταση αυτής της χρήσης. Αφορούν τους παράγοντες που εμπλέκονται στην παραγωγή τεχνολογίας, τους παράγοντες που ελέγχουν τον σύννομο χαρακτήρα της ανάπτυξης και χρήσης τεχνολογίας και δεδομένων, αλλά και τα μέσα επιβολής κανόνων και αρχών.

Η συζήτηση για την κατεύθυνση της Κοινωνίας της Τεχνητής Νοημοσύνης δεν είναι εύκολη, ήδη εκ του γεγονότος ότι πολλοί αισθάνονται ότι η όποια απόπειρα ελέγχου έχει υπερκεραστεί από την πραγματικότητα. Μια τέτοια συζήτηση μπορεί να αποτελέσει μια ευκαιρία να επαναπροσδιορίσουμε τις προτεραιότητές μας, θέτοντας στο επίκεντρο την ανθρώπινη ευημερία και την ευδαιμονία, όπως γράφει ο Γιάννης Ασσαέλ στη συμβολή του στην παρούσα μελέτη, επιχειρώντας να αναδείξει τα ουσιώδη μέσα στον «θόρυβο».

12.1. Τεχνητή Νοημοσύνη: Διακυβέρνηση, ασφάλεια και κοινωνική συμπερίληψη

12.1.1. Το διεθνές τοπίο της Τεχνητής Νοημοσύνης

Ζούμε σε μια εποχή όπου η δημόσια συζήτηση για την Τεχνητή Νοημοσύνη συχνά ταλαντεύεται ανάμεσα σε δύο άκρα, τον άκριτο ενθουσιασμό και την απαισιοδοξία. Αναγνωρίζοντας τη βαρύτητα των προκλήσεων που αφορούν την ασφάλεια και τη διακυβέρνηση, καλούμαστε να διαχειριστούμε το τοπίο της ΤΝ με αίσθημα ευθύνης και συμπερίληψης. Είναι σημαντικό να διασφαλίσουμε ότι τα οφέλη αυτής της τεχνολογίας θα είναι προσβάσιμα σε όλους και ότι η ανάπτυξή της θα σέβεται τις θεμελιώδεις ανθρώπινες αξίες. Η ΤΝ δεν είναι απλώς ένα εργαλείο παραγωγής περιεχομένου, αλλά αποτελεί έναν θεμελιώδη επιταχυντή της επιστημονικής ανακάλυψης που μπορεί να αναβαθμίσει ουσιαστικά την ποιότητα της ζωής μας.

Η συνεισφορά της ΤΝ στις επιστήμες αποτελεί ίσως την πιο ελπιδοφόρα πτυχή αυτής της τεχνολογικής εξέλιξης. Στη Βιολογία, για παράδειγμα, η χρήση συστημάτων όπως το AlphaFold μας επέτρεψε να προβλέψουμε τη δομή των πρωτεϊνών, λύνοντας ένα πρόβλημα που απασχολούσε τους επιστήμονες για μισό αιώνα. Αυτό σημαίνει ότι μπορούμε πλέον να κατανοήσουμε καλύτερα τις ασθένειες και να σχεδιάσουμε νέα φάρμακα με ταχύτητα που φάνταζε αδιανόητη στο παρελθόν. Αντίστοιχα, στην Επιστήμη των Υλικών, η ΤΝ βοηθά στην ανακάλυψη νέων κρυσταλλικών δομών για αποδοτικότερες μπαταρίες, ενώ στον τομέα της Πυρηνικής Σύντηξης συμβάλλει στον έλεγχο του πλάσματος στους αντιδραστήρες. Ακόμη και στις Ανθρωπιστικές Επιστήμες, η τεχνολογία αυτή μας βοηθά να ανακτήσουμε και να διαβάσουμε κατεστραμμένα αρχαία κείμενα, αποδεικνύοντας ότι η τεχνολογία μπορεί να λειτουργήσει ως γέφυρα μεταξύ του μέλλοντος και της ιστορικής μας κληρονομιάς και υπενθυμίζοντάς μας πως οι Φυσικές και οι Ανθρωπιστικές Επιστήμες είναι άρρηκτα συνδεδεμένες.

Για να διακρίνουμε, λοιπόν, το σήμα μέσα στον «θόρυβο» και να χαράξουμε αποτελεσματική πολιτική, είναι σημαντικό να βασιστούμε σε δεδομένα. Εκθέσεις όπως το «Governing AI for Humanity» του ΟΗΕ, καθώς και μεγάλες έρευνες που καταγράφουν τις απόψεις χιλιάδων ειδικών, όπως η μελέτη των Πανεπιστημίων του Μπέρκλεϊ, της Οξφόρδης και της Βόννης «Thousands of AI Authors on the Future of AI», προσφέρουν πολύτιμη

γνώση. Οι ερευνητές επισημαίνουν τις τεράστιες δυνατότητες της ΤΝ, ενώ παράλληλα θέτουν το πλαίσιο για την αντιμετώπιση ζητημάτων κοινωνικής συνοχής και ασφάλειας. Είναι σημαντικό να αξιοποιήσουμε αυτήν τη γνώση και να την προσαρμόσουμε στην ελληνική πραγματικότητα, διασφαλίζοντας ότι η χώρα μας θα παραμείνει στην αιχμή της καινοτομίας με υπευθυνότητα.

12.1.2. Αξιολόγηση και ασφάλεια συστημάτων Τεχνητής Νοημοσύνης

Η διασφάλιση της ανθεκτικότητας και της αξιοπιστίας των συστημάτων ΤΝ αποτελεί κεντρικό πυλώνα του διεθνούς διαλόγου. Η πρόσφατη Διεθνής Επιστημονική Έκθεση για την Ασφάλεια της Προηγμένης ΤΝ (2026) τονίζει την ανάγκη για διεθνή συνεργασία στην αξιολόγηση των ικανοτήτων των μοντέλων αιχμής, επισημαίνοντας ότι η ασφάλεια δεν είναι μια στατική κατάσταση, αλλά μια δυναμική διαδικασία που απαιτεί συνεχή τεχνική επιτήρηση. Οι προβληματισμοί που εκφράζονται σε παγκόσμιο επίπεδο αφορούν πρακτικά ζητήματα ευθυγράμμισης των συστημάτων με τις ανθρώπινες αξίες. Στην περίπτωση της Ελλάδας, το ζήτημα της ασφάλειας λαμβάνει μια επιπλέον διάσταση λόγω της γλώσσας. Καθώς τα περισσότερα συστήματα ΤΝ έχουν εκπαιδευτεί πρωτίστως σε αγγλικό περιεχόμενο, δημιουργείται η ανάγκη για βαθύτερη κατανόηση της απόδοσής τους στην ελληνική γλώσσα και πολιτισμικό πλαίσιο.

Τα συστήματα ΤΝ εκπαιδεύονται σε μεγάλους όγκους δεδομένων, τα οποία λειτουργούν ως καθρέφτης της κοινωνίας. Ωστόσο, στην περίπτωση της ελληνικής γλώσσας, τα διαθέσιμα δεδομένα είναι περιορισμένα, γεγονός που εγκυμονεί κινδύνους μεροληψίας και σημαίνει ότι ο πολιτισμός και η κουλτούρα μας ενδέχεται να μην αποτυπώνονται με ακρίβεια. Επιπλέον, η απόδοση ενός συστήματος ΤΝ στο ίδιο πρόβλημα μπορεί να είναι διαφορετική αν διατυπωθεί στα ελληνικά αντί για τα αγγλικά. Καθίσταται, λοιπόν, σημαντική η δημιουργία αξιολογήσεων, ώστε να μετρήσουμε την απόδοση των συστημάτων ΤΝ στα ελληνικά. Επιπλέον, τίθεται ένα ζήτημα ασφαλείας, καθώς οι δικλίδες που σχεδιάστηκαν για την αποτροπή κακόβουλων χρήσεων στα αγγλικά δεν εγγυώνται αναγκαστικά την ίδια προστασία σε όλες τις γλώσσες. Δεν αρκεί να μεταφράσουμε τα ξένα ορόσημα (benchmarks), αλλά πρέπει να δημιουργήσουμε νέα, που να αντικατοπτρίζουν ουσιαστικά την ελληνική γλώσσα και τον πολιτισμό μας.

Για την αντιμετώπιση αυτών των σύνθετων προκλήσεων, κάποιες χώρες έχουν προσεγγίσει το πρόβλημα προχωρώντας στη δημιουργία εθνικών Ινστιτούτων Ασφάλειας Τεχνητής Νοημοσύνης (AI Safety Institutes – AISI). Αυτοί οι φορείς λειτουργούν ως κέντρα τεχνικής αριστείας για την αξιολόγηση και τον έλεγχο συστημάτων ΤΝ. Μέσω της στενής συνεργασίας με ομάδες της ακαδημαϊκής κοινότητας, αναπτύσσουν τις απαραίτητες αξιολογήσεις και θεσπίζουν πρωτόκολλα ελέγχου, διασφαλίζοντας ότι η ΤΝ που ενσωματώνεται στη δημόσια διοίκηση και την αγορά είναι όχι μόνο

τεχνολογικά προηγμένα, αλλά και κοινωνικά ασφαλής. Ακολουθώντας αυτήν την τάση, μια πιθανή λύση για τη θωράκιση των δικών μας γλωσσικών και πολιτισμικών αναγκών θα μπορούσε να είναι η δημιουργία ενός Ελληνικού Ινστιτούτου Ασφάλειας.

12.1.3. Εμπιστοσύνη στην πληροφορία: Είμαστε η τελευταία γενιά που πίστεψε σε οπτικοακουστικό περιεχόμενο;

Η ραγδαία εξέλιξη της Παραγωγικής ΤΝ έχει φέρει στο προσκήνιο έναν από τους εντονότερους προβληματισμούς της ερευνητικής κοινότητας παγκοσμίως: την εγκυρότητα του οπτικοακουστικού περιεχομένου. Η ευκολία με την οποία μπορούν πλέον να δημιουργηθούν πειστικά *deepfakes* έχει επηρεάσει την εμπιστοσύνη όλων μας, ενώ υπάρχουν αρκετά παραδείγματα που έχουν χρησιμοποιηθεί για να αποσταθεροποιήσουν τον δημόσιο διάλογο.

Σε αυτό το πλαίσιο, η ερευνητική κοινότητα εξετάζει λύσεις όπως η υδατογράφηση (*watermarking*), δηλαδή η ενσωμάτωση μη ορατών σημάτων στα παραγόμενα αρχεία που τα χαρακτηρίζουν ως τεχνητά. Ωστόσο, η προσέγγιση αυτή λειτουργεί συμπληρωματικά και δεν αποτελεί καθολική λύση. Αφενός, υπάρχει κατακερματισμός προτύπων μεταξύ των παρόχων συστημάτων ΤΝ και, αφετέρου, στα μοντέλα ανοιχτού κώδικα, οι δικλίδες αυτές μπορούν να παρακαμφθούν ή να αφαιρεθούν από κακόβουλους χρήστες.

Συνεπώς, αντί να προσπαθούμε να πιστοποιήσουμε τι είναι ψεύτικο, μια μάχη άνιση, είναι στρατηγικά προτιμότερο να εστιάσουμε στην πιστοποίηση του αυθεντικού περιεχομένου. Μια λύση προς αυτό είναι η Πιστοποίηση Προέλευσης Περιεχομένου (*Content Provenance*), όπως, π.χ., το C2PA, το οποίο προσφέρει μια λύση βασισμένη σε κρυπτογραφικές ψηφιακές υπογραφές, μια τεχνολογία που χρησιμοποιούμε ήδη καθημερινά για την ασφάλεια των ιστοσελίδων και την πιστοποίηση της ταυτότητας στα e-mails μας. Κατά τη στιγμή της δημιουργίας ενός αρχείου (π.χ., λήψη φωτογραφίας), δημιουργείται μια μοναδική ψηφιακή σύνοψη των δεδομένων, η οποία κρυπτογραφείται με το ιδιωτικό κλειδί της συσκευής. Αυτή η υπογραφή ενσωματώνεται στο αρχείο και το συνοδεύει. Οποιοσδήποτε χρήστης μπορεί στη συνέχεια να χρησιμοποιήσει το αντίστοιχο δημόσιο κλειδί για να επαληθεύσει ότι το περιεχόμενο προέρχεται από τη δηλωμένη πηγή και δεν έχει αλλοιωθεί.

Φυσικά, κάθε τεχνολογική λύση γεννά νέα ερωτήματα. Η χρήση ψηφιακών υπογραφών εγκυμονεί τον κίνδυνο διαρροής πληροφοριών για τον δημιουργό, γεγονός που μπορεί να είναι προβληματικό σε συνθήκες πολιτικής έντασης, ενώ τίθεται και το ζήτημα της διακυβέρνησης των Αρχών που εκδίδουν ή ανακαλούν τα ψηφιακά πιστοποιητικά. Ωστόσο, από επιστημονικής σκοπιάς, τα προβλήματα είναι για να λύνονται. Είναι ιδιαίτερα θετικό ότι η Ελλάδα συμμετέχει ενεργά σε αυτόν τον διεθνή διάλογο, παρακολουθώντας

στενά τις τεχνικές εξελίξεις και συνεισφέροντας στην υιοθέτηση στρατηγικών που θα μπορούσαν να επαναφέρουν την εμπιστοσύνη στο περιεχόμενο.

12.1.4. Ενδυνάμωση του οικοσυστήματος και εκπαίδευση

Τέλος, κομμάτι του διεθνούς διαλόγου είναι η σημασία της δημοκρατικοποίησης της πρόσβασης και της διασφάλισης της ψηφιακής κυριαρχίας. Σε ένα διαρκώς μεταβαλλόμενο γεωπολιτικό τοπίο, η ανάπτυξη ενός ενεργού και ανθεκτικού εγχώριου οικοσυστήματος καινοτομίας αποτελεί ανάγκη για κάθε σύγχρονο κράτος. Για να επιτευχθεί αυτό, απαιτείται μια ολιστική προσέγγιση που ξεκινά από τις υποδομές, περνά μέσα από την έρευνα και την επιχειρηματικότητα, και καταλήγει στον άνθρωπο και την εκπαίδευση.

Θεμέλιο αυτής της προσπάθειας αποτελούν οι ισχυρές υπολογιστικές υποδομές. Η ΕΕ πραγματοποιεί ήδη σημαντικά βήματα προς την ενίσχυση της τεχνολογικής της αυτονομίας με πρωτοβουλίες όπως τα «AI Factories». Παρόλο που ο ιδιωτικός τομέας διατηρεί σημαντικό προβάδισμα σε πόρους, η αξία των δημόσιων υποδομών είναι ανεκτίμητη, καθώς διασφαλίζουν τον δημοκρατικό και ανοικτό χαρακτήρα της έρευνας. Στο πλαίσιο αυτό, η δημιουργία του υπερυπολογιστή «Δαίδαλος» αποτελεί ένα εξαιρετικό επίτευγμα και ορόσημο για την ελληνική επιστημονική κοινότητα. Με την τεχνολογία να εξελίσσεται εκθετικά, η πρόβλεψη για νέες υποδομές είναι σημαντικό να αποτελεί μόνιμο κομμάτι του στρατηγικού σχεδιασμού, ώστε να διατηρηθεί η ανταγωνιστικότητα της εγχώριας έρευνας.

Παράλληλα, η Ελλάδα έχει θέσει ισχυρές βάσεις για την ανάπτυξη του οικοσυστήματός της, διαθέτοντας ένα ταχύτατα αναπτυσσόμενο περιβάλλον νεοφυούς επιχειρηματικότητας και προσελκύοντας ερευνητικά τμήματα (R&D hubs) κορυφαίων τεχνολογικών εταιρειών, μεταξύ άλλων Tesla, Samsung, Meta και Pfizer, τα οποία δημιουργούν θέσεις εργασίας υψηλής εξειδίκευσης. Προκειμένου να ενισχυθεί περαιτέρω αυτή η δυναμική, κρίνεται σκόπιμη η θέσπιση επιπρόσθετων κινήτρων για την προσέλκυση διεθνών κόμβων καινοτομίας, καθώς και για την προσέλκυση ή τον επαπατρισμό διεθνούς ταλέντου.

Για τη μεγιστοποίηση του οφέλους, ιδιαίτερης σημασίας είναι η διασύνδεση αυτών των πρωτοβουλιών με τα ελληνικά Ανώτατα Εκπαιδευτικά Ιδρύματα και τα Ερευνητικά Κέντρα. Η καλλιέργεια μιας κουλτούρας αμφίδρομης ροής ταλέντου και τεχνογνωσίας μεταξύ ακαδημαϊκού χώρου και βιομηχανίας επιτρέπει την έρευνα σε πραγματικά προβλήματα και τη διάχυση της γνώσης.

Τέλος, απαραίτητο συστατικό για τη βιωσιμότητα αυτού του οικοσυστήματος είναι η εκπαίδευση στην ορθή και αποδοτική χρήση των συστημάτων ΤΝ. Η πρώτη επαφή ξεκινά από την Πρωτοβάθμια Εκπαίδευση, με την καλλιέργεια της υπολογιστικής σκέψης όχι απλώς ως τεχνικής δεξιότητας,

αλλά ως εργαλείο για τη μελέτη και την κατανόηση όλων των επιστημονικών πεδίων. Παράλληλα, επεκτείνεται φυσιολογικά μέσα από τη διάβιου μάθηση. Η συνεχής επιμόρφωση βοηθά το εργατικό δυναμικό να προσαρμοστεί και να αξιοποιήσει την ΤΝ για την εξέλιξή του, με τελικό στόχο μια τεχνολογία που να είναι προσβάσιμη και χρήσιμη σε όλους.

12.1.5. Επίλογος

Είναι εξαιρετικά ενδιαφέρον το πλήθος των φιλοσοφικών ερωτημάτων που καλούμαστε να απαντήσουμε με αφορμή την τεχνολογική πρόοδο. Η συζήτηση αυτή μπορεί να αποτελέσει μια ευκαιρία να επαναπροσδιορίσουμε τις προτεραιότητές μας, θέτοντας στο επίκεντρο την ανθρώπινη ευημερία και την ευδαιμονία. Ο στρατηγικός αυτός στόχος οφείλει να κατευθύνει στην ανάπτυξη συστημάτων που, αντί να εστιάζουν στη μονοπώληση της ανθρώπινης προσοχής, θα συμβάλλουν ουσιαστικά στη βελτίωση της ποιότητας ζωής, στην ψυχική υγεία και την προσωπική ανάπτυξη. Προϋπόθεση για την επίτευξη αυτού του στόχου αποτελεί η υπεύθυνη, ασφαλής και έμπιστη ανάπτυξη και διακυβέρνηση της τεχνολογίας.

Η ΤΝ είναι ένα εργαλείο με πρωτοφανή μετασχηματιστική ισχύ και η χώρα μας έχει αποδείξει ότι διαθέτει τη βούληση και το ανθρώπινο δυναμικό να το αξιοποιήσει και να οικοδομήσει μια κοινωνία ανθεκτική στις προκλήσεις της νέας εποχής. Η Ελλάδα, αξιοποιώντας το διεθνώς αναγνωρισμένο ταλέντο της εντός και εκτός συνόρων, το πολιτισμικό της υπόβαθρο και τη σύγχρονη δυναμική της, έχει σημαντικό ρόλο να διαδραματίσει στον διεθνή διάλογο, προωθώντας ένα ανθρωποκεντρικό μοντέλο.

Βιβλιογραφία

- Bengio, Y., Clare, S., Prunkl, C., Andriushchenko, M., Bucknall, B., Murray, M., Bommasani, R., Casper, S., Davidson, T., Douglas, R., Duvenaud, D., Fox, P., Gohar, U., Hadshar, R., Ho, A., Hu, T., Jones, C., Kapoor, S., Kasirzadeh, A., ... Mindermann, S. (2026). *International AI Safety Report 2026*. <https://internationalaisafetyreport.org/publication/international-ai-safety-report-2026>
- Grace, K., Sandkühler, J. F., Stewart, H., Weinstein-Raun, B., Thomas, S., Stein-Perlman, Z., Salvatier, J., Brauner, J., & Korzekwa, R. C. (2025). Thousands of AI authors on the future of AI. *Journal of Artificial Intelligence Research*, 84, Article 9. <https://doi.org/10.1613/jair.1.19087>
- Coalition for Content Provenance and Authenticity (C2PA) (2025). *C2PA technical specification (version 2.3)*. https://c2pa.org/specifications/specifications/2.3/specs/C2PA_Specification.html
- Organisation for Economic Co-operation and Development (OECD) (2025). *Governing with artificial intelligence: The state of play and way forward in core government functions*. oecd.org/content/dam/oecd/en/publications/reports/2025/06/governing-with-artificial-intelligence_398fa287/795de142-en.pdf
- United Nations (UN) (2024). *Governing AI for humanity: Final report*. High-level Advisory Body on Artificial Intelligence. <https://www.un.org/en/ai-advisory-body>

12.2. Καινοτομία με κάθε κόστος; Και η κλιματική κρίση;

Η προσέγγιση πρέπει να είναι σταθμισμένη και ολιστική και να απαντά στο ερώτημα εάν επιδιώκουμε την ανάπτυξη της Τεχνητής Νοημοσύνης και την ανάπτυξη με την Τεχνητή Νοημοσύνη με κάθε κόστος.

Από περιβαλλοντικής σκοπιάς, η ανάπτυξη και η λειτουργία μεγάλων συστημάτων ΤΝ απαιτούν τεράστια υπολογιστική ισχύ, και άρα υψηλή κατανάλωση ηλεκτρικής ενέργειας. Η ΕΕ υπήρξε σαφής στη γενική στόχευση: Σύμφωνα με τις Κατευθυντήριες Γραμμές Δεοντολογίας για Αξιόπιστη ΤΝ, «η διαδικασία ανάπτυξης, εγκατάστασης και χρήσης του συστήματος, καθώς και ολόκληρη η αλυσίδα εφοδιασμού του, θα πρέπει να αξιολογείται σε αυτό το πλαίσιο, π.χ. μέσω κριτικής εξέτασης της χρήσης πόρων και της κατανάλωσης ενέργειας κατά τη διάρκεια της εκπαίδευσης, και να επιλέγονται λιγότερο επιβλαβείς επιλογές». Παρά τις διακηρύξεις, πληθαίνουν οι επισημάνσεις για την ενέργεια που απαιτείται για τη λειτουργία κέντρων δεδομένων, την εκπαίδευση των μοντέλων²⁸⁶ και εν γένει για τη χρήση της ΤΝ. Παράλληλα, θέτουν εκ νέου και μετ' επιτάσεως το ζήτημα της εξισορρόπησης μεταξύ των στόχων της καινοτομίας και της «βιώσιμης και φιλικής προς το περιβάλλον ΤΝ».

Πώς μπορούμε να ενσωματώσουμε βιώσιμα συστήματα ΤΝ μέσα σε μια Κυκλική Οικονομία; Η υιοθέτηση λύσεων ΤΝ θα πρέπει να σταθμίζεται με το οικονομικό και περιβαλλοντικό κόστος. Συνεπώς, ένα ερώτημα που πρέπει να θέτουμε είναι εάν υπάρχουν λύσεις εκτός ΤΝ που μπορούν να επιφέρουν συγκρίσιμο αποτέλεσμα χωρίς περιβαλλοντική επιβάρυνση. Με άλλα λόγια, να υπάρχει ένας υπεύθυνος και τεκμηριωμένος λόγος χρήσης ΤΝ. Η τεκμηρίωση θα πρέπει να συμπεριλαμβάνει και μια αξιολόγηση και εκτίμηση των περιβαλλοντικών επιπτώσεων.

²⁸⁶ Βλ. α.α. τη σχετική μελέτη του ΟΟΣΑ. Organisation for Economic Co-operation and Development (OECD). (2022). *Measuring the environmental impacts of artificial intelligence compute and applications: The AI footprint*. OECD Digital Economy Papers, No. 341. <https://doi.org/10.1787/7babf571-en> Ωστόσο, οι εκτιμήσεις αυτές αφορούν την περίοδο πριν από την εκρηκτική διάδοση της Παραγωγικής ΤΝ. Σε μια πρόσφατη (2025) έκθεση η Greenpeace Γερμανίας προειδοποίησε ότι η ζήτηση ηλεκτρικής ενέργειας για την ΤΝ, οι εκπομπές, η χρήση νερού και οι ανάγκες σε πρώτες ύλες αυξάνονται με ταχείς ρυθμούς, και ότι η ζήτηση ηλεκτρικής ενέργειας των κέντρων δεδομένων ΤΝ θα μπορούσε να είναι 11 φορές υψηλότερη το 2030 σε σχέση με το 2023, εκτός αν οι κυβερνήσεις παρέμβουν. Βλ. Greenpeace International. (2026, April 2). *The energy and environmental impact of AI and how it undermines democracy*. <https://www.greenpeace.org/international/story/82486/ai-energy-environment-democracy/>

Η σταθμισμένη και λελογισμένη χρήση αφορά και την Παραγωγική ΤΝ. Η προσφυγή του μεμονωμένου χρήστη, των εταιρειών, των δημόσιων υπηρεσιών στα Μεγάλα Γλωσσικά Μοντέλα συνεπάγεται σημαντικό περιβαλλοντικό κόστος. Αυτό που διαφοροποιεί την Παραγωγική ΤΝ είναι η πυκνότητα ισχύος που απαιτεί: «Μπορεί να καταναλώνει επτά ή οκτώ φορές περισσότερη ενέργεια από έναν τυπικό υπολογιστικό φόρτο εργασίας», επισημαίνει ο Noman Bashir, προσθέτοντας ότι ο καθημερινός χρήστης δεν το σκέφτεται ιδιαίτερα αυτό, καθώς η ευκολία χρήσης της Παραγωγικής ΤΝ και η έλλειψη πληροφόρησης σχετικά με τις περιβαλλοντικές επιπτώσεις των ενεργειών του «σημαίνουν ότι, ως χρήστης, δεν έχω ιδιαίτερο κίνητρο να περιορίσω τη χρήση της παραγωγικής ΤΝ».²⁸⁷

Η αντιμετώπιση των αυξημένων απαιτήσεων ενέργειας και η ενεργειακή επιβάρυνση δεν θέτουν μόνο το ζήτημα της προστασίας του περιβάλλοντος και της λελογισμένης χρήσης πόρων σε μια εποχή κλιματικής κρίσης. Τίθενται ηθικά και εν γένει δικαιιοπολιτικά ζητήματα που αφορούν την περιβαλλοντική και καταναλωτική δικαιοσύνη και εν γένει την ισότητα.

Το κόστος ανάπτυξης και χρήσης της ΤΝ αυξάνεται. Η αυξημένη κατανάλωση ενέργειας που απαιτείται για την εκπαίδευση και τη λειτουργία προηγμένων συστημάτων ΤΝ συνδέεται άμεσα με ζητήματα ισότητας και δικαιοσύνης. Οι υψηλές ενεργειακές και κατ' αποτέλεσμα οικονομικές απαιτήσεις περιορίζουν καταρχάς τη δυνατότητα ανάπτυξης και αξιοποίησης τέτοιων συστημάτων κυρίως σε ισχυρές εταιρείες, εύρωστα ερευνητικά κέντρα και οικονομικά ανεπτυγμένα κράτη. Υπάρχει συνεπώς ο κίνδυνος της περαιτέρω διεύρυνσης κοινωνικών, οικονομικών, γνωστικών και γεωπολιτικών ανισοτήτων. Χώρες, οργανισμοί, μικρότερες επιχειρήσεις και κοινωνικές ομάδες με περιορισμένους πόρους αποκλείονται από την ισότιμη συμμετοχή στην παραγωγή και αξιοποίηση αυτής της τεχνολογίας, γεγονός που εντείνει το ήδη υφιστάμενο ψηφιακό χάσμα.

Ένα περαιτέρω ζήτημα αφορά την άνιση κατανομή της ενεργειακής επιβάρυνσης. Οι περιβαλλοντικές επιπτώσεις, π.χ. μέσω της εγκατάστασης ενεργοβόρων υποδομών, της κατανάλωσης φυσικών πόρων ή της επιβάρυνσης ενεργειακών δικτύων, μπορεί να επιβαρύνουν δυσανάλογα τοπικές κοινωνίες ή και χώρες, οι οποίες δεν απολαμβάνουν κατ' ανάγκην και τα αντίστοιχα οφέλη. Τα ζητήματα που εγείρονται αφορούν, εν τέλει, το ερώτημα: Ποιοι ωφελούνται από την τεχνολογική πρόοδο και ποιοι επωμίζονται το κόστος της;

Μπορεί να υπάρξει μια περιβαλλοντική διακυβέρνηση της ΤΝ; Το ζήτημα αυτό συζητά ο Δημήτρης-Φραγκίσκος Λέκκας στη συμβολή του στην παρούσα μελέτη.

²⁸⁷ Zewe, A. (2025, January 17). Explained: Generative AI's environmental impact. MIT News, Massachusetts Institute of Technology. <https://news.mit.edu/2025/explained-generative-ai-environmental-impact-0117>

12.3. Τεχνητή Νοημοσύνη και Κυκλική Οικονομία: Ψηφιακός μετασχηματισμός της περιβαλλοντικής διακυβέρνησης στην Ευρώπη και τη Μεσόγειο

Ο 21ος αιώνας χαρακτηρίζεται από μια σύνθετη περιβαλλοντική κρίση, η οποία δεν περιορίζεται στην κλιματική αλλαγή και την αύξηση της μέσης παγκόσμιας θερμοκρασίας, αλλά περιλαμβάνει, μεταξύ άλλων, την επιτάχυνση ακραίων καιρικών φαινομένων, την απώλεια βιοποικιλότητας, την εξάντληση φυσικών πόρων, τη ρύπανση υδάτων και εδαφών. Η Μεσόγειος αποτελεί χαρακτηριστική περίπτωση περιοχής υψηλής περιβαλλοντικής πίεσης, καθώς παρουσιάζεται έντονη τουριστική δραστηριότητα και σημαντική αστικοποίηση παράκτιων ζωνών, με αποτέλεσμα να καταγράφονται σημαντικές ποσότητες θαλάσσιων απορριμμάτων και πλαστική ρύπανση, δημιουργώντας σημαντικές περιβαλλοντικές πιέσεις.

Η κρίση αυτή συνδέεται δομικά με το γραμμικό μοντέλο οικονομίας («take-make-dispose»), το οποίο βασίζεται στην εξαγωγή πρώτων υλών, στην κατανάλωση και στην τελική διάθεση, και η αντιμετώπισή της απαιτεί έναν συστημικό μετασχηματισμό των μοντέλων παραγωγής και κατανάλωσης. Η Ευρωπαϊκή Πράσινη Συμφωνία και η Ευρωπαϊκή Στρατηγική για την Κυκλική Οικονομία προτείνουν μια ουσιαστική αλλαγή και βασίζονται σε ένα αναγεννητικό οικονομικό μοντέλο που στοχεύει στον σεβασμό των φυσικών πόρων και την πρόληψη της δημιουργίας αποβλήτων. Το θεωρητικό πλαίσιο της Κυκλικής Οικονομίας (Circular Economy – CE), όπως παρουσιάζεται από το Ellen MacArthur Foundation, βασίζεται σε τρεις θεμελιώδεις αρχές:

- α) απαλοιφή των αποβλήτων και της ρύπανσης,
- β) διατήρηση προϊόντων και υλικών σε χρήση και
- γ) αναγέννηση φυσικών συστημάτων.

Παράλληλα, η Ευρωπαϊκή Πράσινη Συμφωνία θέτει ως στρατηγικό στόχο την επίτευξη κλιματικής ουδετερότητας έως το 2050, ενσωματώνοντας το Circular Economy Action Plan ως βασικό πυλώνα. Η Πράσινη Συμφωνία αναγνωρίζει ρητά ότι η πράσινη μετάβαση πρέπει να συνοδεύεται από ψηφιακό μετασχηματισμό. Σε αυτό το πλαίσιο, η Τεχνητή Νοημοσύνη αναδεικνύεται σε κρίσιμο εργαλείο συστημικής διαχείρισης της πολυπλοκότητας

που χαρακτηρίζει τόσο τα περιβαλλοντικά προβλήματα όσο και την εφαρμογή της Κυκλικής Οικονομίας.

Η Κυκλική Οικονομία είναι εγγενώς ένα πολυδιάστατο σύστημα, το οποίο ενσωματώνει τη βιωσιμότητα σε όλα τα επίπεδα: περιβάλλον, κοινωνία και οικονομία. Απαιτεί την παρακολούθηση πολυεπίπεδων ροών υλικών, τον συντονισμό της παραγωγής και της κατανάλωσης προϊόντων, αλλά και της χρήσης φυσικών πόρων, την πρόβλεψη μελλοντικών ροών πόρων, τον συντονισμό των ρυθμιστικών Αρχών, την ανάλυση κύκλου ζωής προϊόντων και τη διαχείριση αβεβαιότητας σε διάφορες διαστάσεις. Τα παραπάνω χαρακτηριστικά καθιστούν την εφαρμογή της CE εξαιρετικά απαιτητική και οριακά ανέφικτη χωρίς ψηφιακά εργαλεία υψηλής υπολογιστικής ισχύος.

12.3.1. Διαχείριση πολυπλοκότητας και μη γραμμικών συστημάτων

Τα συστήματα και οι διαδικασίες για την υποστήριξη του κυκλικού μοντέλου περιλαμβάνουν εξ ορισμού πολλαπλούς δρώντες (multi-actor systems/approaches), ροές υλικών διαφορετικής ποιότητας, χρονικές υστερήσεις και αβεβαιότητα. Με διαφορετικό ρυθμό ανακυκλώνεται και ανακτάται, για παράδειγμα, το αλουμίνιο από μια συσκευασία και με διαφορετικό από ένα όχημα. Η παραδοσιακή γραμμική ανάλυση αδυνατεί να αποδώσει πλήρως αυτήν την πολυπλοκότητα. Τα μοντέλα TN μπορούν να υποστηρίξουν την πολυπλοκότητα που εμπεριέχεται, καθώς με την εν γένει δυναμική δομή τους έχουν τη δυνατότητα να αναλύουν μεγάλες βάσεις δεδομένων (Big Data), να μοντελοποιούν πολύπλοκες σχέσεις μεταξύ μεταβλητών, να προβλέπουν μελλοντικές ροές υλικών και να υποστηρίζουν προγνωστικά μοντέλα.

12.3.2. Χαρτογράφηση ροών υλικών και πόρων

Η κυκλικότητα προϋποθέτει γνώση του «πού βρίσκεται τι». Οι νέες τεχνολογίες, σε συνδυασμό με Διαδίκτυο των Πραγμάτων (Internet of Things – IoT) και ψηφιακά διαβατήρια προϊόντων (Digital Product Passports), επιτρέπουν την ικνηλασιμότητα σε πραγματικό χρόνο, τη βελτιστοποίηση logistics, τον εντοπισμό σημείων απώλειας αξίας ή και σημείων δημιουργίας αξίας, καθώς παρέχονται δυνατότητες αξιοποίησης πληροφοριών που δεν μπορούσαν να προσδιοριστούν και να αποτυπωθούν μέχρι τώρα. Κατ' επέκταση, δημιουργούνται οι προϋποθέσεις για τη διατήρηση αξίας των υλικών εντός του συστήματος. Η ΕΕ, στο πλαίσιο του Circular Economy Action Plan και του Ecodesign for Sustainable Products Regulation (ESPR), εισάγει το Digital Product Passport (DPP), που περιλαμβάνει πληροφορίες σύνθεσης υλικών, καταγράφει την επισκευασιμότητα, παρακολουθεί το περιβαλλοντικό αποτύπωμα και υποστηρίζει την ανάκτηση υλικών.

12.3.3. Προληπτική συντήρηση και επιμήκυνση κύκλου ζωής

Η προληπτική συντήρηση (predictive maintenance – PdM) αποτελεί μία από τις πλέον ώριμες και αποδεδειγμένες εφαρμογές της ΤΝ στη βιομηχανία και τις υποδομές. Σε αντίθεση με τη διορθωτική (reactive) ή την προγραμματισμένη (preventive) συντήρηση, η προληπτική συντήρηση βασίζεται σε συνεχή συλλογή δεδομένων μέσω αισθητήρων (IoT), ανάλυση μεγάλων δεδομένων (Big Data), αλγορίθμους μηχανικής μάθησης, προγνωστικά μοντέλα φθοράς και αστοχίας.

Η προληπτική συντήρηση με χρήση ΤΝ δεν αποτελεί απλώς τεχνική βελτίωση λειτουργίας. Αποτελεί δομικό μηχανισμό Κυκλικής Οικονομίας, καθώς συμβάλλει άμεσα στην επιμήκυνση της διάρκειας ζωής προϊόντων και εξοπλισμού, στη μείωση πρόωρης αντικατάστασης, σε σημαντική μείωση αποβλήτων και κατανάλωση πρώτων υλών και ενισχύει την εφαρμογή μοντέλων product-as-a-service, που είναι δομικό στοιχείο στο πλαίσιο της Κυκλικής Οικονομίας. Συνεπώς, η προληπτική συντήρηση είναι ένας από τους πιο ώριμους και άμεσα εφαρμόσιμους πυλώνες AI-enabled circular economy.

12.3.4. Βελτιστοποίηση σχεδιασμού προϊόντων (Eco-design & Circular Design)

Η μετάβαση στην Κυκλική Οικονομία δεν μπορεί να επιτευχθεί μόνο μέσω της διαχείρισης και αξιοποίησης των αποβλήτων. Το 70–80% των περιβαλλοντικών επιπτώσεων ενός προϊόντος προκαθορίζεται στο στάδιο του σχεδιασμού του. Συνεπώς, η ουσιαστική κυκλικότητα απαιτεί μετατόπιση από τη διαχείριση τέλους κύκλου ζωής (end-of-life management) προς τον σχεδιασμό πρόληψης επιπτώσεων (design for circularity).

Η ΤΝ ενσωματώνεται πλέον στη φάση σχεδιασμού προϊόντων, προσφέροντας εργαλεία προγνωστικής ανάλυσης, βελτιστοποίησης υλικών και μοντελοποίησης εναλλακτικών σεναρίων κύκλου ζωής. Η παραδοσιακή Ανάλυση Κύκλου Ζωής (Life Cycle Assessment – LCA) βασίζεται σε στατικά δεδομένα και συντελεστές εκπομπών στα στάδια ζωής ενός προϊόντος (from cradle to grave). Η ΤΝ μετατρέπει την LCA σε δυναμικό εργαλείο σχεδιασμού, αξιοποιώντας και αναλύοντας μεγάλες βάσεις δεδομένων περιβαλλοντικών επιπτώσεων (Ecoinvent, GaBi κλπ.), προβλέποντας μελλοντικά επίπεδα εκπομπών με βάση διαφορετικά σενάρια χρήσης, προσομοιώνοντας εναλλακτικά υλικά και σχεδιαστικές επιλογές.

Η ΤΝ επιτρέπει ανάλυση ιδιοτήτων υλικών σε μεγάλη κλίμακα και σύγκριση περιβαλλοντικών δεικτών, αξιοποιώντας τεχνολογικές δυνατότητες όπως machine learning για ανακάλυψη νέων βιώσιμων υλικών, ανάλυση trade-offs μεταξύ ανθεκτικότητας και ανακυκλωσιμότητας, βελτιστοποίηση συνδυασμού υλικών. Η βελτιστοποίηση σχεδιασμού προϊόντων μέσω ΤΝ μεταφέρει

την κυκλικότητα από το τέλος του κύκλου ζωής στην αρχή του, ενισχύει την αποδοτικότητα πόρων, μειώνει τον ενσωματωμένο άνθρακα και δύναται να υποστηρίξει νέα επιχειρηματικά μοντέλα.

Η ΤΝ δεν λειτουργεί απλώς ως εργαλείο ανάλυσης, αλλά ως μηχανισμός ανασχεδιασμού του ίδιου του παραγωγικού συστήματος. Η πρόκληση δεν είναι εάν η ΤΝ θα ενσωματωθεί στην Κυκλική Οικονομία, αλλά πώς θα ενσωματωθεί με τρόπο βιώσιμο, διαφανή και κοινωνικά δίκαιο, σε ευθυγράμμιση με τους στόχους της Ευρωπαϊκής Πράσινης Συμφωνίας και των περιφερειακών περιβαλλοντικών πλαισίων, όπως το Πρόγραμμα των Ηνωμένων Εθνών για το Περιβάλλον/Μεσογειακό Σχέδιο Δράσης (UNEP/MAP).

Παρά το σημαντικό μετασχηματιστικό δυναμικό της ΤΝ για την Κυκλική Οικονομία, η διεθνής επιστημονική βιβλιογραφία υπογραμμίζει ότι η ενσωμάτωσή της συνοδεύεται από κρίσιμους κινδύνους και θεσμικές προκλήσεις. Οι προκλήσεις αυτές δεν είναι απλώς τεχνολογικές, αλλά αφορούν ζητήματα ενεργειακής βιωσιμότητας, διαφάνειας, διακυβέρνησης δεδομένων, κοινωνικής ισότητας και κανονιστικής εποπτείας.

Ένα από τα πλέον συζητημένα ζητήματα αφορά το ενεργειακό αποτύπωμα των συστημάτων ΤΝ. Η εκπαίδευση και η χρήση μεγάλων μοντέλων βαθιάς μάθησης (deep learning) απαιτεί σημαντική υπολογιστική ισχύ, η οποία μεταφράζεται σε αυξημένη κατανάλωση ηλεκτρικής ενέργειας και εκπομπές CO₂.

Η εκπαίδευση Μεγάλων Γλωσσικών Μοντέλων έχει πολύ μεγάλο αποτύπωμα άνθρακα. Λαμβάνοντας υπ' όψιν αυτήν τη διάσταση, έχει προταθεί η έννοια «Green AI», η οποία υποστηρίζει ότι η αποδοτικότητα και η ενεργειακή κατανάλωση πρέπει να αποτελούν κριτήρια αξιολόγησης ερευνητικών αποτελεσμάτων στην ΤΝ. Σε ένα πλαίσιο Κυκλικής Οικονομίας, όπου ο στόχος είναι η μείωση εκπομπών και η αποδοτική χρήση πόρων, η αντίφαση αυτή είναι εμφανής: Η ΤΝ μπορεί να βελτιστοποιεί τη χρήση υλικών, αλλά ταυτόχρονα να αυξάνει τη ζήτηση ενέργειας μέσω data centers. Η αντιμετώπιση αυτής της πρόκλησης απαιτεί χρήση Αναnewώσιμων Πηγών Ενέργειας για υπολογιστικές υποδομές, βελτιστοποίηση αλγορίθμων ως προς ενεργειακή αποδοτικότητα και ανάπτυξη δεικτών «AI sustainability performance».

12.4. Η Τεχνητή Νοημοσύνη πήρε το όπλο της;

Στο πλαίσιο της συζήτησης για τις μείζονος χαρακτήρα επιπτώσεις της χρήσης της Τεχνητής Νοημοσύνης, δεν θα μπορούσαμε παρά να συμπεριλάβουμε ένα ζήτημα με σημαντικές ηθικές και νομικές διαστάσεις, που, αν και δεν είναι νέο, επανήλθε στο προσκήνιο: το ζήτημα της χρήσης συστημάτων ΤΝ σε πολεμικές επιχειρήσεις. Όπως σημειώνεται στο Editorial του Nature,²⁸⁸ οι επιθέσεις στο Ιράν²⁸⁹ μάς υπενθύμισαν πόσο κοντά βρίσκεται η έρευνα στην ΤΝ στην πρώτη γραμμή του πολέμου.

Χωρίς αμφιβολία, τα αυτόνομα όπλα (drones, σε συνδυασμό με συστήματα ΤΝ) μπορούν να προσφέρουν μεγάλο τακτικό πλεονέκτημα, καθώς είναι οι «μηχανές» που εμπλέκονται με τον «εχθρό», ενώ οι άνθρωποι-στρατιώτες παραμένουν ασφαλείς. Τα άλλα σημαντικά «πλεονεκτήματα», η αστραπιαία ταχύτητα και η ακρίβεια, συνιστούν, ωστόσο, ακριβώς εκείνες τις χρήσεις που εγείρουν σοβαρά ηθικά και δικαιολογικά ερωτήματα.

Καταρχάς, τίθεται ένα ζήτημα αξιοπιστίας, το οποίο, όμως, έχει ουσιαστικές, αλλά και, όπως βάσιμα εικάζεται, τραγικές συνέπειες. Το Nature, επικαλούμενο ερευνητές από την OpenAI και την Google, επισημαίνει ότι ακόμη και οι πιο προηγμένες τεχνολογίες, γνωστές ως μοντέλα Frontier AI, δεν είναι ακόμη ικανές να λειτουργούν με αξιοπιστία ή εντός του υφιστάμενου Δικαίου του πολέμου.

Η ενσωμάτωση οπλικών συστημάτων που υποστηρίζονται από ΤΝ διευκολύνει την αντικειμενοποίηση των ανθρώπινων στόχων, οδηγώντας σε αυξημένη ανοχή απέναντι στις παράπλευρες απώλειες. Μπορούμε να αποδεχούμε τα ελλείμματα που έχουν εντοπιστεί σε σχέση με χρήσεις ΤΝ σε άλλα πεδία, όπως τα σφάλματα, οι προκαταλήψεις, τα κενά ελέγχου και διακυβέρνησης, και στις πολεμικές χρήσεις αυτών των τεχνολογιών; Ποιο επίπεδο αυτοματισμού (άρα και σφάλματος) και ανάθεσης αποφάσεων

²⁸⁸ Nature Editorial. (2026, March 10). Stop the use of AI in war until laws can be agreed. *Nature*, 651 (8105), 282. <https://doi.org/10.1038/d41586-026-00762-y>

²⁸⁹ Η πρώτη καταγεγραμμένη χρήση πλήρως αυτόνομων οπλικών συστημάτων χρονολογείται από το 2021 στη Λιβύη. Κρίσιμη εξέλιξη υπήρξε ο πόλεμος στην Ουκρανία, όπου και οι δύο πλευρές έχουν χρησιμοποιήσει αυτόνομα οπλικά συστήματα, ενώ ανέπτυξαν και δοκίμασαν νέα.

σε αυτόνομα συστήματα ΤΝ είμαστε εν προκειμένω διατεθειμένοι να αποδεχτούμε ως κοινωνίες;

Είναι επαρκές να θέσουμε απλά το αίτημα του ανθρώπινου ελέγχου ή της ανθρώπινης παρέμβασης που θα διορθώσει το σφάλμα ή θα αναιρέσει την ενέργεια; Μπορεί πράγματι στις δεδομένες συνθήκες πολλαπλής και πολυδιάστατης χρονικής και ψυχικής πίεσης²⁹⁰ να υπάρξει ουσιαστικός ανθρώπινος έλεγχος, ώστε να απενεργοποιηθούν τέτοια συστήματα αν συνειδητοποιήσουμε το λάθος; Έχουν και σε ποιον βαθμό τα πρόσωπα που είναι επιφορτισμένα με την επίβλεψη τέτοιων συστημάτων τον χρόνο και την αυτονομία να παρέμβουν; Η Elke Shwartz επισημαίνει ότι η προκατάληψη αυτοματισμού και η τεχνολογική διαμεσολάβηση αποδυναμώνουν την ηθική αυτενέργεια των χειριστών συστημάτων στόχευσης με δυνατότητες ΤΝ, μειώνοντας την ικανότητά τους για ηθική λήψη αποφάσεων.²⁹¹

Η Elke Schwarz θέτει και ένα άλλο εξόχως ηθικό ζήτημα: Η ενσωμάτωση οπλικών συστημάτων που υποστηρίζονται από ΤΝ διευκολύνει την αντικειμενοποίηση των ανθρώπινων στόχων, οδηγώντας σε αυξημένη ανοχή απέναντι στις παράπλευρες απώλειες. Χωρίς αμφιβολία, οι ηθικές αρχές και οι κανόνες διεξαγωγής ενός πολέμου ισχύουν και για τη χρήση τέτοιων όπλων. Είναι, όμως, οι κανόνες αυτοί επαρκείς για να αντιμετωπίσουν αυτά τα καινοφανή ζητήματα; Ένα κρίσιμο πεδίο είναι αυτό του εντοπισμού και της απόδοσης της ευθύνης. Λόγω της εγγενούς αδιαφάνειας, της περιορισμένης προβλεψιμότητας και της ελλειμματικής επεξηγησιμότητας, αλλά και εξαιτίας του κατανεμημένου τρόπου με τον οποίο αναπτύσσονται αυτά τα ζητήματα, δεν είναι ευχερής η ανασύνθεση της αλυσίδας αποφάσεων και ενεργειών που οδήγησαν σε καταστροφικά αποτελέσματα, ακούσια ή εκούσια.

Ακόμη κι αν ο νομοθέτης, ευλόγως και για λόγους αρμοδιότητας, εξαιρεί από τις συνήθεις ρυθμιστικές πρωτοβουλίες την εθνική άμυνα,²⁹² θα πρέπει σε επίπεδο διεθνούς κοινότητας να τεθεί το ζήτημα της ανεκτής χρήσης των αυτόνομων όπλων, αλλά και τα ζητήματα της σχετικής ευθύνης. Εν προκειμένω, θα πρέπει να ενεργοποιηθεί τόσο η επιστημονική κοινότητα όσο και η Κοινωνία των Πολιτών.

²⁹⁰ Αναφέρει χαρακτηριστικά το Editorial του Nature: «Κάθε δευτερόλεπτο μετρά στη δίνη της μάχης, και όσοι αναπτύσσουν αυτήν την τεχνολογία θεωρούν ότι έχει τη δυνατότητα να πλήττει ταχύτερα ή να ανταποκρίνεται στα εισερχόμενα πυρά πιο γρήγορα απ' ό,τι οι άνθρωποι, με ελάχιστη ανάγκη άμεσης επικοινωνίας με τα κέντρα διοίκησης».

²⁹¹ Schwarz, E. (2018). *Death Machines: The Ethics of Violent Technologies* (pp. 132, 155 ff.). Manchester University Press.

²⁹² Σύμφωνα με το άρθρο 2 παρ. 3 της AI Act, αυτή δεν εφαρμόζεται σε συστήματα ΤΝ τα οποία δεν διατίθενται στην αγορά, ούτε τίθενται σε λειτουργία στην Ένωση, όταν τα «στοιχεία εξόδου» χρησιμοποιούνται στην Ένωση αποκλειστικά για στρατιωτικούς και αμυντικούς σκοπούς ή σκοπούς εθνικής ασφάλειας, ανεξάρτητα από το είδος της οντότητας που εκτελεί τις εν λόγω δραστηριότητες.

12.5. Είναι θέμα σχεδιασμού;

Ένα άλλο μείζον διακύβευμα που στη φάση αυτή καλείται να αντιμετωπίσει τόσο η Πολιτεία όσο και η κοινωνία είναι τα προβλήματα χειραγώγησης που ανακύπτουν από τις εφαρμογές Τεχνητής Νοημοσύνης, σε συνδυασμό με τον «εθιστικό σχεδιασμό» (addictive design), που χαρακτηρίζεται μεταξύ των άλλων από πρακτικές όπως το infinite scroll (αδιάκοπη κύλιση), το autoplay (αυτόματη αναπαραγωγή), τα push notifications, εξατομικευμένα recommender systems (συστήματα συστάσεων), την καλλιέργεια του λεγόμενου FOMO (Fear of Missing Out – «φόβος της παράλειψης»). Ήδη το 2023 το Ευρωπαϊκό Κοινοβούλιο είχε επισημάνει ότι στη σημερινή οικονομία της προσοχής, ορισμένες τεχνολογικές εταιρείες χρησιμοποιούν τον σχεδιασμό και τη λειτουργικότητα των συστημάτων για να εκμεταλλεύονται την ευαλωτότητα των χρηστών και των καταναλωτών, με στόχο να αποσπούν την προσοχή τους και να μεγιστοποιούν τον χρόνο που περνούν στις ψηφιακές πλατφόρμες και τα χρήματα που δαπανούν εκεί. Το κρίσιμο ζήτημα, που συνιστά πλέον και αντικείμενο διερεύνησης από την ΕΕ,²⁹³ αλλά και αντικείμενο δικαστικών διαφορών, είναι ότι ο «εθιστικός σχεδιασμός» μπορεί να προκαλέσει ουσιώδη βλάβη στους καταναλωτές και να είναι επιζήμιος για τη σωματική και ψυχολογική τους υγεία.

Λόγω της εγγενώς διεθνοποιημένης φύσης των πλατφορμών, υποδομών, συστημάτων και μοντέλων, οι σχετικές απαντήσεις είναι, ωστόσο, αμφίβολο εάν μπορούν να δοθούν σε επίπεδο μεμονωμένου κράτους. Η Πράξη για την Τεχνητή Νοημοσύνη, μέσω της κατάταξης των μεθόδων χειραγώγησης της συμπεριφοράς στις απαγορευμένες πρακτικές, προσφέρει μια καταρχήν ασπίδα προστασίας. Η προστασία αυτή θα πρέπει να ενισχυθεί μέσω ρυθμίσεων που αφορούν τις αθέμιτες εμπορικές πρακτικές και τα δικαιώματα των καταναλωτών. Περαιτέρω, είναι αναγκαίο να προωθηθούν και να εφαρμοστούν πολιτικές πρωτοβουλίες και βιομηχανικά πρότυπα για την ασφάλεια ήδη από τον σχεδιασμό σε ψηφιακές υπηρεσίες και προϊόντα ιδίως για ανηλίκους-παιδιά, τα οποία είναι, όπως έχει αποδειχθεί, ιδιαίτερα ευάλωτα.

²⁹³ Ήδη από το 2024 η Ευρωπαϊκή Επιτροπή διενεργεί ελέγχους σε πλατφόρμες ψηφιακών κοινωνικών δικτύων για τους κινδύνους του «εθιστικού σχεδιασμού».

Η αναφορά μας στον «εθιστικό σχεδιασμό» κρίθηκε αναγκαία για να αναδειχθεί μια κρίσιμη παράμετρος της αντιμετώπισης των ζητημάτων που θέτει η χρήση της ΤΝ, αυτή της ανάγκης διασφάλισης των εγγυήσεων θεμιτής και σύννομης χρήσης μέσω του σχεδιασμού. Ήδη επιστήμονες του κλάδου εστιάζουν στα πεδία της *fairness by design*, που συμπεριλαμβάνει την *explainability by design* (*explainable AI*). Ο «ηθικός σχεδιασμός» των σχετικών υπηρεσιών πρέπει να προωθηθεί ως προεπιλογή, ενώ είναι σημαντικό να δημιουργηθεί ένας κατάλογος ορθών πρακτικών σχεδιαστικών χαρακτηριστικών που δεν είναι εθιστικά ή χειριστικά και να διασφαλίσει ότι οι χρήστες έχουν πλήρη έλεγχο και μπορούν να προβαίνουν σε συνειδητές και ενημερωμένες ενέργειες στο Διαδίκτυο, χωρίς να αντιμετωπίζουν υπερφόρτωση πληροφοριών ή να υπόκεινται σε υποσυνείδητες επιρροές.²⁹⁴

Ο «ηθικός σχεδιασμός» των συστημάτων ΤΝ έγκειται στην απαίτηση τα συστήματα αυτά να αναπτύσσονται, να εγκαθίστανται και να χρησιμοποιούνται κατά τρόπο που να συμμορφώνεται με τις ηθικές αρχές του σεβασμού της ανθρώπινης αυτονομίας, της πρόληψης της βλάβης, της δικαιοσύνης και της επεξηγησιμότητας. Το αίτημα αυτό συνδέεται με τη συζήτηση για υπεύθυνη έρευνα και καινοτομία (*Responsible Research and Innovation*) που συνίσταται σε μια διαφανή, διαδραστική διαδικασία, μέσω της οποίας οι κοινωνικοί φορείς και οι ερευνητές καθίστανται αμοιβαία ανταποκρινόμενοι ο ένας στον άλλον, με στόχο την (ηθική) αποδοχή, τη βιωσιμότητα και την κοινωνική επιθυμητότητα της διαδικασίας καινοτομίας και των εμπορεύσιμων προϊόντων της, ώστε να καταστεί δυνατή η ορθή ενσωμάτωση των επιστημονικών και τεχνολογικών εξελίξεων στην κοινωνία.

²⁹⁴ Βλ. την έκθεση του Ευρωπαϊκού Κοινοβουλίου: European Parliament. (2023). *Report on addictive design of online services and consumer protection in the EU single market (2023/2043(INI))* (A9-0340/2023). https://www.europarl.europa.eu/doceo/document/A-9-2023-0340_EN.htm

12.6. Κάποιες τελικές σκέψεις...

Όπως εντοπίσαμε σε αρκετά σημεία αυτής της μελέτης, η –συνάα απεριόριστη– αισιοδοξία ότι η ΤΝ συνιστά ένα νέο «Κέραρ της Αμαλθείας» που θα επιλύσει τα ανθρώπινα προβλήματα αντιπαρτίθεται σε μία –ενίοτε όλως καθηλωτική και δυστοπική– προοπτική μιας ανεξέλεγκτης Τεχνητής Νοημοσύνης, που φέρεται να ξεπερνάει την ανθρωπίνη, και στους φόβους που διατυπώνονται αναφορικά με την ικανότητα των μηχανών να αυτοβελτιώνονται, οδηγώντας την κοινωνία, τις δομές και τις σχέσεις της σε μη αναστρέψιμη μεταλλαγή.

Η συζήτηση πρέπει να εκκινεί, αλλά δεν πρέπει να εξαντλείται στη γενική προσέγγιση για την ΤΝ. Η γενική επίκληση αυτού του όρου²⁹⁵ εξυπηρετεί μεν μια –πιο εύληπτη– ρητορική, αλλά θολώνει την εικόνα και τη συζήτηση για τα ζητήματα, καθώς οι συζητήσεις για τη διακυβέρνηση της ΤΝ στοχεύουν σε μια αόριστη έννοια, αντί σε συγκεκριμένες δυνατότητες, ενώ οι συζητήσεις για τις πολιτικές και τις ρυθμίσεις γίνονται αφηρημένες, αντί να εστιάζουν σε συγκεκριμένες εφαρμογές. Δυνατότητες, ευκαιρίες και κίνδυνοι γενικεύονται λανθασμένα και εν τέλει δεν εντοπίζονται επαρκώς. Η ουσιαστική συζήτηση δεν σταματά στο «Τι πρέπει να κάνουμε για/με την ΤΝ;», αλλά να εκτείνεται στη διερεύνηση του συγκεκριμένου συστήματος και του συγκεκριμένου πλαισίου. Το ζητούμενο είναι τόσο απλό όσο και δυσχερώς επιτεύξιμο: Πρέπει να συζητήσουμε τα επιμέρους προβλήματα, τις απειλές και τους κινδύνους χωρίς να μας διαφεύγει η γενική εικόνα και στόχευση.

Το ερώτημα δεν είναι μόνο εάν και πώς μπορούμε να διάγουμε τη ζωή μας στην εποχή των αλγορίθμων, ώστε αυτή να είναι αξιοβίωτη ή –έστω– ανεκτή, αλλά και πώς θα κατανεμηθεί το προσδοκώμενο όφελος στην κοινωνία και στις γενεές. Από ποια αφετηρία και με ποιους στόχους θα επιδιώξουμε να ορίσουμε το μέλλον; Ποια είναι η κοσμοαντίληψη και το κοσμοείδωλό μας; Το ζήτημα δεν τίθεται (μόνο) γενικά και αφηρημένα, εάν δηλαδή ο άνθρωπος θα είναι σε θέση να ελέγξει κάτι που είναι πιο ευφυές από εκείνον, αλλά πολύ συγκεκριμένα, εάν και πώς είναι εφικτή η άσκηση ελέγχου και επιρροής στους –κυρίως ιδιωτικούς– φορείς ιδιωτικής τεχνολογικής

²⁹⁵ Wood, N. G., Robbins, S., Zagarra Berodt, E., Graf von Westerholt, A., Behrndt, M., Budig, H., & Kloock-Schreiber, D. (2026). Stop saying "AI". *Philosophy & Technology*, 39, Article 106. <https://doi.org/10.1007/s13347-026-01114-4>

και πληροφοριακής ισχύος, που καθορίζουν τη ροή και αξιοποίηση της πληροφορίας και τις τεχνολογικές εξελίξεις σε παγκόσμιο επίπεδο, σε μια πραγματικότητα που είναι διαρκώς μεταβαλλόμενη, αβέβαια, πολύπλοκη.

Χιλιάδες χρόνια τώρα οι άνθρωποι πασχίζουμε να προβλέψουμε, να ελέγξουμε, να διαχειριστούμε το μέλλον, όπως γράφει η Jennifer M. Gidley στο ομώνυμο βιβλίο της.²⁹⁶ Το τελευταίο διάστημα δεν είναι διόλου τυχαία η έμφαση σε όρους όπως ανθεκτικότητα στο μέλλον, προετοιμασία για το μέλλον, προόραση. Η ΤΝ μπορεί, πράγματι, να παραγάγει προ-γνώσεις που θα βοηθήσουν σε αυτό. Το ποιο μέλλον, από τα πολλά πιθανά, σαγηνευτικά ή απειλητικά, θα επιχειρήσουμε να επιδιώξουμε παραμένει ανθρώπινη, κοινωνική και πολιτική επιλογή.

Ανθρώπινη επιλογή είναι και η επιβεβαίωση και διασφάλιση της ανθρωπίνης πρωτοκαθεδρίας και υπεροχής. Η αντίληψη αυτή υπερβαίνει την ανθρωποκεντρική προσέγγιση και απαιτεί την εγγύηση ότι όπου διακυβεύονται η ανθρώπινη αξιοπρέπεια και αυτονομία, η ανθρώπινη αυτενέργεια πρέπει να διατηρείται και να προστατεύεται. Με άλλα λόγια, θα πρέπει να διασφαλίζεται ότι η ΤΝ θα ενισχύει τις ανθρώπινες δυνατότητες και δεν θα υποκαθιστά την ανθρώπινη κρίση σε ζητήματα θεμελιώδους σημασίας. Ταυτόχρονα, επιβάλλεται μια συζήτηση για την καλλιέργεια μιας νέας νοοτροπίας σύζευξης τεχνολογικής και οικονομικής ανάπτυξης για την εποχή της ΤΝ με την προστασία της δημοκρατίας και των δικαιωμάτων, την οποία θα μπορούσαμε να αποκαλέσουμε «κράτος δικαίου, δημοκρατία και ανθρώπινα δικαιώματα by design».²⁹⁷

²⁹⁶ Gidley, J. M. (2022). *Το Μέλλον*. Πανεπιστημιακές Εκδόσεις Κρήτης.

²⁹⁷ Ένα είδος τεχνολογίας-μηχανικής που είναι σύμφωνη με τις βασικές αρχές όπως αποτυπώνονται σε ένα δημοκρατικό Σύνταγμα, ένα είδος constitution-compliant engineering.



diaNEOsis



dianeosis_org



diaNEOsis



diaNEOsis



diaNEOsis

διαNEOsis
ΟΡΓΑΝΙΣΜΟΣ ΕΡΕΥΝΑΣ & ΑΝΑΛΥΣΗΣ