



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΛΟΠΟΝΝΗΣΟΥ
ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΑΣ, ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

Διπλωματική Εργασία για την απόκτηση του μεταπτυχιακού
τίτλου σπουδών “Επιστήμη Υπολογιστών”

**Η τεχνητή νοημοσύνη από την πολιτική στην
κοινωνική προοπτική: χρήση, ρύθμιση και
αποτίμηση από τον δημόσιο τομέα**

Ανδρονίκη Μαρία Τζομάκα
Α.Μ. 2022202102018

Επιβλέπων:

Καθηγητής Χρήστος Τρυφονόπουλος
Πανεπιστήμιο Πελοποννήσου

Τρίπολη, Μάρτιος 2025

Η εξεταστική επιτροπή

Χρήστος Τρυφονόπουλος
Καθηγητής
Πανεπιστήμιο Πελοποννήσου

Κώστας Βασιλάκης
Καθηγητής
Πανεπιστήμιο Πελοποννήσου

Δημοσθένης Κυριαζής
Καθηγητής
Πανεπιστήμιο Πειραιώς

Ανδρονίκη Μαρία Τζομάκα
Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστήμιο Πελοποννήσου

Copyright © , Ανδρονίκη Μαρία Τζομάκα 2025
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευτεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Πανεπιστημίου Πελοποννήσου.

Ignorant, quem portum petat, nullus suus ventus est.

Seneca, Ad Lucilium Epistulae Morales LXXI

Ευχαριστίες

Θέλω να ευχαριστήσω ορισμένους ανθρώπους που συνάντησα σε διαφορετικά στάδια αυτής της ερευνητικής διαδρομής και την καθόρισαν σε κρίσιμες καμπές της. Τους Κ.Σ. για την παρότρυνση, Χ.Τ. για την εμπιστοσύνη, Β.Μ. για τη σταθερότητα, Γ.Π. για την προσήνεια, Α.Π. για τη διεισδυτικότητα, Κ.Κ. για τον νέο ορίζοντα και Γ.Μ. για την υποστηρικτικότητα.

Δεν θα μιλήσω για το επιστημονικό βάθος και τις γνώσεις τους, άλλωστε με τη βοήθεια διαφόρων γλωσσικών μοντέλων, θα μπορούσε κάποιος και μόνος του να ανακαλύψει μεγάλο μέρος τους. Στάθηκα στα παραπάνω χαρακτηριστικά γιατί αυτές οι ανθρώπινες αξίες ήταν που μου πρόσφεραν τη σιγουριά και την έμπνευση να ξεκλειδώσω τον συναρπαστικό κόσμο της τεχνητής νοημοσύνης, με τρόπο που κανένα εργαλείο της δεν θα μπορούσε να μου μεταδώσει.

Τέλος, ένα παντοτινό ευχαριστώ στον Πάνο που με παρακινεί να ζαναμαθαίνω να μαθαίνω.

Περίληψη

Η παρούσα εργασία εξετάζει την επίδραση της Τεχνητής Νοημοσύνης στη δημόσια σφαίρα, αναδεικνύοντας τόσο τις προκλήσεις όσο και τα οφέλη που προκύπτουν από τη χρήση της. Η Τεχνητή Νοημοσύνη ως σύνολο τεχνολογιών και εφαρμογών έχει ήδη προκαλέσει ραγδαίες εξελίξεις σε ποικίλες πτυχές της δημόσιας και ιδιωτικής ζωής, με θετικές και αρνητικές όψεις. Οι σοβαρές προκλήσεις που επιφυλάσσει στα πεδία των θεμελιωδών δικαιωμάτων, του κράτους δικαίου, της δημοκρατίας και του περιβάλλοντος αντισταθμίζονται από ταχέως αναπτυσσόμενα οφέλη σε όλο το φάσμα της κοινωνικής και οικονομικής ζωής. Η έρευνα επικεντρώνεται στην Τεχνητή Νοημοσύνη από τη σκοπιά της ρύθμισης, της χρήσης και της κοινωνικής προοπτικής της, με έμφαση στις θεσμικές παρεμβάσεις και τις γεωπολιτικές επιπτώσεις της. Η εργασία εξετάζει την ανάγκη καθιέρωσης ενός σαφούς και κοινά αποδεκτού ορισμού της Τεχνητής Νοημοσύνης, αναδεικνύει πώς αυτή η συζήτηση παραμένει ανοιχτή και κρίσιμη, καθώς ο ορισμός επηρεάζει άμεσα το ρυθμιστικό πλαίσιο, τις πολιτικές και τα θεσμικά όρια στις πρακτικές εφαρμογές της τεχνολογίας.

Ιδιαίτερη έμφαση δίνεται στον δημόσιο τομέα, ο οποίος αναγνωρίζει την αξία των συστημάτων Τεχνητής Νοημοσύνης κατά την άσκηση του έργου του αλλά, κατόπιν πολλών περιστατικών που σημειώθηκαν, αντιλαμβάνεται την κλιμάκωση των κινδύνων που ελλοχεύουν. Στο πλαίσιο αυτό παρακολουθούμε πώς αναπτύσσει εργαλεία, όπως η Εκτίμηση Αλγοριθμικού Αντικτύπου που εφαρμόζεται στον Καναδά, και πώς θέτει προϋποθέσεις που συνδέονται με τη διακυβέρνηση δεδομένων και αλγορίθμων ως συστατικών στοιχείων της Τεχνητής Νοημοσύνης προκειμένου να διαχειριστεί τους κινδύνους, όπως προτείνει η χρήση του ALTAI (Assessment List for Trustworthy Artificial Intelligence), ενός εργαλείου αυτοαξιολόγησης που έχει αναπτυχθεί από την Ομάδα Εμπειρογνομόνων Υψηλού Επιπέδου (AI HLEG) της Ευρωπαϊκής Επιτροπής για να βοηθήσει τους φορείς να διασφαλίσουν την αξιοπιστία της ΤΝ.

Ένα από τα κεντρικά ζητήματα της εργασίας είναι η θεσμική ρύθμιση της Τεχνητής Νοημοσύνης, η οποία δεν αποτελεί απλώς τεχνικό αλλά κυρίως γεωπολιτικό διακύβευμα. Η παραπάνω διαπίστωση έχει κινητοποιήσει διεθνείς και υπερεθνικούς οργανισμούς, όπως η Ευρωπαϊκή Ένωση, κυβερνήσεις, ερευνητικούς και ακαδημαϊκούς φορείς να εργαστούν, ως ρυθμιστές, για μια χρηστή και δεοντολογική αξιοποίηση των συστημάτων τεχνητής νοημοσύνης. Συγχρόνως αναδεικνύει πώς η σύνδεση της Τεχνητής Νοημοσύνης με άλλους ρυθμιζόμενους τομείς, όπως η διακυβέρνηση δεδομένων, η ψηφιακή αγορά και η κυβερνοασφάλεια, ενισχύει την εδραίωση του ρυθμιστικού πλαισίου τονίζοντας ιδιαίτερα τη σημασία της Πράξης για τη Διακυβέρνηση Δεδομένων και την Πράξη για τα Δεδομένα που αποσκοπεί στην ελεύθερη ροή δεδομένων και στη δημιουργία μιας ενιαίας αγοράς βιομηχανικών δεδομένων.

Η έρευνα επίσης αναδεικνύει τη σημασία των τεχνικών και δεοντολογικών προτύπων ως βασικών μηχανισμών για την εξασφάλιση της υπεύθυνης ανάπτυξης και χρήσης της Τεχνητής Νοημοσύνης, εξετάζει διεξοδικά διεθνείς και ευρωπαϊκές πρωτοβουλίες τυποποίησης και διερευνά πώς τα πρότυπα μπορούν να ενσωματωθούν στα συστήματα μέτρησης της απόδοσης της ΤΝ, διασφαλίζοντας ότι η τεχνολογία παραμένει σύμφωνη με τις κοινωνικές αξίες, το κράτος δικαίου και τις δημοκρατικές αρχές. Ειδική έμφαση δίνεται στους δείκτες που σχετίζονται με τα θεμελιώδη δικαιώματα, τη δημοκρατία και την προστασία του κράτους δικαίου, καθώς και στη σημασία των

παρατηρητηρίων Τεχνητής Νοημοσύνης ως μηχανισμών παρακολούθησης και χάραξης πολιτικής. Με αυτόν τον τρόπο συμβάλλει στην προοπτική της μεγιστοποίησης του κοινωνικού οφέλους από την αξιόπιστη, υπεύθυνη και ανθρωποκεντρική χρήση της.

Η συμβολή της εργασίας έγκειται στην πρωτότυπη σύνθεση και αξιολόγηση των ανθρωποκεντρικών δεικτών επίδοσης, οι οποίοι σχετίζονται με την προστασία των θεμελιωδών δικαιωμάτων. Η ερευνητική μεθοδολογία βασίστηκε στη συγκριτική ανάλυση υφιστάμενων δεικτών και στη δημιουργία μιας ενοποιημένης προσέγγισης για την κατάρτιση δεικτών που διασφαλίζουν την ανθρωποκεντρική ανάπτυξη της Τεχνητής Νοημοσύνης, την κοινωνική συνοχή και τη δικαιοσύνη.

Η έρευνα βασίστηκε στη συστηματική διερεύνηση βασικών δεικτών επίδοσης (KPIs) της Τεχνητής Νοημοσύνης, με στόχο την αποτίμηση της αξιοπιστίας, της ανθρωποκεντρικής διάστασης και του κοινωνικού αντίκτυπου αυτών των συστημάτων. Η εργασία παρουσιάζει μια πρωτότυπη καταγραφή και συγκριτική ανάλυση υφιστάμενων δεικτών από διεθνείς οργανισμούς, ερευνητικά κέντρα και ευρωπαϊκούς θεσμούς, ενώ ταυτόχρονα προτείνει νέα μεθοδολογικά εργαλεία για τη μέτρηση της κοινωνικής και ανθρωποκεντρικής διάστασης της Τεχνητής Νοημοσύνης.

Εν κατακλείδι, η εργασία συμβάλλει στην κατανόηση των ρυθμιστικών προκλήσεων, προτείνοντας τρόπους για την προώθηση μιας υπεύθυνης και αξιόπιστης χρήσης της Τεχνητής Νοημοσύνης. Μέσα από μια πολυδιάστατη προσέγγιση, καταδεικνύει ότι η επιτυχής ενσωμάτωσή της απαιτεί όχι μόνο τεχνικές λύσεις, αλλά και βαθιά κατανόηση των κοινωνικών, ηθικών και πολιτικών της προεκτάσεων.

Abstract

This study examines the impact of Artificial Intelligence on the public sphere, highlighting both the challenges and benefits that arise from its use. As a set of technologies and applications, Artificial Intelligence has already triggered rapid developments in various aspects of public and private life, with both positive and negative implications. The serious challenges it poses in the fields of fundamental rights, the rule of law, democracy, and the environment are counterbalanced by its rapidly growing benefits across the spectrum of social and economic life.

The study focuses on Artificial Intelligence from the perspective of regulation, application, and social implications, with an emphasis on institutional interventions and geopolitical factors. The research explores the need to establish a clear and universally accepted definition of Artificial Intelligence, underscoring how this ongoing debate is crucial, as it directly influences regulatory frameworks, policies, and the institutional limits of Artificial Intelligence applications.

Particular emphasis is placed on the public sector, which recognizes the value of Artificial Intelligence systems in performing its functions. However, following multiple incidents, it also acknowledges the increasing risks associated with Artificial Intelligence. In this context, the study examines how the public sector develops tools, such as the Algorithmic Impact Assessment applied in Canada, and how it sets conditions related to data governance and algorithmic accountability, which are essential components of Artificial Intelligence. These efforts aim to manage potential risks, in line with recommendations for using ALTAI (Assessment List for Trustworthy Artificial Intelligence)—a self-assessment tool developed by the High-Level Expert Group on Artificial Intelligence (AI HLEG) of the European Commission to help organizations ensure its trustworthiness.

A central issue addressed in this thesis is the institutional regulation of Artificial Intelligence, which is not merely a technical matter but primarily a geopolitical challenge. This realization has prompted international and supranational organizations, such as the European Union, as well as governments, research institutions, and academic bodies, to act as regulators in promoting the responsible and ethical deployment of Artificial Intelligence systems. At the same time, the study highlights how Artificial Intelligence is interconnected with other regulated sectors, such as data governance, the digital market, and cybersecurity, strengthening the regulatory framework. Special attention is given to the Data Governance Act and the Data Act, which aim to facilitate the free flow of data and create a unified market for industrial data.

The research also underscores the importance of technical and ethical standards as key mechanisms for ensuring the responsible development and deployment of Artificial Intelligence. It thoroughly examines international and European standardization initiatives and explores how standards can be integrated into Artificial Intelligence performance measurement systems, ensuring that the technology remains aligned with social values, the rule of law, and democratic principles. Special emphasis is placed on indicators related to fundamental rights, democracy, and legal protections, as well as the significance of Artificial Intelligence observatories as mechanisms for monitoring and policymaking. In this way, the study contributes to the maximization of social benefits through the reliable, responsible, and human-centric use of Artificial Intelligence.

The contribution of this work lies in the original synthesis and evaluation of human-centered performance indicators (KPIs), particularly those related to fundamental rights compliance. The methodology is based on a comparative analysis of existing indicators and the creation of a unified approach for establishing metrics that ensure Artificial Intelligence's human-centered development, social cohesion, and fairness.

The research was based on the systematic investigation of key performance indicators (KPIs) of Artificial Intelligence, aiming to assess the reliability, human-centered dimension, and social impact of these systems. This study presents an original documentation and comparative analysis of existing indicators from international organizations, research centers, and European institutions, while also proposing new methodological tools for measuring the social and human-centered dimensions of Artificial Intelligence.

In conclusion, the present thesis contributes to the understanding of regulatory challenges, offering strategies for the promotion of responsible and trustworthy Artificial Intelligence. Through a multidimensional approach, it demonstrates that the successful integration of Artificial Intelligence requires not only technical solutions but also a profound understanding of its social, ethical, and political implications.

Πίνακας περιεχομένων

Εισαγωγή	2
1 Επισκόπηση και ορισμοί της Τεχνητής Νοημοσύνης	7
1.1 Επισκόπηση	7
1.2 Ιστορικό της Τεχνητής Νοημοσύνης ως επιστήμης	9
1.3 Ιστορική αναδρομή σε ορισμούς TN και σχετικές έννοιες	14
1.4 Ορισμός της TN στην Ευρωπαϊκή Ένωση	18
1.4.1 Ορισμός TN: Ανακοίνωση της Επιτροπής, Απρίλιος 2018	18
1.4.2 Ορισμός TN: Ομάδα εμπειρογνομόνων (HLEG), Απρίλιος 2019	18
1.4.3 Ορισμός TN: Πράξη για την TN, Απρίλιος 2021	18
1.4.4 Αναγκαιότητα κοινά αποδεκτού ορισμού για ασφάλεια δικαίου	19
1.4.5 Σύνοψη των ορισμών TN: ΕΕ, ΗΠΑ και Κίνα	20
1.5 Ορισμός της TN από τον ΟΟΣΑ	21
1.6 Συμπέρασμα	24
2 Εφαρμογές της TN στον δημόσιο τομέα και προκλήσεις	25
2.1 Πρακτική εφαρμογή της TN στον δημόσιο τομέα	25
2.1.1 “Νομικό μαύρο κουτί” έναντι “τεχνικού μαύρου κουτιού”	26
2.1.2 Εσωτερική υλοποίηση έναντι εξωτερικής υλοποίησης	26
2.1.3 Οργανωσιακές ικανότητες	27
2.1.4 Δεδομένα: εσωτερικά έναντι εξωτερικών	27
2.1.4.1 Εσωτερικά δεδομένα	27
2.1.4.2 Εξωτερικά δεδομένα	27
2.1.5 Εργασίες	28
2.1.5.1 Επαναλαμβανόμενες εργασίες	28
2.1.5.2 Προγνωστικές μη εκτελέσιμες εργασίες	28
2.1.5.3 Προγνωστικές εκτελέσιμες εργασίες	28
2.1.6 Αντίκτυπος της χρήσης ευφύων συστημάτων υποστήριξης	28
2.2 Η TN στην υπηρεσία της κοινοβουλευτικής λειτουργίας	30
2.3 Το παράδοξο της ρύθμισης της TN και οι εκλογές στο παγκόσμιο τοπίο το 2024	32
2.4 Αλγοριθμική μεροληψία και κράτος δικαίου: the Dutch scandal	34
2.4.1 The Dutch scandal: συνοπτικό ιστορικό	35
2.4.2 Ανάλυση πραγματικών περιστατικών από την Επιτροπή της Βενετίας	36
2.4.3 Συμπέρασμα της Επιτροπής της Βενετίας	38
2.5 Εθνικές στρατηγικές και εργαλεία αλγοριθμικής λογοδοσίας	38
2.6 Εκτίμηση Αλγοριθμικού Αντικτύπου της Ομοσπονδιακής Κυβέρνησης του Καναδά	40
2.7 Συμπέρασμα	45

3	Γιατί και πώς ρυθμίζεται θεσμικά η ΤΝ	49
3.1	Παγκόσμια επισκόπηση πρωτοβουλιών ρύθμισης της ΤΝ	49
3.2	Γιατί και σε ποιους τομείς είναι αναγκαία η ρύθμιση	52
3.3	Η Ευρώπη πρωτοστατεί στη ρύθμιση με γεωπολιτικούς στόχους	54
3.3.1	Το σκεπτικό που οδήγησε την Ευρώπη στη ρύθμιση της ΤΝ	54
3.3.2	Επιδίωξη για πρωτοπορία της Ευρώπης	55
3.4	Γεωπολιτική και τεχνολογία	56
3.4.1	“Εθνικισμός” μέσω της τεχνολογίας	57
3.4.2	Brussels effect (Φαινόμενο των Βρυξελλών)	58
3.4.3	Digital Services Act και Digital Markets Act	59
3.4.4	Ευρωπαϊκή στρατηγική για τα δεδομένα (A European strategy for data) .	60
3.4.4.1	Ευρωπαϊκή διακυβέρνηση δεδομένων (European data governance)	61
3.4.4.2	Κανονισμός για τα Δεδομένα (Data Act)	62
3.5	Διαβούλευση και πιλοτικά εργαλεία που προηγήθηκαν της πρότασης της Πράξης για την ΤΝ	63
3.6	Πράξη για την ΤΝ: Συμφραζόμενα, ιστορικό και ρύθμιση	64
3.6.1	Συμφραζόμενα	64
3.6.2	Ιστορικό	66
3.6.2.1	Η θέση της Ευρωπαϊκής Επιτροπής	67
3.6.2.2	Παραδείγματα βαθμών κινδύνου	67
3.6.2.3	Η θέση του Συμβουλίου της ΕΕ	68
3.6.2.4	Η Θέση του Κοινοβουλίου της ΕΕ	69
3.6.3	Ρύθμιση	71
3.6.3.1	Πεδίο εφαρμογής	71
3.6.3.2	Μοντέλα GPAI/θεμελίωσης και απαιτήσεις διαφάνειας	71
3.6.3.3	Απαιτήσεις για συστήματα τεχνητής νοημοσύνης υψηλού κινδύνου	72
3.6.3.4	Εξαιρέσεις για την επιβολή του νόμου	73
3.6.3.5	Απαγορευμένες πρακτικές	73
3.6.3.6	Κυρώσεις	74
3.6.3.7	Τι έπεται;	74
3.7	Ζητήματα κυβερνοαφάλειας στην Πράξη για την ΤΝ (άρθρο 15)	74
3.8	Συμπέρασμα	76
4	Παρατηρητήρια, πρότυπα και δείκτες προς μια αξιόπιστη και ανθρωποκεντρική ΤΝ	77
4.1	Αξιόπιστη και ανθρωποκεντρική τεχνητή νοημοσύνη	77
4.1.1	Κατευθυντήριες γραμμές δεοντολογίας για αξιόπιστη ΤΝ	78
4.1.2	Αξιόπιστη τεχνητή νοημοσύνη	79
4.1.3	Ανθρωποκεντρική ΤΝ και η αρχή του σεβασμού της ανθρώπινης αυτονομίας	80
4.1.4	Ο κατάλογος αυτοαξιολόγησης για την αξιόπιστη ΤΝ (ALTAI)	80
4.2	Ειδικότερα για τα Θεμελιώδη Δικαιώματα	80
4.2.1	Τα θεμελιώδη δικαιώματα στην αιτιολογική έκθεση της Πράξης για την ΤΝ	81
4.2.2	FRA - European Agency for Fundamental Rights	83
4.3	Παρατηρητήρια: τα κείμενα της ΕΕ προβλέπουν	86
4.3.1	Ο ρόλος και η λειτουργία ενός Παρατηρητηρίου ΤΝ στην ΕΕ	87
4.4	Τυποποίηση και πρότυπα	89
4.4.1	Η σειρά προτύπων IEEE 7000 για δεοντολογικό σχεδιασμό και διακυβέρνηση της ΤΝ	90
4.4.1.1	IEEE: Σύντομο ιστορικό και ο ρόλος του στη δεοντολογία για τον κοινωνικό αντίκτυπο της ΤΝ	90

4.4.1.2	Το πρότυπο IEEE 7010-2020 για την αξιολόγηση των επιπτώσεων της ΤΝ στην ανθρώπινη ευημερία	91
4.4.2	Τυποποίηση και πρότυπα για την ΤΝ στην ΕΕ	92
4.5	Βασικοί δείκτες επίδοσης (KPIs) για την ΤΝ	94
4.5.1	Επισκόπηση και τεκμηρίωση των βασικών δεικτών επίδοσης μέσα από ευρωπαϊκά κείμενα	94
4.5.2	Μεθοδολογία έρευνας για την εξεύρεση βασικών δεικτών επίδοσης-KPIs	95
4.5.3	Ψηφιακή πυξίδα για την ψηφιακή δεκαετία 2020-30 και δείκτες DESI . .	96
4.5.3.1	Βασικοί δείκτες επίδοσης	97
4.5.3.2	Ποιες είναι οι πηγές των δεδομένων και πώς διαμορφώνεται ο δείκτης DESI για την ΤΝ;	98
4.5.4	Δείκτες στην έρευνα “European enterprise survey on the use of technologies”	100
4.5.5	Βασικοί Δείκτες Επίδοσης στην Πρόταση Κανονισμού “Πράξη για την ΤΝ”	100
4.5.5.1	Δείκτες επίδοσης της χρήσης της ΤΝ στο κείμενο της “Πράξης για την ΤΝ”	100
4.5.5.2	Αναφορά στους Στόχους Βιώσιμης Ανάπτυξης των Ηνωμένων Εθνών	101
4.5.5.3	Δείκτες επίδοσης της υλοποίησης της “Πράξης για την ΤΝ” . .	101
4.5.6	Δείκτες του AI WATCH	102
4.5.6.1	AI Watch, National Strategies on AI: A European Perspective, 2022	102
4.5.6.2	AI Watch Index 2021	103
4.5.7	Δείκτες του STAFF WORKING PAPER: AI Act Impact Assessment . .	104
4.5.8	AI Index, Stanford Institute for Human-Centered AI (HAI)	104
4.5.9	Δείκτες στον Ελληνικό Νόμο 4961/2022	105
4.5.10	Δείκτες που αντλούμε από την ταξινόμηση των συστημάτων ΤΝ του ΟΟΣΑ	106
4.6	Συμπέρασμα	108
5	Συμπεράσματα	119

Κατάλογος σχημάτων

1	René Magritte, <i>Time Transfixed</i> (1938) Art Institute of Chicago, © 2018 C. Herscovici, London / Artists Rights Society (ARS), New York https://www.artic.edu/artworks/34181/la-duree-poignardee-time-transfixed	1
1.1	Χρονολόγιο θεωριών TN, Mission AI, Springer, 2023	10
1.2	Η TN ως επιστημονικός κλάδος (πρώτο κύμα), Mission AI, Springer, 2023	13
1.3	Τι είναι Τεχνητή Νοημοσύνη: διάγραμμα ροής, Karen Hao, MIT Technology Review, 2018	15
1.4	A high-level conceptual view of an AI system, OECD, 2019	22
1.5	Detailed conceptual view of an AI System, OECD, 2019	22
1.6	Updates to the OECD’s definition of an AI system, OECD, Nov. 2023	23
2.1	(1/3) Πιλοτικά έργα Τεχνητής Νοημοσύνης για δημόσιες υπηρεσίες, Omar Ballester, 2021, An Artificial Intelligence Definition and Classification Framework for Public Sector Applications, ACM	46
2.2	(2/3) Πιλοτικά έργα Τεχνητής Νοημοσύνης για δημόσιες υπηρεσίες, Omar Ballester, 2021, An Artificial Intelligence Definition and Classification Framework for Public Sector Applications, ACM	47
2.3	(3/3) Πιλοτικά έργα Τεχνητής Νοημοσύνης για δημόσιες υπηρεσίες, Omar Ballester, 2021, An Artificial Intelligence Definition and Classification Framework for Public Sector Applications, ACM	48
3.1	A ‘risk-based’ approach	66
4.1	Τα δικαιώματα του Χάρτη Θεμελιωδών Δικαιωμάτων που άπτονται της Πράξης TN	82
4.2	Θεμελιώδη δικαιώματα στην τεχνητή νοημοσύνη: Τι πρέπει να ληφθεί υπόψη, FRA	84
4.3	Απόσπασμα από την “Περίληψη των κινδύνων της TN για τα θεμελιώδη δικαιώματα”, Παράρτημα μελέτης DG CNECT	86
4.4	Η επιρροή που άσκησε το EAD του IEEE σε άλλες πρωτοβουλίες ρύθμισης της TN	91
4.5	Σύνοψη των δεικτών του AI Watch Index 2021	103
4.6	Βασικές διαστάσεις του πλαισίου του ΟΟΣΑ για την ταξινόμηση των συστημάτων τεχνητής νοημοσύνης	106

Κατάλογος πινάκων

2.1	Πλαίσιο ταξινόμησης για τη χρήση εφαρμογών ΤΝ από τον δημόσιο τομέα	30
3.1	DSA & DMA: Βασικές πτυχές και διαφορές, Sérvulo	59
4.1	DESI σε συνδυασμό με EUROSTAT: Δείκτες για την ΤΝ	99
4.2	Δείκτες για την παρακολούθηση της υλοποίησης της πρότασης από το «Νομοθετικό Δημοσιονομικό Δελτίο» της προτεινόμενης Πράξης για την ΤΝ	102
4.3	Δείκτες παρακολούθησης και αξιολόγησης (μετάφραση του “Table 13: Indicators for monitoring and evaluation” του εγγράφου Staff Working Paper, AI Act Impact Assessment)	104
4.4	Συγκεντρωτικός πίνακας Βασικών Δεικτών Επίδοσης για την ΤΝ (KPIs for AI) .	117



Εικόνα 1. René Magritte, *Time Transfixed* (1938)

Art Institute of Chicago, © 2018 C. Herscovici, London / Artists Rights Society (ARS), New York
<https://www.artic.edu/artworks/34181/la-duree-poignardee-time-transfixed>

Εισαγωγή

Στην ενότητα αυτή παρουσιάζουμε τον σκοπό, τους στόχους και τη δομή της εργασίας.

Σκοπός εργασίας

Η παρούσα εργασία επιχειρεί να καταδείξει την αναγκαιότητα ρύθμισης της χρήσης της Τεχνητής Νοημοσύνης (TN) αλλά και το πώς οι πρωτοβουλίες ρύθμισης, διεθνώς, αν και πηγάζουν από κοινούς προβληματισμούς και οικουμενικές αρχές, εντούτοις περνούν από τα διαφορετικά φίλτρα στοχεύσεων, προτεραιοτήτων και γεωπολιτικών επιδιώξεων που θέτει η εκάστοτε δύναμη, κυρίαρχο κράτος, φόρουμ ή οργανισμός και κατ' αυτόν τον τρόπο αποκρυσταλλώνονται και υλοποιούνται διαφοροποιημένα. Ως επίκεντρο της μελέτης επελέγη η περίπτωση ρύθμισης από την Ευρωπαϊκή Ένωση καθώς η συγγραφή της εργασίας συνέπεσε με τη μακρά περίοδο διαπραγματεύσεων επί της προταθείσας “Πράξης για την TN”, που εικάζεται ότι θα έχει παγκόσμια επιρροή όπως θα αναλυθεί υπό το πρίσμα του “Φαινομένου των Βρυξελλών”, και μας ενδιαφέρει για τον επιπλέον λόγο ότι ως κράτος-μέλος η Ελλάδα υπάγεται άμεσα σε αυτήν. Σκοπός λοιπόν ήταν να κατανοηθεί η αφετηρία, η διαδικασία αλλά και οι προκλήσεις του εγχειρήματος, κάτι που αποδεικνύουν οι διακυμάνσεις στην εξέλιξή του, ειδικά από την έλευση στο προσκήνιο του Chat GPT-4, μέχρι τις παραμονές της δημοσίευσης του επίσημου πλέον κειμένου του Κανονισμού, τον Ιούλιο 2024, το οποίο διευκρινίζεται ότι δεν αποτέλεσε μέρος της ερευνηθείσας ύλης.

Εν συνεχεία ο ερευνητικός σκοπός εμπλουτίστηκε με τον συσχετισμό του δημοσίου τομέα με την τεχνητή νοημοσύνη μέσα από το ρόλο του χρήστη, που κατέλαβε επίσης σημαντικό μέρος της έρευνας που πραγματοποιήθηκε και επιχείρησε να καλύψει τόσο τις δυνατότητες χρήσης όσο και τις περιπτώσεις κατάχρησης με κρούσματα αδιαφάνειας και μεροληψίας.

Επιπλέον, διακριτός σκοπός υπήρξε να ερευνηθεί το πλήθος των διεργασιών που προηγήθηκαν πριν ολοκληρωθεί το κείμενο της ευρωπαϊκής πρότασης της Πράξης για την TN, ακριβώς για να εμπεδωθεί καλύτερα η πολυπλοκότητα αυτού που αποκαλούμε ρύθμιση, αρχής γενομένης από τον ορισμό αυτού που επιθυμούμε να ρυθμιστεί. Πρόκειται για ζυμώσεις που έγιναν μέσα σε μια περίοδο ρευστότητας και ταχύτατης τεχνολογικής εξέλιξης με επίγνωση του επεκτατικού αποτελέσματος πέραν των γεωγραφικών ορίων που επιφέρει στις δικαιοδοσίες, της επελθούσας όσμωσης του φυσικού με τον ψηφιακό κόσμο που προκαλεί η πληθώρα των δεδομένων, δηλαδή των ψηφιοποιημένων πληροφοριών, και της συνακόλουθης αναγκαιότητας για διαχείριση και αξιοποίηση αυτών, όχι μόνο αυτοτελώς αλλά και ως πρώτης ύλης για τη λειτουργία των συστημάτων TN.

Τέλος ο σκοπός της εργασίας επεκτάθηκε στην τεκμηρίωση μιας τεχνητής νοημοσύνης αξιόπιστης και ανθρωποκεντρικής που εξυπηρετεί το δημόσιο συμφέρον, και στην αναζήτηση των οργανωσιακών, τεχνικών και απολογιστικών μέσων που μπορούν να συμβάλουν σε αυτή την κατεύθυνση.

Στόχοι εργασίας

Η εργασία, στο πλαίσιο του περιγραφέντος σκοπού, συγκέντρωσε ορισμούς της TN από ποι-

κίλα θεσμικά κείμενα και στη συνέχεια, περιπτώσεις θεσμικού πλαισίου για τη ρύθμιση αυτής, σε συνάρτηση με τους περιρρέοντες γεωπολιτικούς ανταγωνισμούς. Για το λόγο αυτό μάλιστα επεκτάθηκε στην αποτύπωση της ρύθμισης άλλων τομέων, ειδικά από την ΕΕ που επιχειρεί να εδραιώσει τη ρυθμιστική πρωτοκαθεδρία της παγκοσμίως, οι οποίοι είναι σύμφυτοι με την ανάπτυξη της ΤΝ. Τέτοιοι τομείς, όπως η διακυβέρνηση των δεδομένων, προσωπικών και μη, των ψηφιακών υπηρεσιών και των όρων λειτουργίας της ψηφιακής αγοράς, καθώς και ο ρόλος της κυβερνοασφάλειας στην ευρωπαϊκή ρύθμιση της ΤΝ, αναλύθηκαν επίσης σε επίπεδο στρατηγικών και νομοθετικών προβλέψεων. Ειδικά για τη χρήση των δεδομένων επιδιώχθηκε να εκτεθεί το πόσο είναι σημαντική η Πράξη για τη Διακυβέρνηση Δεδομένων¹ (όπως περιγράφεται στον Κανονισμό 2022/868 και αφορά την περαιτέρω χρήση, εντός της Ένωσης, ορισμένων κατηγοριών δεδομένων που βρίσκονται στην κατοχή φορέων του δημόσιου τομέα) και η Πράξη για τα Δεδομένα² (πρόταση Κανονισμού 2022/68) που αποσκοπεί στη δημιουργία μιας ενιαίας αγοράς για τα βιομηχανικά δεδομένα που παράγονται από συνδεδεμένες συσκευές (Διαδίκτυο των πραγμάτων) και στην ελεύθερη ροή δεδομένων μη προσωπικού χαρακτήρα μεταξύ επιχειρήσεων, της πανεπιστημιακής κοινότητας, και του δημόσιου τομέα.

Παρουσιάζοντας λοιπόν τη ρύθμιση της ΤΝ, τέθηκε ως εξίσου σημαντικός στόχος να εξεταστούν τα μέτρα που ρυθμίζουν τη χρήση και την πρόσβαση σε δεδομένα εντός της Ευρωπαϊκής Ένωσης, τα οποία δεν είναι “δεδομένα προσωπικού χαρακτήρα” όπως τα εννοεί ο ΓΚΠΔ³. Πρόκειται για μέτρα που πρέπει να συνδυαστούν με την Πράξη για τη Διακυβέρνηση των Δεδομένων που έχει στόχο να θεσπίσει κοινά πρότυπα για την περαιτέρω χρήση των δεδομένων σε όλους τους τομείς και την αντιμετώπιση των αθέμιτων εμπορικών πρακτικών μαζί με την έλλειψη διαθεσιμότητας δεδομένων. Ήταν σημαντικό να τεκμηριωθεί ότι υπάρχει αλληλεπίδραση μεταξύ της Πράξης για τα Δεδομένα και του ΓΚΠΔ της ΕΕ, δεδομένου ότι και οι δύο αποτελούν νομοθεσία που σχετίζεται με τα δεδομένα, αλλά και να καταδειχθεί ότι έχουν βασικές αποκλίσεις. Η Πράξη για τα Δεδομένα διαφέρει από τον ΓΚΠΔ, κυρίως στο ότι ο ΓΚΠΔ ρυθμίζει αποκλειστικά τα δεδομένα προσωπικού χαρακτήρα (δεδομένα που μπορούν να ταυτοποιήσουν φυσικά πρόσωπα), η Πράξη για τα Δεδομένα ρυθμίζει επίσης τα μη προσωπικά δεδομένα, συμπεριλαμβανομένων σχεδόν κάθε είδους ψηφιοποιημένων πληροφοριών, είτε τα δεδομένα είναι “ανώνυμα” είτε όχι⁴.

Γενικότερα παρουσιάστηκε εκτενώς η χρήση της ΤΝ από τον δημόσιο τομέα, προκειμένου να προαγάγει το έργο του, τις υπηρεσίες του αλλά και την κοινοβουλευτική του λειτουργία. Ειδικά στην περίπτωση του δημόσιου τομέα ως χρήστη, αναζητήθηκαν οι τρόποι αξιοποίησης της ΤΝ για την άσκηση κυβερνητικής πολιτικής, εκτέθηκαν χαρακτηριστικά παραδείγματα περιστατικών αδιαφάνειας και μεροληψίας αλλά και εργαλεία μετριασμού των πιθανών κινδύνων, όπως η Εκτίμηση Αλγοριθμικού Αντικτύπου που εφαρμόζει το Καναδικό κράτος και προβλέπει ο Ελληνικός νόμος 4961/22. Στο ίδιο πλαίσιο εξετάστηκε ο κίνδυνος νόθευσης των δημοκρατικών διαδικασιών μέσα από τη χρήση βαθυπαραποίησης (deep fakes) κατά τη διάρκεια των επάλληλων προεκλογικών εκστρατειών και εκλογικών διαδικασιών κατά το έτος 2024.

Στόχος της εργασίας υπήρξε επίσης η έρευνα και η συγκέντρωση του σχετικού νομικού πλαισίου και ακολούθως η εξέταση του πώς επιδιώχθηκε να εξισορροπηθούν η καινοτομία, η τεχνική πρόοδος, η νομοθετική παρέμβαση και η δικαιική προστασία παραδεδωμένων αρχών και αναγνωρισμένων δικαιωμάτων, μέσα σε ένα αναπτυσσόμενο τεχνολογικό περιβάλλον, με τρόπο που να προάγει την ατομική και την κοινωνική ευημερία. Αυτό συνοψίζεται εύστοχα στο λεγόμενο “δίλημμα του Collingridge” (1980) όπως αναφέρθηκε και στη Βουλή των Ελλήνων⁵ κατά την ψήφιση

¹<https://eur-lex.europa.eu/eli/reg/2022/868/oj>

²<https://tinyurl.com/ydr9z52r>

³Γενικός κανονισμός για την προστασία δεδομένων (ΓΚΠΔ), Κανονισμός (ΕΕ) 2016/679, (<https://eur-lex.europa.eu/EL/legal-content/summary/general-data-protection-regulation-gdpr.html>)

⁴<https://www.lexology.com/library/detail.aspx?g=cb288792-b02f-4d06-a56b-a29785a6e0b0>

⁵Η χώρα πολιτογραφείται στο μέλλον, Δελτία Τύπου, Υπουργείο Ψηφιακής Διακυβέρνησης, 25 Ιουλίου 2022, <https://mindigital.gr/archives/3442>, Το “δίλημμα του Collingridge” αναφέρεται στη στάθμιση της κατάλληλης στιγμής για τη νομοθετική ρύθμιση μιας καινοφανούς τεχνολογίας, που δεν

του Ν. 4961/2022 (25.7.2022).

Το επιστέγασμα όλων των παραπάνω είναι ότι αν και γνωρίζουμε πως η χρήση της τεχνητής νοημοσύνης προσφέρει τεράστιες δυνατότητες σε όλους τους τομείς, δεν γνωρίζουμε πώς ακριβώς να την χρησιμοποιούμε υπεύθυνα⁶. Οδηγό για την ολοκλήρωση της παρούσας εργασίας αποτέλεσε η αναζήτηση κοινωνικών, ανθρωποκεντρικών και περιβαλλοντικών δεικτών που αποτυπώνουν την υπεύθυνα χρήση της και των φορέων που μπορούν να τους εφαρμόσουν, αλλά και η εκ των προτέρων εξασφάλιση μιας αξιόπιστης ΤΝ δια της τυποποίησης και της κατάρτισης προτύπων με άξονα τη δεοντολογία και τη συμπληρωματικότητα με τις ανθρωπιστικές επιστήμες. Εις επίρρωσιν αυτού, η παρούσα εργασία στόχο έχει να ρίξει φως στην προοπτική για αξιόπιστη και ανθρωποκεντρική ΤΝ μέσα από την ανάπτυξη μετρικών που προσεγγίζουν με νέο τρόπο αυτή τη νέα πραγματικότητα, ειδικότερα όταν η χρήση πρόκειται να γίνει από τον δημόσιο τομέα και μέσα σε ένα πλαίσιο εύρυθμης λειτουργίας της δημοκρατίας και του κράτους δικαίου. Συνολικότερος στόχος, συνεπώς, υπήρξε να μετατοπιστεί το ενδιαφέρον σε ανθρωποκεντρικούς δείκτες επίδοσης των συστημάτων ΤΝ, δηλαδή να γίνει η αποτίμηση των τελευταίων μέσα από την κατάκτηση βαθύτερης και ουσιαστικότερης γνώσης για την ανθρώπινη φύση και την οπτική μιας ευημερίας για όλους, των θεμελιωδών δικαιωμάτων και του κράτους δικαίου.

Δομή εργασίας

Για να κατανοηθεί η διαδρομή προς την προοπτική κατάλληλων μοντέλων διακυβέρνησης της τεχνητής νοημοσύνης επιλέχθηκε στην παρούσα εργασία, αρχικά, να παρατεθούν επιγραμματικά οι επιστημονικοί ορισμοί της και το επιστημονικό και τεχνολογικό ιστορικό υπόβαθρό της (**Κεφάλαιο 1**). Η αποκρυστάλλωση κοινά αποδεκτού ορισμού είναι θεμελιώδης για την ασφάλεια δικαίου και αναλύεται διεξοδικά υπό το πρίσμα της ΕΕ και ειδικότερα της πρότασης της Πράξης για την ΤΝ, δεδομένου ότι θα αποτελέσει το πρώτο πλήρες ρυθμιστικό πλαίσιο με οικουμενικό χαρακτήρα και επιβολή κυρώσεων για τους παραβάτες. Όσον αφορά την ιστορική αναδρομή στη δημιουργία του επιστημονικού κλάδου της ΤΝ, εστίασαμε περισσότερο στην παρουσίαση του πρώτου κύματος καθώς ανάγεται σε πιο απομακρυσμένη, άρα και λιγότερο οικεία χρονική περίοδο.

Στη συνέχεια της εργασίας (**Κεφάλαιο 2**) διαπιστώθηκε η ανάγκη για ρύθμιση του ρόλου του ίδιου του δημόσιου τομέα ως φορέα ανάπτυξης και χρήστη συστημάτων τεχνητής νοημοσύνης. Στην εργασία συνεπώς αποτυπώνονται ποικίλες δυνατότητες της τεχνολογίας αυτής και οι προεκτάσεις της προς τομείς που την ανέδειξαν σε ένα πολύτιμο εργαλείο πολιτικής διακυβέρνησης και δημόσιας διοίκησης. Η εργασία πραγματεύεται επίσης το πώς ο δημόσιος τομέας σταδιακά συνειδητοποιεί αφενός, την αξία της χρήσης συστημάτων τεχνητής νοημοσύνης κατά την άσκηση του έργου του και αφετέρου, τους κινδύνους που ελλοχεύουν σε αυτό το πεδίο. Στο πλαίσιο αυτό εκτίθενται κάποια αντιπροσωπευτικά παραδείγματα χρήσεων και προκλήσεων. Επίσης δειγματοληπτικά επισημαίνει προϋποθέσεις που θέτουν δημόσιοι φορείς ως χρήστες της ΤΝ και ενδεικτικά απαριθμούνται ορισμένα εργαλεία που έχουν χρησιμοποιηθεί για μετριασμό των κινδύνων, τα οποία συνδέονται με τη διακυβέρνηση δεδομένων και αλγορίθμων, ως συστατικών στοιχείων της ΤΝ.

Ακολούθως (**Κεφάλαιο 3**) αναλύθηκε το πώς οι παραπάνω διαπιστώσεις έχουν κινητοποιήσει τον δημόσιο τομέα, τους διεθνείς οργανισμούς, την Ευρωπαϊκή Ένωση, τις κυβερνήσεις αλλά και τη δημόσια σφαίρα γενικότερα, όπως μη κυβερνητικούς οργανισμούς, την κοινωνία των πολιτών και εκπροσώπους της ακαδημαϊκής και ερευνητικής κοινότητας, να αναζητήσουν τρόπους ρύθμισης για δεοντολογική και σύννομη χρήση των συστημάτων τεχνητής νοημοσύνης. Η αναζήτηση

θα είναι πολύ πρόωπη ούτε πολύ όψιμη: Λίλιαν Μήτρου, Η “ρύθμιση” της Τεχνητής Νοημοσύνης ή το δίλημμα του Collingridge, <https://www.iit.demokritos.gr/wp-content/uploads/2020/06/Dr.-Lilian-Mitrou-The-regulation-of-Artificial-Intelligence-or-the-Collingridge-dilemma.pdf>

⁶<https://www.tortoisemedia.com/intelligence/tortoiseairesponsibility/>

κατάλληλου μοντέλου διακυβέρνησης διαμορφώνεται από το πρίσμα μιας διαρκώς αυξανόμενης επίγνωσης του ότι η χρήση της ΤΝ εκφεύγει των παραδοσιακών ορίων δικαιοδοσιών και συνόρων, δημόσιας και ιδιωτικής σφαίρας, φυσικού και ψηφιακού κόσμου. Έχει μάλιστα προκαλέσει έντονο γεωπολιτικό ανταγωνισμό αλλά και επέκταση σε περαιτέρω ύλη που αποδεικνύεται κρίσιμη για την άρτια επίδοση της ΤΝ, όπως η κυβερνοασφάλεια και η διακυβέρνηση των δεδομένων.

Τέλος (**Κεφάλαιο 4**), εξετάζεται ο ρόλος που μπορεί να έχει η ΤΝ στον πολιτικό σχεδιασμό και στη χάραξη πολιτικής μέσα από τη λεπτομερή καταγραφή που μπορούν να διενεργούν αρμόδια παρατηρητήρια. Αποτυπώνεται επίσης η προσπάθεια ενσωμάτωσης των θεμελιωδών δικαιωμάτων στα εναρμονισμένα πρότυπα και παρουσιάζεται το εγχείρημα των φορέων τυποποίησης να αναπτύξουν πρότυπα που να συμβάλλουν στην προστασία των θεμελιωδών δικαιωμάτων, ευαισθητοποιώντας του τεχνοεπιστήμονες σε ζητήματα ηθικής, δεοντολογίας και δικαίου ατομικών ελευθεριών. Επίσης παρουσιάζονται βασικοί δείκτες επιδόσεων που συνδέονται με την επίδραση της ΤΝ στον δημόσιο τομέα. Σκιαγραφείται η τρέχουσα διαμόρφωση τέτοιων δεικτών και αναδεικνύεται η σημασία της επικαιροποίησης και προσαρμογής των βασικών δεικτών επίδοσης (Key Performance Indicators) της ΤΝ, στην προοπτική του δημοσίου συμφέροντος και στην τεκμηριωμένη και μετρήσιμη προαγωγή της ευημερίας μέσω της ανθρωποκεντρικής χρήσης της.

Κλείνοντας, (**Κεφάλαιο 5**) παρουσιάζονται τα συμπεράσματα.

Κεφάλαιο 1

Επισκόπηση και ορισμοί της Τεχνητής Νοημοσύνης

- Για να κατανοηθεί η διαδρομή προς την προοπτική κατάλληλων μοντέλων διακυβέρνησης της τεχνητής νοημοσύνης θα παρατεθούν επιγραμματικά οι επιστημονικοί ορισμοί της και το επιστημονικό και τεχνολογικό ιστορικό υπόβαθρό της.
- Όσον αφορά την ιστορική αναδρομή στη δημιουργία του επιστημονικού κλάδου της ΤΝ, θα παρουσιαστεί το πρώτο κύμα καθώς ανάγεται σε πιο απομακρυσμένη, άρα και λιγότερο οικεία χρονική περίοδο.
- Επίσης θα αναλυθεί διεξοδικά η αποκρυστάλλωση κοινά αποδεκτού ορισμού υπό το πρίσμα της ΕΕ και ειδικότερα της πρότασης της Πράξης για την ΤΝ, δεδομένου ότι πρόκειται για το πρώτο πλήρες ρυθμιστικό πλαίσιο με οικουμενικό χαρακτήρα και πρόβλεψη κυρώσεων για τους παραβάτες.

1.1 Επισκόπηση

Η τεχνητή νοημοσύνη, ένας όρος που επινοήθηκε τη δεκαετία του 1950, βρίσκεται συνεχώς στα πρωτοσέλιδα από την έλευση του GPT-4 και άλλων παρόμοιων συστημάτων. Με μια σύντομη αναδρομή στο πρόσφατο παρελθόν διαπιστώνουμε ότι οι μεγάλες επιχειρήσεις επενδύουν σε αυτή ασταμάτητα και αυτό τους αποφέρει τεράστια κέρδη. Πρόκειται για μια τεχνολογία που ενσωματώνεται όλο και περισσότερο στην καθημερινότητά μας μέσω των αναζητήσεων στο Google, των ροών στο Facebook, της χρήσης της ψηφιακής βοηθού Siri της Apple και των συστάσεων από το Amazon και το Netflix, για να αναφέρουμε κάποια χαρακτηριστικά παραδείγματα.

Πολλές εταιρείες χρησιμοποιούν την ΤΝ για την εξατομίκευση των υπηρεσιών, την ενημέρωση των προϊόντων και τη βελτιστοποίηση των επιχειρηματικών λειτουργιών. Την ίδια στιγμή μαγνητίζει το ενδιαφέρον του κοινού μέσα από δυστοπικές κινηματογραφικές απεικονίσεις όπου κακόβουλα υπολογιστικά συστήματα έχουν καταλάβει τον πλήρη έλεγχο της ατομικής και κοινωνικής ζωής. Συγχρόνως πρωταγωνιστεί στη δημόσια συζήτηση με αντιπαραθέσεις σχετικά με τα θανατηφόρα αυτόνομα όπλα, τα αυτόνομα οχήματα, την τεχνολογία αναγνώρισης προσώπου και κυρίως την απειλή απώλειας θέσεων εργασίας.

Πώς όμως ξεκινήσαμε από επαναστατικές χρήσεις της ΤΝ προκειμένου τα κράτη και η δημόσια διοίκηση να λαμβάνουν αποφάσεις που μεγιστοποιούν τα οφέλη και μετριάζουν τις πιθανές αδικίες ή βλάβες, σε τομείς όπως η απασχόληση, οι προσλήψεις, η χρηματοδότηση και η προνοιακή πολιτική, η υγειονομική περίθαλψη και οι κρίσιμες υποδομές, και φτάσαμε στην έννοια του “υπαρξιακού κινδύνου” (x-risk, existential risk) που έχει αρχίσει να οικοδομείται γύρω από

τη χρήση της TN¹; Ποια είναι η διάθεση ρύθμισης της υπερ-ευφυούς τεχνητής νοημοσύνης, που πολλοί θεωρούν ότι πλησιάζει συντομότερα από ό,τι προβλεπόταν, τονίζοντας τον κίνδυνο τα αυτόνομα συστήματα να αναπτύξουν δικούς τους επιμέρους στόχους, που κλιμακώνονται από τη χειραγώγηση ανθρώπων μέχρι τη διακινδύνευση της ανθρώπινης ύπαρξης;

Η εργασία αποσκοπεί να φωτίσει το μέγεθος του διακυβεύματος της ρύθμισης της TN όταν μέχρι και οι ίδιοι οι επιχειρηματίες που επενδύουν σε αυτή εμφανίζονται να προβληματίζονται σχετικά με την υπεύθυνη και ηθική χρήση της. Χαρακτηριστική ήταν η ανοιχτή επιστολή που πολλοί από αυτούς, μεταξύ των οποίων ήταν ο Elon Musk και ο συνιδρυτής της Apple Steve Wozniak, προσυπέγραψαν. Η επιστολή δημοσιεύτηκε στις 22 Μαρτίου 2023 από το Future of Life Institute, και καλούσε² όλα τα εργαστήρια τεχνητής νοημοσύνης να σταματήσουν για τουλάχιστον 6 μήνες την εκπαίδευση συστημάτων τεχνητής νοημοσύνης πιο ισχυρών από το GPT-4. Τελικά υπογράφηκε από 30.000 ερευνητές.

Μάλιστα προσωπικότητες της έρευνας για την TN, όπως ο Bengio και ο πρωτοπόρος της βαθιάς μάθησης Geoff Hinton, υποστήριζαν πρόσφατα ότι το χρονοδιάγραμμα για την υπερ-ευφυή τεχνητή νοημοσύνη είναι συντομότερο από ό,τι προβλεπόταν, τονίζοντας τον κίνδυνο τα αυτόνομα συστήματα να αναπτύξουν τους δικούς τους επιμέρους στόχους, από τη χειραγώγηση ανθρώπων μέχρι τη διακινδύνευση της ανθρώπινης ύπαρξης³. Σε απάντηση στην αυξανόμενη πίεση, οι ανώτερες ηγεσίες των Google DeepMind, Anthropic και OpenAI υπέγραψαν μια ηπιότερη δήλωση από το Κέντρο για την Ασφάλεια της Τεχνητής Νοημοσύνης (Center for AI Safety), στην οποία αναφέρεται ότι “ο μετριασμός του κινδύνου εξαφάνισης από την Τεχνητή Νοημοσύνη θα πρέπει να αποτελεί παγκόσμια προτεραιότητα μαζί με άλλους κινδύνους κοινωνικής κλίμακας, όπως οι πανδημίες και οι πυρηνικοί πόλεμοι”⁴.

Ωστόσο, μια διαφορετική ανάγνωση της παραπάνω κίνησης του Future of Life Institute ήταν η διαπίστωση ότι πολλοί από τους υποτιθέμενους υπογράφαντες είτε το διέψευσαν, είτε κατηγορήθηκαν για αποπροσανατολισμό της κοινής γνώμης προς μια δυστοπική κινδυνολογία αντί για την επικέντρωση στους ορατούς και υπαρκτούς κινδύνους της TN, όπως οι ρατσιστικές ή σεξιστικές προκαταλήψεις που αποδεδειγμένα αναπαράγονται⁵. Επιπλέον σε άρθρο⁶ που δημοσιεύτηκε στις 28 Σεπτεμβρίου 2023 αναφέρεται ότι πολλοί από τους υπογράφαντες, όταν ερωτήθηκαν έξι μήνες αργότερα, δήλωσαν ότι δεν περίμεναν πραγματικά πως θα ισχύσει ποτέ κάτι τέτοιο και απεναντίας ορισμένοι είχαν επιταχύνει τη σχετική έρευνα αντί να την αναστείλουν.

Αυτή ακριβώς η αντίφαση είναι που πυροδοτεί περισσότερο τη συζήτηση για το λεγόμενο “ξέπλυμα” της ηθικής της τεχνητής νοημοσύνης. Υποστηρίζεται από μια μερίδα ότι η συζήτηση περί ηθικής TN ενδεχομένως αποτελεί αφήγημα που εξυπηρετεί όσους δεν θέλουν να δεσμεύονται από ένα αυστηρό νομικό πλαίσιο και προτιμούν μία χαλαρή ρύθμιση της TN σε επίπεδο ηθικών δεσμεύσεων και αυτορρύθμισης από τον ιδιωτικό τομέα⁷.

Ειδικότερα ως προς τα επιχειρήματα περί υπαρξιακού κινδύνου “x-risk”, οι επικριτές τους υποστηρίζουν ότι παραμένουν εικασίες. Για παράδειγμα, ο Francois Chollet, ερευνητής της Google AI και ένας από τους κύριους αρχιτέκτονες των TensorFlow και Keras, υποστήριξε ότι: “Δεν υπάρχει κανένα μοντέλο ή τεχνική τεχνητής νοημοσύνης που θα μπορούσε να αποτελέσει κίνδυνο εξαφάνι-

¹Nathan Benaich, Air Street Capital, State of AI Report 2023, σελ. 141, 12 Οκτωβρίου 2023, <https://www.stateof.ai/>

²<https://futureoflife.org/open-letter/pause-giant-ai-experiments/>, 22 Μαρτίου 2023

³AI Experts Call For Policy Action to Avoid Extreme Risks, TIME, 24 Οκτωβρίου 2023, <https://time.com/6328111/open-letter-ai-policy-action-avoid-extreme-risks/>

⁴<https://www.safe.ai/statement-on-ai-risk>

⁵<https://www.theguardian.com/technology/2023/mar/31/ai-research-pause-elon-musk-chatgpt>, 1 Απριλίου 2023

⁶<https://www.wired.com/story/fast-forward-elon-musk-letter-pause-ai-development/>, 28 Σεπτεμβρίου 2023

⁷Wagner Ben, “Ethics As An Escape From Regulation, From “Ethics-Washing” To Ethics-Shopping?.” BEING PROFILED: COGITAS ERGO SUM, Amsterdam University Press, 2018, 84-89, https://mediarep.org/bitstream/handle/doc/14203/Being-profiled_84-88_Wagner_Ethics_as_an_escape.pdf?sequence=1

σης για την ανθρωπότητα... ούτε καν αν προεκτείνουμε τις δυνατότητες πολύ μακριά στο μέλλον”⁸. Ο Yann LeCun (καθ. NYU, επιστημονικός επικεφαλής στη Meta) υποστήριξε ότι υπερεκτιμούμε την ωριμότητα των σημερινών συστημάτων τεχνητής νοημοσύνης, λέγοντας ότι “μέχρι να έχουμε ένα στοιχειώδες σχέδιο για τεχνητή νοημοσύνη ακόμη και σε επίπεδο σκύλου (πόσο μάλλον σε ανθρώπινο επίπεδο), το να συζητάμε πώς θα την κάνουμε ασφαλή είναι πρόωρο”⁹.

Ως αντίλογος στον υπαρξιακό κίνδυνο επισημαίνεται επίσης¹⁰ η ύπαρξη ηθικών κατευθυντηρίων γραμμών, όπως αυτές που εκτίθενται στο έγγραφο της Ευρωπαϊκής Επιτροπής “Κατευθυντήριες γραμμές δεοντολογίας για αξιόπιστη τεχνητή νοημοσύνη”¹¹ που συνέταξε η Ομάδα Εμπειρογνομόνων Υψηλού Επιπέδου το 2019, οι οποίες περιέχουν ισχυρές αρχές και διευκρινίσεις ευθυγραμμισμένες με τις κοινωνικές προσδοκίες και την τρέχουσα κατάσταση της συζήτησης για τη δεοντολογία της ΤΝ. Τονίζεται ότι αυτές οι συστάσεις δεοντολογίας λειτουργούν σε συνδυασμό με τη νομοθεσία της ΕΕ, αναγνωρίζοντας την ανάγκη για πρόβλεψη τόσο δικαίου όσο και δεοντολογίας στον τομέα της ΤΝ. Οι κατευθυντήριες γραμμές χρησιμεύουν ως σημείο αναφοράς για τη διαμόρφωση νέας νομοθεσίας, για την ερμηνεία των υφιστάμενων νόμων και για τη σύσταση δεοντολογικών πρακτικών πέραν των νομικών απαιτήσεων.

Ειδικότερα για τις (ηθικές) κατευθυντήριες γραμμές της ΕΕ διευκρινίζεται ότι παρέχουν ένα ολοκληρωμένο πρότυπο και ένα κοινό εννοιολογικό πλαίσιο σε αντίθεση με μια πληθώρα αρχών ηθικής ΤΝ που έχουν διατυπωθεί και προκαλούν συγκεχυμένη και κατακερματισμένη προσέγγιση.

1.2 Ιστορικό της Τεχνητής Νοημοσύνης ως επιστήμης

Η δυσκολία στον ορισμό της ΤΝ συνδέεται στενά με την εξελισσόμενη φύση της ως επιστήμης και για τον λόγο αυτό μια ιστορική επισκόπηση της ανάπτυξης της είναι απαραίτητη για την καλύτερη κατανόησή της. Στην επισκόπηση μας θα εστιάσουμε στην παρουσίαση του πρώτου κύματος της ΤΝ καθώς ανάγεται σε πιο απομακρυσμένη, άρα και λιγότερο οικεία χρονική περίοδο. Επιπλέον επειδή τότε τέθηκαν οι βάσεις της συμβολικής ΤΝ πάνω σε κανόνες που είχαν προδιατυπωθεί από ανθρώπους¹² και αυτό προκάλεσε μια εννοιολογική μεταφορά (metaphore)¹³ με αποτέλεσμα την ανθρωπομορφική προσέγγιση της ΤΝ, γενεσιουργό αιτία πολλών παρερμηνειών με απώτατη τον υπαρξιακό κίνδυνο του ανθρώπου. Κρίθηκε συνεπώς σημαντικό να επισκεφθούμε τις, από ιστορική άποψη, επιστημονικές ρίζες αυτής της λογικής πλάνης ως προς τη φύση της ΤΝ.

Η γέννηση της τεχνητής νοημοσύνης ως επιστημονικού κλάδου ανάγεται στο 1956, κατά τη διάρκεια ενός θερινού σχολείου στο Dartmouth College στο New Hampshire των ΗΠΑ από μία ομάδα επιστημόνων. Μεταξύ αυτών ήταν ο McCarthy που θεωρείται ότι καθιέρωσε τον όρο “τεχνητή νοημοσύνη” και ο Minsky, που την όρισε ως επιστήμη του σχεδιασμού μηχανών ικανών να κάνουν πράγματα που απαιτούν νοημοσύνη όταν γίνονται από ανθρώπους. Ωστόσο, η ανάπτυξη της ΤΝ είχε βαθιές ρίζες που προηγήθηκαν της επίσημης καθιέρωσής της ως επιστημονικού πεδίου. Η ιστορία της τεχνητής νοημοσύνης μπορεί να χωριστεί σε τρεις διακριτές φάσεις¹⁴. Αυτές οι τρεις φάσεις καταδεικνύουν πώς η αντίληψη για τη φύση της ΤΝ εξελίχθηκε με την πάροδο του χρόνου, ξεκινώντας από μύθους, θρύλους και θεωρητικές προσεγγίσεις και καταλήγοντας σε ένα σαφώς καθορισμένο επιστημονικό πεδίο.

⁸<https://twitter.com/fchollet/status/1664507791913140225>

⁹<https://www.cnbc.com/2023/06/15/ai-is-not-even-at-dog-level-intelligence-yet-meta-ai-chief.html>

¹⁰Floridi Luciano, Establishing the rules for building trustworthy AI, Nat Mach Intell 1, 261–262, 2019, <https://doi.org/10.1038/s42256-019-0055-y>

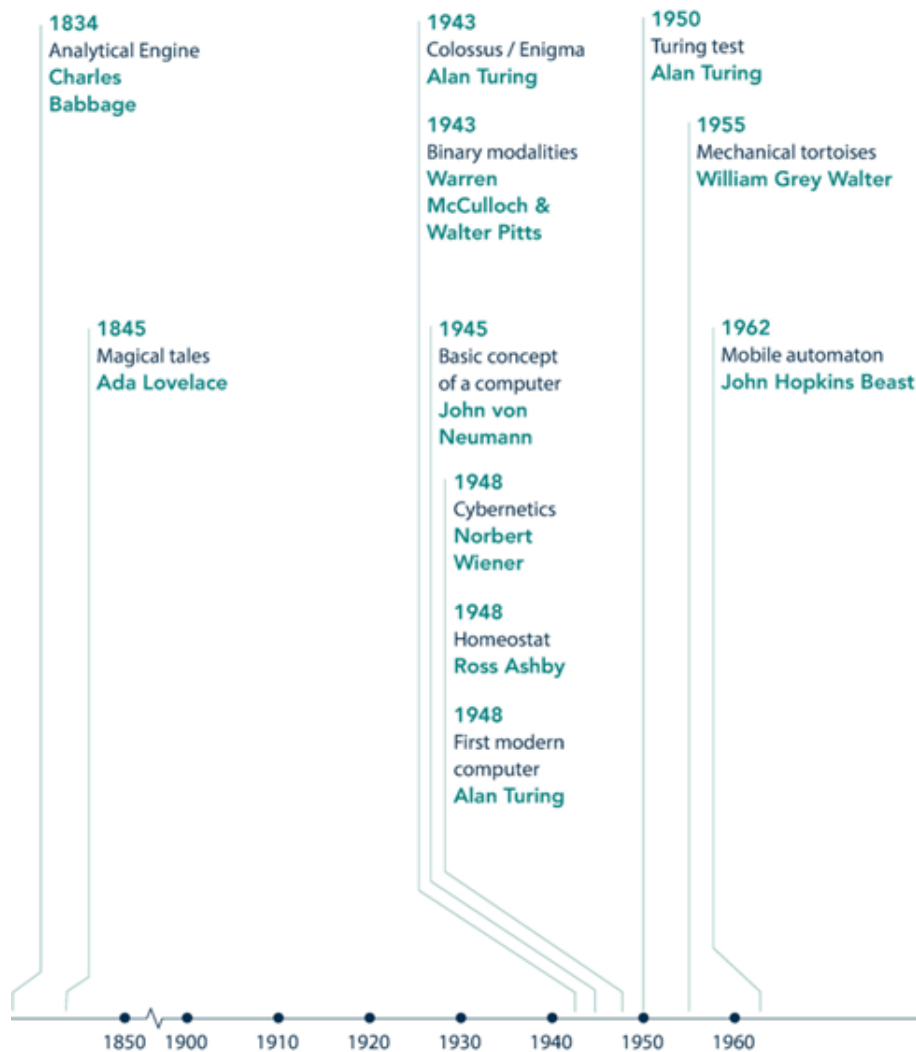
¹¹<https://digital-strategy.ec.europa.eu/en/node/1950>

¹²Λίλιαν Μήτρου (επιμ.), ΜΠΟΡΕΙ Ο ΑΛΓΟΡΙΘΜΟΣ ..., 2023, Πανεπιστημιακές εκδόσεις Κρήτης, σελ. 256

¹³Johanna Teresa Wallenborn, AI as a flying blue brain? How metaphors influence our visions of AI, 17 May 2022, <https://www.hiig.de/en/ai-metaphors/>

¹⁴Sheikh, H., Prins, C., Schrijvers, E. (2023). Artificial Intelligence: Definition and Background. In: Mission AI. Research for Policy. Springer, Cham. https://doi.org/10.1007/978-3-031-21448-6_2

Θα μπορούσαμε να δούμε μια πρώτη φάση της ΤΝ στους διάφορους μύθους και λαϊκούς θρύλους για τεχνητή ζωή. Πρόκειται για αρχαίες και μεσαιωνικές αναπαραστάσεις τεχνητών οντοτήτων καθώς και δοξασίες για σκεπτόμενες μηχανές. Στον ρου της ιστορίας, διάφοροι πολιτισμοί και μυθολογίες έχουν περιγράψει μη ανθρώπινες νοήμονες οντότητες που μοιάζουν με αυτό που σήμερα αναγνωρίζουμε ως τεχνητή νοημοσύνη. Πολλοί αρχαίοι ελληνικοί μύθοι απεικόνιζαν τέτοιους χαρακτήρες όπως ο Τάλως, δημιούργημα του Δαίδαλου που περιγράφεται σαν ένα ρομπότ σε καθημερινή περιπολία που προστάτευε την Κρήτη, εκσφενδονίζοντας πέτρες στα πλοία που πλησίαζαν. Ο Τάλως αποτέλεσε έμπνευση για την ονομασία ενός σύγχρονου ρομποτικού εξωσκελετού που σχεδιάστηκε για τον αμερικανικό στρατό¹⁵. Παρόμοια θέματα εμφανίζονται στις μυθολογίες άλλων πολιτισμών, όπως για παράδειγμα ο εσθονικός θρύλος για το Kratt, ένα μαγικό πλάσμα φτιαγμένο από άχυρο και οικιακά αντικείμενα. Μάλιστα η εθνική στρατηγική της Εσθονίας για την ΤΝ έχει ονομαστεί National AI “Kratt” Strategy¹⁶.



Εικόνα 1.1. Χρονολόγιο θεωριών ΤΝ, Mission AI, Springer, 2023

Η επόμενη φάση στην εξέλιξη της σκέψης για την τεχνητή νοημοσύνη σηματοδοτήθηκε από την καθιέρωση των θεωρητικών βάσεων για τους υπολογιστές και οδήγησε στην ανάπτυξη της ΤΝ ως διακριτού επιστημονικού κλάδου. Εμφανίστηκε κατά την εποχή των σημαντικών στοχαστών

¹⁵[https://en.wikipedia.org/wiki/TALOS_\(uniform\)](https://en.wikipedia.org/wiki/TALOS_(uniform))

¹⁶<https://e-estonia.com/new-e-estonia-factsheet-national-ai-kratt-strategy/>

όπως ο Γαλιλαίος, ο Νεύτων και ο Ντεκάρτ των οποίων η κοσμοθεωρία συνοδεύτηκε από την κατασκευή διαφόρων νέων μηχανών στις οποίες θα μπορούσε να μεταλαμπαδευτεί νόηση. Αυτή η περίοδος ήταν που σηματοδότησε τη μετάβαση από τις μυθολογικές στις μηχανικές προσεγγίσεις σχετικά με τη δημιουργία της ΤΝ. Αξιοσημείωτος σταθμός αυτής της εποχής υπήρξε η ανακάλυψη του προγραμματισμού από την Ada Lovelace, κόρη του Λόρδου Βύρωνα (1815-52), πρωτοπόρο μαθηματικό. Σε συνεργασία με τον Charles Babbage, ο οποίος σχεδίασε την “Αναλυτική μηχανή” το 1834, μια συσκευή που αποσκοπούσε στη δημιουργία μαθηματικών και αστρονομικών πινάκων, η Lovelace δημοσίευσε τον πρώτο αλγόριθμο που δημιουργήθηκε για να εφαρμοστεί σε αυτή την υπολογιστική μηχανή και διέβλεψε τις δυνατότητες επέκτασης των μαθηματικών υπολογισμών σε πραγματικές εφαρμογές.

Η μετάβαση στην επόμενη περίοδο, κατά την οποία ξεκινούν πλέον συγκεκριμένες σκέψεις για υλοποίηση ΤΝ, σφραγίστηκε από το θεατρικό έργο “R.U.R.” (Rossum’s Universal Robots) του Τσέχου Karel Čapek το 1920. Έμπνευσή του υπήρξε ο εβραϊκός μύθος για ένα μηχανικό άνθρωπο (γκόλεμ) στη μεσαιωνική Πράγα. Με το έργο αυτό εισήγαγε τον όρο “ρομπότ”, λέξη που προέρχεται από την παλαιά σλαβική λέξη “rabota” και σημαίνει καταναγκαστική εργασία. Το έργο αντανακλά έναν συχνό και έμφυτο στην ανθρωπότητα φόβο για την τεχνητή νοημοσύνη καθώς ανθρωπόμορφοι συνθετικοί εργάτες (ανθρωποειδή) επαναστατούν εναντίον των δημιουργών τους, οδηγώντας τελικά στην πτώση του ανθρώπινου είδους. Το έργο αυτό συνέπεσε με την περίοδο που ξεκινούσε προοδευτικά ένα πιο συγκεκριμένο όραμα για τη δημιουργία ΤΝ.

Σε αυτή την τελική φάση η ΤΝ αναδύεται ως επιστημονικός κλάδος όπου οι ερευνητές άρχισαν να διερευνούν την ανάπτυξη ευφώνων μηχανών, θέτοντας τα θεμέλια για τον τομέα της ΤΝ όπως τον γνωρίζουμε σήμερα. Η έννοια της τεχνητής νοημοσύνης (ΤΝ) ως σκεπτόμενου υπολογιστή συνέβη παράλληλα με τη θεωρητική θεμελίωση και την κατασκευή των πρώτων υπολογιστών. Οι βασικές εξελίξεις κατά τη διάρκεια αυτής της περιόδου σημειώνονται κατά τον Δεύτερο Παγκόσμιο Πόλεμο όταν παρουσιάστηκε η ανάγκη για γρήγορους υπολογισμούς για τον εντοπισμό των τροχιών κατά τη διάρκεια των αεροπορικών επιδρομών. Η έρευνα αυτή έδωσε επίσης το έναυσμα για τον τομέα της κυβερνητικής, ο οποίος έθεσε ερωτήματα σχετικά με την αυτοματοποίηση και τον ανθρώπινο έλεγχο.

Οι McCulloch και Pitts σε άρθρο τους το 1943 αξιοποίησαν την εργασία του Turing και προσομοίωσαν τη λειτουργία του νευρικού συστήματος με ένα δίκτυο τυπικών νευρώνων. Αυτοί οι λογιστικοί νευρώνες ήταν στην πραγματικότητα διασυνδεδεμένα ηλεκτρονικά αυτόματα με όριο λειτουργίας 0/1. Συνέδεσαν τους υπολογιστές με τη θεωρία των νευρικών συνάψεων και της λογικής και απέδειξαν τη δυαδική λειτουργία σε διάφορους τομείς, συνδέοντας το “αληθές ή το ψευδές” της λογικής με τις καταστάσεις “on ή off” των νευρώνων και τις τιμές “0 και 1” των υπολογιστών. Αν και αυτό το έργο δεν είχε άμεσο αποτέλεσμα, ενέπνευσε αργότερα τον Johann von Neumann να δημιουργήσει την κλασική αρχιτεκτονική των υπολογιστών, εξισώνοντας τη λειτουργία ανθρώπων και μηχανών και υποστηρίζοντας ότι ο υπολογιστής αποτελείται από στοιχεία όπως ο κεντρικός επεξεργαστής, η μνήμη και οι συσκευές εισόδου-εξόδου.

Σημαντικός θεμελιωτής της θεωρίας της ΤΝ ήταν ο Norbert Wiener που επινόησε τον όρο “κυβερνητική” το 1948 ως “τη μελέτη του ελέγχου και της επικοινωνίας στα ζώα και τις μηχανές”. Οι βασικές έννοιες περιλάμβαναν τις αρχές του ελέγχου, της ανατροφοδότησης και της ικανότητας προσαρμογής της μελλοντικής συμπεριφοράς στην εμπειρία του παρελθόντος.

Ο Alan Turing, μαθηματικός και επικεφαλής της ομάδας που αποκρυπτογράφησε τη γερμανική κρυπτογραφική συσκευής Enigma, το 1943, θεωρείται πατέρας της ΤΝ εισάγοντας το 1950 το “παιχνίδι μίμησης”, το οποίο αργότερα έγινε γνωστό ως Turing test. Αυτό το πείραμα περιελάμβανε έναν υπολογιστή που προσποιούνταν τον άνθρωπο, με τον υπολογιστή να περνάει το τεστ όταν ένας άνθρωπος δεν μπορούσε να διακρίνει εάν επρόκειτο για υπολογιστή ή για άλλο άνθρωπο.

Με αυτές τις προόδους, οι επιστήμονες ήταν έτοιμοι να περάσουν από τη θεωρητική σκέψη στην πρακτική ανάπτυξη της τεχνολογίας ΤΝ στα εργαστήρια. Η επίσημη έναρξη αυτής της προ-

σπάθειας, που ονομάζεται πρώτο κύμα, εντοπίζεται στο έτος 1956, μέσω του θερινού ερευνητικού προγράμματος Dartmouth Summer Research Project on Artificial Intelligence. Με πρόσκληση του John McCarthy, στον οποίο αποδίδεται η επινόηση του όρου “τεχνητή νοημοσύνη” το 1956, και κατόπιν αίτησής του για χρηματοδότηση από το Ίδρυμα Rockefeller, μια ομάδα δέκα εκ περιτροπής επιστημόνων (συνολικά ήταν περισσότεροι) συγκεντρώθηκε για διάστημα οχτώ εβδομάδων και πρότειναν ότι κάθε πτυχή της ανθρώπινης νοημοσύνης θα μπορούσε να περιγραφεί με ακρίβεια ώστε να προσομοιωθεί με μια μηχανή. Το έργο αυτό καθιέρωσε την ερευνητική ατζέντα για την τεχνητή νοημοσύνη. Στην πρωτοβουλία αυτή συμμετείχαν εξέχουσες προσωπικότητες όπως οι John McCarthy, Marvin Minsky, Herbert Simon, John Nash και Arthur Samuel.

Κατά τη διάρκεια αυτής της περιόδου, η έρευνα για την TN επικεντρώθηκε στη δημιουργία προγραμμάτων για επιτραπέζια παιχνίδια, ιδίως ντάμα. Ο Arthur Samuel ανέπτυξε ένα πρόγραμμα ντάμας που τελικά κατάφερε να νικήσει τον άνθρωπο δημιουργό του. Ωστόσο, ο Wiener πίστευε ότι οι άνθρωποι θα έχαναν το ενδιαφέρον τους για τα παιχνίδια όταν οι υπολογιστές θα μπορούσαν να ξεπεράσουν τις ανθρώπινες επιδόσεις. Τα προγράμματα τεχνητής νοημοσύνης άρχισαν να αντιμετωπίζουν λογικά και εννοιολογικά προβλήματα. Για παράδειγμα, αναπτύχθηκε μια “Μηχανή Λογικής Θεωρίας” για να αποδείξει τα λογικά θεωρήματα του Bertrand Russell και κατάφερε να αποδείξει έναν σημαντικό αριθμό από αυτά.

Παρά τις επιτυχίες που ακολούθησαν, όπως το πρόγραμμα επίλυσης προβλημάτων με γενικό τρόπο “General Problem Solver” και το ρομπότ “Shakey the Robot” στο Ερευνητικό Ινστιτούτο του Στάνφορντ, οι πρακτικές εφαρμογές της Τεχνητής Νοημοσύνης αντιμετώπιζαν ζητήματα αξιοπιστίας και ακρίβειας. Για παράδειγμα, ένα σύστημα μετάφρασης ρωσικής επικοινωνίας παρήγαγε χιουμοριστικά λάθη. Αυτοί οι πρακτικοί περιορισμοί, σε συνδυασμό με τις περικοπές χρηματοδότησης και την έλλειψη ανακαλύψεων, συνέβαλαν σε αυξανόμενη απαισιοδοξία στη δεκαετία του 1970. Αυτή έφερε τον πρώτο “χειμώνα της τεχνητής νοημοσύνης”, ο οποίος σταμάτησε την πρόοδο του πρώτου κύματος της έρευνας για την τεχνητή νοημοσύνη.

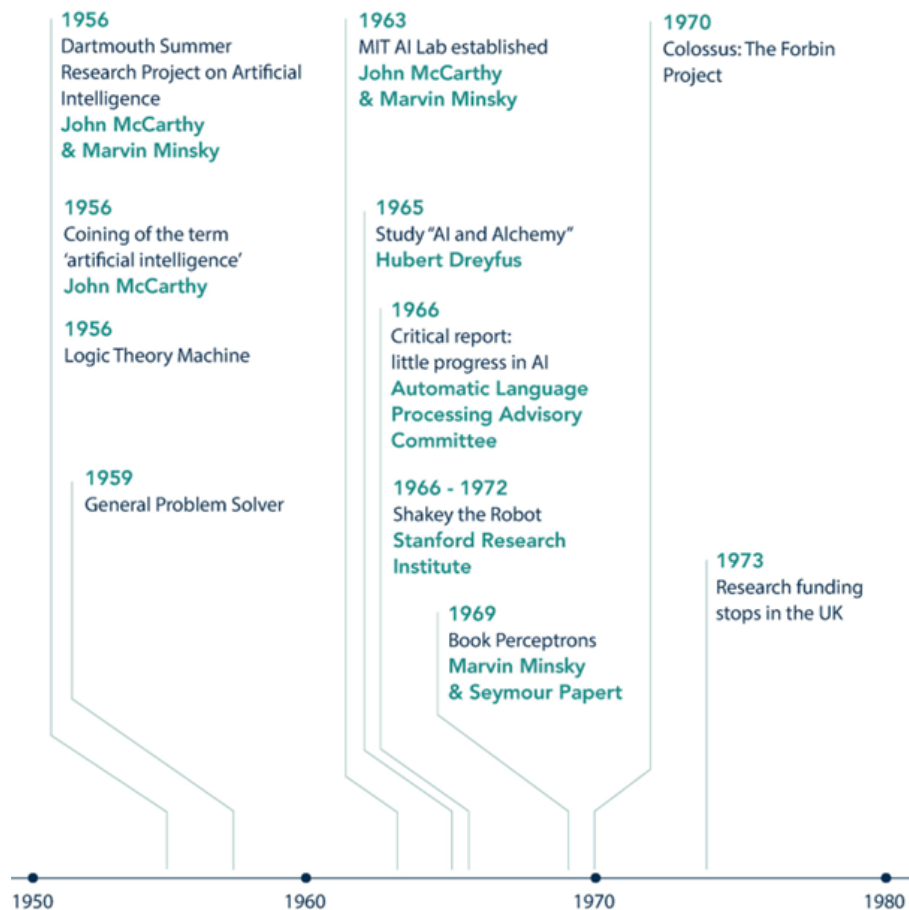
Κατά τη διάρκεια του πρώτου κύματος της τεχνητής νοημοσύνης (TN) στις δεκαετίες του 1950 και 1960, δύο διαφορετικές προσεγγίσεις της TN κέρδισαν την προβολή τους: η συμβολική ή λογική τεχνητή νοημοσύνη και τα τεχνητά νευρωνικά δίκτυα (ANNs – Connectionism).

Η **συμβολική** ή λογική τεχνητή νοημοσύνη βασίζεται στην κωδικοποίηση λογικών κανόνων με τύπους όπως “Αν X, τότε Y”. Εδράζεται στη χρήση ρητών κανόνων που εκφράζονται με ανθρώπινα σύμβολα για να καθοδηγούν τη συλλογιστική του υπολογιστή. Μια από τις βασικές εφαρμογές αυτής της προσέγγισης εμφανίστηκε τη δεκαετία του 1970 με τη μορφή των “συστημάτων εμπειρογνομόνων”. Τα συστήματα αυτά έχουν σχεδιαστεί για να προσομοιώνουν την ανθρώπινη εμπειρογνομοσύνη εφαρμόζοντας προκαθορισμένους κανόνες.

Τα **τεχνητά νευρωνικά δίκτυα (ANN)** είναι η δεύτερη προσέγγιση, γνωστή ως συνδεδετισμός (connectionism), και είναι εμπνευσμένα από τη λειτουργία των νευρώνων στον ανθρώπινο εγκέφαλο. Τα ANN αποτελούνται από τεχνητούς νευρώνες οργανωμένους σε δίκτυα για τη λήψη και τη μετάδοση πληροφοριών. Αυτή η προσέγγιση δεν βασίζεται σε ρητά καθορισμένους κανόνες, αλλά αντίθετα μαθαίνει από μεγάλες ποσότητες δεδομένων για να εντοπίζει μοτίβα και να κάνει προβλέψεις.

Στο πρώτο κύμα της τεχνητής νοημοσύνης, η συμβολική τεχνητή νοημοσύνη ήταν η κυρίαρχη προσέγγιση. Η αντιπαλότητα μεταξύ των κοινοτήτων της συμβολικής TN και του συνδεδετισμού, μαζί με τις ανησυχίες σχετικά με τους περιορισμούς και την αξιοπιστία των νευρωνικών δικτύων, συνέβαλε στην έναρξη του πρώτου “χειμώνα της TN”, που χαρακτηρίστηκε από μειωμένη χρηματοδότηση. Παρόλα αυτά, τα νευρωνικά δίκτυα θα επανεμφανίζονταν τελικά ως κυρίαρχη δύναμη στον τομέα της TN, ιδίως με την ανάπτυξη τεχνικών βαθιάς μάθησης τις επόμενες δεκαετίες.

Το δεύτερο κύμα ξεκίνησε τη δεκαετία του 1980, εν μέρει λόγω του διεθνούς ανταγωνισμού μεταξύ Ιαπωνίας, ΗΠΑ και Ευρώπης. Αυτό δημιούργησε τα συστήματα εμπειρογνομόνων, τις πρώτες σημαντικές εμπορικές εφαρμογές της TN. Το τρίτο κύμα ξεκίνησε τη δεκαετία του 1990 με



Εικόνα 1.2. Η ΤΝ ως επιστημονικός κλάδος (πρώτο κύμα), Mission AI, Springer, 2023

σημαντικά επιτεύγματα στη συμβολική ΤΝ, αλλά απέκτησε σωστή δυναμική μόνο μερικά χρόνια αργότερα λόγω των εξελίξεων στον τομέα της μηχανικής μάθησης και του υποπεδίου της, της βαθιάς μάθησης.

Είναι σαφές ότι η ταχεία εξάπλωση της ΤΝ τα τελευταία χρόνια έχει τις ρίζες της στην επιστημονική έρευνα. Μεγάλες εταιρείες όπως η Google έσπευσαν στη συνέχεια να προσλάβουν ταλαντούχους ερευνητές, αλλά αυτοί που ευθύνονται για τις σημαντικότερες ανακαλύψεις είναι οι επιστήμονες στα πανεπιστήμια. Εκτός από τα ακαδημαϊκά ορόσημα, δύο άλλοι παράγοντες διέπουν την άνοδο και την εφαρμογή της τεχνητής νοημοσύνης. Αφενός έχουμε την αύξηση της επεξεργαστικής ισχύος, όπως συνοψίζεται στο νόμο του Moore, σύμφωνα με τον οποίο ο αριθμός των τρανζίστορ σε ένα τσιπ διπλασιάζεται περίπου κάθε δύο χρόνια, όπως παρατηρείται σταθερά στη βιομηχανία υπολογιστών εδώ και δεκαετίες. Αυτό σημαίνει ότι όλο και περισσότερη υπολογιστική ισχύς γίνεται διαθέσιμη, ενώ οι τιμές συνεχίζουν να μειώνονται. Εξ ου και το γεγονός ότι τα σημερινά smartphones ξεπερνούν την υπολογιστική ισχύ των καλύτερων υπολογιστών που υπήρχαν μόλις πριν από μερικές δεκαετίες.

Αφετέρου, ένα περαιτέρω άλμα στην εν λόγω ισχύ προήλθε από τη βιομηχανία τσιπ, χρησιμοποιώντας μονάδες επεξεργασίας γραφικών (GPU) αντί για τις κλασικές κεντρικές μονάδες επεξεργασίας (CPU). Οι GPU αναπτύχθηκαν αρχικά για πολύπλοκα γραφικά στη βιομηχανία παιχνιδιών, αλλά στη συνέχεια διαπιστώθηκε ότι επιτρέπουν πολύ περισσότερους παράλληλους υπολογισμούς και στα συστήματα τεχνητής νοημοσύνης, ενώ από το 2015 άρχισαν να χρησιμοποιούνται και οι μονάδες επεξεργασίας τανυστών (tensor processing units, TPU) που έχουν σχεδιαστεί ειδικά για εφαρμογές ΜΛ.

Συνεπώς η τριάδα που απαρτίζεται από τα επιστημονικά επιτεύγματα, την αυξημένη υπολο-

γιστική ισχύ και την αύξηση των δεδομένων επέτρεψε στην Τεχνητή Νοημοσύνη να εκτοξευθεί πρόσφατα και αυτό οφείλεται κυρίως στην εφαρμογή της μηχανικής μάθησης ως μέρους της προσέγγισης των νευρωνικών δικτύων και περαιτέρω στην ανάπτυξη της βαθιάς μάθησης¹⁷.

1.3 Ιστορική αναδρομή σε ορισμούς TN και σχετικές έννοιες

Μια περιεκτική σύνοψη του τι είναι και πού βρίσκεται η Τεχνητή Νοημοσύνη σήμερα, δίνει η Karen Hao¹⁸, συντάκτρια τεχνητής νοημοσύνης στο MIT Technology Review. Γράφει ότι η τεχνητή νοημοσύνη αναφέρεται σε μηχανές που μπορούν να μαθαίνουν, να σκέφτονται και να ενεργούν από μόνες τους. Μπορούν να παίρνουν τις δικές τους αποφάσεις όταν έρχονται αντιμέτωπες με νέες καταστάσεις, με τον ίδιο τρόπο που μπορούν οι άνθρωποι και τα ζώα. Επισημαίνει ότι στην παρούσα φάση, η συντριπτική πλειοψηφία των εξελίξεων και εφαρμογών TN αναφέρονται σε μια κατηγορία αλγορίθμων που είναι γνωστή ως μηχανική μάθηση. Αυτοί οι αλγόριθμοι χρησιμοποιούν κυρίως τη στατιστική για να βρουν μοτίβα σε τεράστιες ποσότητες δεδομένων. Στη συνέχεια χρησιμοποιούν αυτά τα μοτίβα για να κάνουν προβλέψεις σχετικά με τον χρήστη, όπως ποιες είναι οι προτιμήσεις του στο Netflix, τι λέει όταν μιλάει στην Alexa ή αν τυχόν έχει κάποιον κακοήθη όγκο με βάση τη μαγνητική του τομογραφία.

Η μηχανική μάθηση, και το υποσύνολό της, η βαθιά μάθηση είναι η βάση πολλών σημαντικών επιτευγμάτων, όπως η αναγνώριση προσώπου, η δημιουργία αληθοφανών φωτογραφιών και φωνής και το AlphaGo¹⁹, το πρόγραμμα που νίκησε τον καλύτερο παίκτη του πολύπλοκου παιχνιδιού Go²⁰. Απώτερη φιλοδοξία είναι να αναπτυχθεί κάτι που να μοιάζει με την ανθρώπινη νοημοσύνη, η οποία συχνά αναφέρεται ως “Τεχνητή Γενική Νοημοσύνη” ή “AGI”, κάτι που ορισμένοι ειδικοί πιστεύουν ότι θα επιτευχθεί σύντομα ενώ άλλοι θεωρούν ότι είναι ακόμη πολύ μακριά.

Η τεχνητή νοημοσύνη μπορεί να έχει κατακτήσει το σκάκι και το Go, αλλά από άλλες απόψεις εξακολουθεί να έχει “νοημοσύνη” νηπίου. Υπό αυτή την έννοια, ο ορισμός της TN εξελίσσεται συνεχώς. Αυτό που θα θεωρούνταν τεχνητή νοημοσύνη στο παρελθόν μπορεί να μην θεωρείται τεχνητή νοημοσύνη σήμερα. Η συντάκτρια καταλήγει ότι εξαιτίας αυτού, τα όρια της TN μπορούν να προκαλέσουν μεγάλη σύγχυση και ο όρος συχνά διαστρεβλώνεται για να συμπεριλάβει κάθε είδους αλγόριθμο ή πρόγραμμα υπολογιστή ενώ αφήνει έναν υπαινιγμό για τον ρόλο που παίζουν οι εταιρείες της Silicon Valley στη διαρκή διάγκωση των δυνατοτήτων της τεχνητής νοημοσύνης για λόγους συμφέροντος. Τέλος, για να διευκολύνει τους αναγνώστες να εκτιμήσουν αν κάποιο σύστημα χρησιμοποιεί TN ή όχι, σχεδίασε ένα διάγραμμα ροής που να χωράει μόλις στη μία όψη ενός φακέλου (εικόνα 1.1).

Για μία λεπτομερέστερη αναδρομή στην έννοια και το ιστορικό της TN, ξεκινούμε με την παραδοχή ότι αν θέλουμε να ενσωματώσουμε την τεχνητή νοημοσύνη στην κοινωνία, πρέπει να κατανοήσουμε τι είναι και πώς έχει εξελιχθεί. Ο ορισμός της τεχνητής νοημοσύνης δεν είναι κάτι απλό και στην πραγματικότητα δεν υπάρχει ένας γενικά αποδεκτός ορισμός της έννοιας. Έχουν διατυπωθεί διάφοροι ορισμοί αλλά ακριβώς αυτή η ποικιλομορφία είναι εγγενής στην πολύπλευρη και διεπιστημονική φύση της τεχνητής νοημοσύνης.

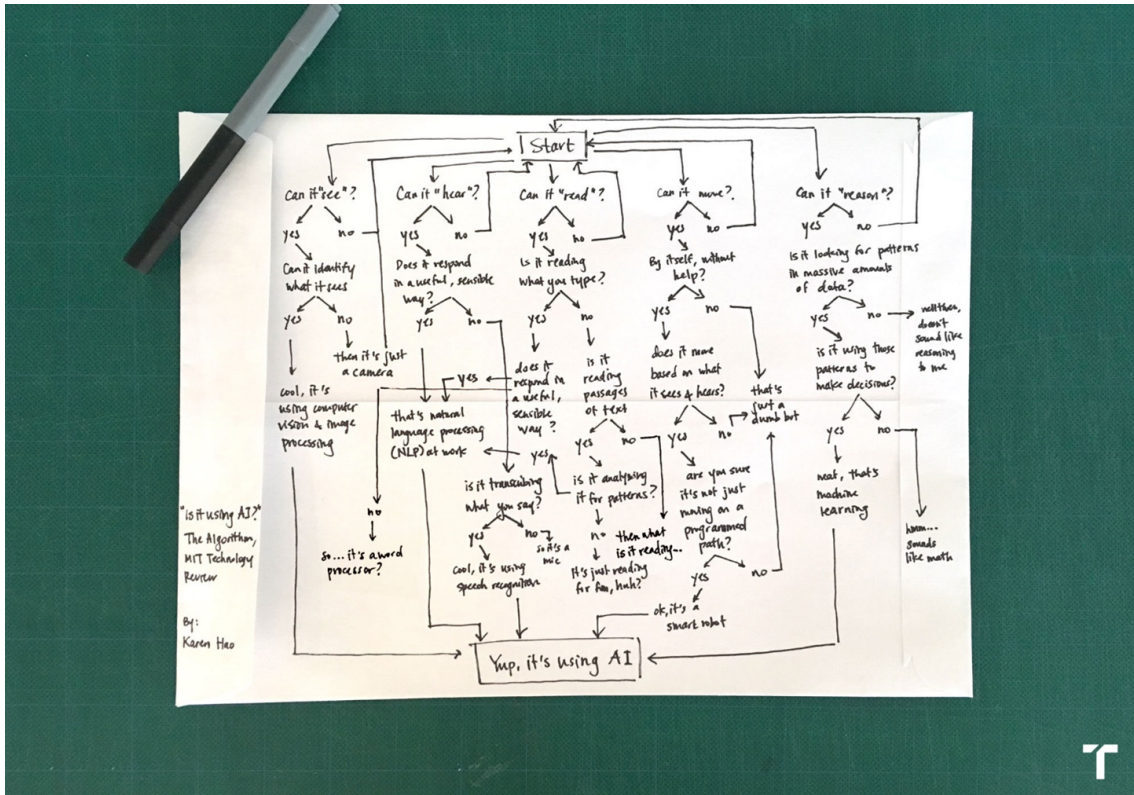
Ο ευρύτερος ορισμός της τεχνητής νοημοσύνης την εξισώνει με τους αλγορίθμους. Ωστόσο, οι αλγόριθμοι προϋπάρχουν και χρησιμοποιούνται ευρέως και εκτός του πεδίου της TN. Ο αλγόριθμος, που ετυμολογικά προέρχεται από τον Πέρση μαθηματικό του 9ου αιώνα Mohammed ibn Musa al-Kharizmi, αναφέρεται σε μια συγκεκριμένη οδηγία για την επίλυση ενός προβλήματος ή

¹⁷ Sheikh, H., Prins, C., Schrijvers, E. (2023). Artificial Intelligence: Definition and Background. In: Mission AI. Research for Policy. Springer, Cham. https://doi.org/10.1007/978-3-031-21448-6_2

¹⁸ Karen Hao, MIT Technology Review, 10 Νοεμβρίου 2018, <https://www.technologyreview.com/2018/11/10/139137/is-this-ai-we-drew-you-a-flowchart-to-work-it-out/>

¹⁹ <https://deepmind.google/technologies/alphago/>

²⁰ ACM Prize in Computing Awarded to AlphaGo Developer David Silver, 1 Απριλίου 2020, <https://www.acm.org/media-center/2020/april/acm-prize-2019>



Εικόνα 1.3. Τι είναι Τεχνητή Νοημοσύνη: διάγραμμα ροής, Karen Hao, MIT Technology Review, 2018

την εκτέλεση ενός υπολογισμού. Ο ορισμός της ΤΝ αποκλειστικά ως χρήση αλγορίθμων θα περιλάμβανε ένα ευρύ φάσμα δραστηριοτήτων, από τη χρήση μιας αριθμομηχανής τσέπης έως την εκτέλεση των οδηγιών μιας συνταγής μαγειρικής.

Αντίθετα, ένας αυστηρότερος ορισμός της τεχνητής νοημοσύνης δίνει έμφαση στους υπολογιστές που μιμούνται την ανθρώπινη νοημοσύνη. Ωστόσο, οι υποστηρικτές αυτής της σχολαστικής άποψης υποστηρίζουν ότι πολλές σημερινές εφαρμογές είναι υπερβολικά απλοϊκές για να θεωρηθούν αληθινή τεχνητή νοημοσύνη. Η χρήση αυτού του ορισμού θα προϋπέθετε την παραδοχή ότι η τεχνητή νοημοσύνη δεν υπάρχει επί του παρόντος, αναιρώντας ουσιαστικά την ίδια την ύπαρξή της.

Ο ορισμός της τεχνητής νοημοσύνης είναι ένα πολύπλοκο εγχείρημα καθώς αυτή περιλαμβάνει τη μίμηση πολύπλοκων ανθρώπινων δεξιοτήτων, οι οποίες δεν είναι πλήρως κατανοητές. Οι συνήθεις ορισμοί της συχνά την περιγράφουν ως μια τεχνολογία που επιτρέπει στις μηχανές να μιμούνται αυτές τις δεξιότητες, ωστόσο η έννοια των πολύπλοκων ανθρώπινων δεξιοτήτων δεν είναι συγκεκριμένη, καθιστώντας τη φύση της τεχνητής νοημοσύνης διαφορετική. Το ίδιο ισχύει αν προσθέσουμε σε αυτές και τη συνθήκη του πολύπλοκου περιβάλλοντος.

Για να δοθεί ένας πιο εύστοχος ορισμός, έχουν προταθεί διάφορες περιγραφές με βάση τις εργασίες που πρέπει να εκτελεστούν (tasks). Αυτοί οι ορισμοί υπογραμμίζουν την ικανότητα της ΤΝ να λειτουργεί κατάλληλα, να επιδεικνύει διορατικότητα²¹, να αντιλαμβάνεται, να επιδιώκει στόχους, να δρομολογεί ενέργειες και να μαθαίνει από την ανατροφοδότηση²². Ένας παρόμοιος ορισμός προτάθηκε από την Ομάδα Εμπειρογνομώνων Υψηλού Επιπέδου για την Τεχνητή Νοη-

²¹Nilsson, N. (2009). The Quest for Artificial Intelligence. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CB09780511819346>

²²Denk Werk. (2018). Artificial Intelligence in Nederland: Zelf Aan Het Stuur, https://denkwerk.online/media/1029/artificial_intelligence_in_nederland_juli_2018.pdf

μοσύνη (AI HLEG) της Ευρωπαϊκής Επιτροπής που αποδίδει την TN ως “συστήματα που επιδεικνύουν ευφυή συμπεριφορά αναλύοντας το περιβάλλον τους και αναλαμβάνοντας δράσεις, με κάποιο βαθμό αυτονομίας, για την επίτευξη συγκεκριμένων στόχων”²³.

Ωστόσο, ακόμη και αυτοί οι ορισμοί που βασίζονται σε καθήκοντα έχουν περιορισμούς, με όρους όπως “κάποιος βαθμός αυτονομίας” να παραμένουν κάπως ασαφείς. Περιλαμβάνουν επίσης φαινόμενα που πολλοί θα δίσταζαν να κατηγοριοποιήσουν ως TN, όπως ένας απλός θερμοστάτης. Αυτή η συσκευή είναι επίσης σε θέση να αντιλαμβάνεται (να μετρά τη θερμοκρασία του δωματίου), να επιδιώκει στόχους (την προγραμματισμένη θερμοκρασία), να ξεκινά ενέργειες (να ρυθμίζει τον θερμοστάτη) και να μαθαίνει από έναν βρόγχο ανατροφοδότησης (να σταματά μόλις επιτευχθεί η προγραμματισμένη θερμοκρασία). Παρ’ όλα αυτά ποιος θα ήταν διατεθειμένος να θεωρήσει ένα θερμοστάτη τεχνητή νοημοσύνη;

Ο δυσνόητος ορισμός της τεχνητής νοημοσύνης οφείλεται κυρίως στη μίμηση της ανθρώπινης νοημοσύνης, μια έννοια που δεν έχει ακόμη κατανοηθεί πλήρως από ερευνητές σε τομείς όπως η ψυχολογία, η επιστήμη της συμπεριφοράς και η νευρολογία. Η κατανόηση της τεχνητής νοημοσύνης απαιτεί την ταυτόχρονη διερεύνηση της ανθρώπινης νοημοσύνης, καθώς και οι δύο είναι στενά συνδεδεμένες μεταξύ τους και συνεξελίσσονται. Ένα κλασικό παράδειγμα είναι το σκάκι, όπου η TN έχει διαπρέψει από τη δεκαετία του 1990, παρόλο που κάποτε πίστευαν ότι η δημιουργία μιας μηχανής που παίζει σκάκι θα διείσδυε στον πυρήνα της ανθρώπινης διανοητικής λειτουργίας. Αυτό αναδεικνύει πόσο σύνθετο είναι να κατανοήσουμε τη φύση της ανθρώπινης νοημοσύνης. Η νίκη του Deep Blue της IBM επί του Garry Kasparov το 1997 οδήγησε αρχικά στην πεποίθηση ότι οι μηχανές θα μπορούσαν να ξεπεράσουν τους ανθρώπους σε διάφορες δραστηριότητες, καθώς το σκάκι θεωρούνταν το αποκορύφωμα της ανθρώπινης νοημοσύνης.

Ωστόσο, αυτή η προσέγγιση έχει έκτοτε αναθεωρηθεί. Το σκάκι αναγνωρίζεται πλέον ως ένα μαθηματικό πρόβλημα με πεπερασμένες εναλλακτικές λύσεις. Τα προγράμματα σκακιού μοιάζουν με αριθμομηχανές τσέπης, ικανά να επιλύουν πολύπλοκους υπολογισμούς που είναι δυσεπίλυτοι ακόμη και από πολύ έξυπνους ανθρώπους, όμως αυτό δεν τους καθιστά μια τεχνητή μορφή της ανθρώπινης νοημοσύνης.

Το “παράδοξο του Moravec”²⁴ επισημαίνει ότι εργασίες που αποτελούν πρόκληση για τους ανθρώπους, όπως το σκάκι, μπορεί να είναι εύκολες για τους υπολογιστές, ενώ φαινομενικά απλές εργασίες για τους ανθρώπους, όπως η αναγνώριση αντικειμένων και οι κινητικές δεξιότητες, δυσκολεύουν τους υπολογιστές. Οι άνθρωποι έχουν εξελιχθεί ώστε να υπερέχουν σε εργασίες απαραίτητες για την επιβίωσή τους σε συνδυασμό με τις αισθητηριακές τους εμπειρίες, καθώς η σωματικότητα και η πνευματικότητα συνυπάρχουν και αλληλοτροφοδοτούνται. Απεναντίας, οι ικανότητες της TN βασίζονται σε υπολογιστικές δεξιότητες και όχι στη βιολογική εξέλιξη.

Ο ορισμός της τεχνητής νοημοσύνης εξελίσσεται παράλληλα με τις διαρκώς εξελισσόμενες δυνατότητες των υπολογιστών, με αποτέλεσμα μόλις επιτευχθεί αυτό που κάποτε μας εντυπωσίαζε, να χαρακτηρίζεται ως “απλώς ένας υπολογισμός” και όχι ως πραγματική νοημοσύνη. Αυτό αποκάλεσε “Φαινόμενο της Τεχνητής Νοημοσύνης” η Pamela McCorduck²⁵ και σημαίνει ότι η TN περιλαμβάνει όποια τεχνολογία μάς εντυπωσιάζει σήμερα, και όταν δεν μας εντυπωσιάζει πλέον, αναφέρεται συνήθως ως “λογισμικό”.

Η αντίληψη ότι οι μηχανές, ιδίως η σύγχρονη τεχνολογία, μπορούν να εκτελούν εργασίες ισόδυναμα με τον άνθρωπο μπορεί να είναι παραπλανητική. Οι Agrawal, Gans και Goldfarb²⁶ προτείνουν εναλλακτικούς όρους όπως “μηχανές πρόβλεψης” για να αναγνωρίσουν ότι η τεχνητή νοημοσύνη αφορά κυρίως την πρόβλεψη αποτελεσμάτων και όχι την αναπαραγωγή όλων των πτυχών

²³High-Level Expert Group on Artificial Intelligence. (2019). A definition of AI: Main capabilities and scientific disciplines. European Commission, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56341

²⁴Moravec, H. (1988), *Mind Children: The future of robot and human intelligence*, Harvard University Press.

²⁵McCorduck, Pamela (2004), *Machines Who Think* (2η έκδοση), Natick, MA: AK Peters, Ltd., ISBN 1-56881-205-1

²⁶Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction machines: The simple economics of Artificial Intelligence*, Harvard Business Press.

της ανθρώπινης νοημοσύνης. Ο Edsger Dijkstra διατύπωσε παραστατικά την προσέγγιση αυτή με το σχόλιο “Το ερώτημα αν οι υπολογιστές μπορούν να σκέφτονται είναι ακριβώς όπως το ερώτημα αν τα υποβρύχια μπορούν να κολυμπούν”²⁷. Αυτό που κάνουν τα συγκεκριμένα σκάφη είναι παρόμοιο με αυτό που οι άνθρωποι ονομάζουν κολύμπι, αλλά το να το ονομάσουμε έτσι θα ήταν λάθος. Η τεχνητή νοημοσύνη φαίνεται να μιμείται την ανθρώπινη νοημοσύνη, αλλά ουσιαστικά λειτουργεί διαφορετικά.

Αυτή η προσέγγιση εξηγεί καλύτερα το παράδοξο Moravec που προαναφέρθηκε. Η αναγνώριση προσώπων είναι εύκολη για τους ανθρώπους, αλλά δύσκολη για τους υπολογιστές. Αυτό συμβαίνει επειδή η αναγνώριση των άλλων ήταν εξελικτικά κρίσιμη για την επιβίωσή μας και έτσι ο εγκέφαλός μας έχει μάθει να το κάνει χωρίς σκέψη. Η ικανότητα να παίζουμε σκάκι δεν ήταν απαραίτητη στην εξέλιξη και επομένως είναι πιο δύσκολο να την κατακτήσουμε καθώς απαιτεί ένα ορισμένο επίπεδο υπολογιστικής ικανότητας. Οι υπολογιστές δεν είναι προϊόν βιολογικής εξέλιξης, οπότε οι ικανότητές τους είναι διαφορετικές από αυτές των ανθρώπων. Μια σημαντική πτυχή αυτής της θεωρίας είναι ότι δεν πρέπει να προσπαθούμε να κατανοήσουμε την ΤΝ ως σύνολο τεχνολογιών, ούτε υπό το πρίσμα της ανθρώπινης νοημοσύνης ούτε με όρους ανθρωπομορφισμού. Ωστόσο, ο όρος “τεχνητή νοημοσύνη” έχει καθιερωθεί και είναι πλέον απίθανο να αντικατασταθεί.

Τέλος, υπάρχει η τάση η τεχνητή νοημοσύνη να ταυτίζεται με την εκάστοτε τελευταία εξέλιξη που συναντούμε σε αυτή την τεχνολογία. Η ΤΝ έχει αποκτήσει τεράστια δυναμική τα τελευταία χρόνια και ένας από τους σημαντικότερους μοχλούς για αυτό υπήρξε η πρόοδος σε έναν συγκεκριμένο τομέα του πεδίου, τη μηχανική μάθηση (ML) και τη βαθιά μάθηση (DL). Οι αλγόριθμοι ML και DL υπερέχουν στην αναγνώριση μοτίβων στα δεδομένα, που συχνά αναφέρονται ως “αλγόριθμοι αυτοεκμάθησης”, και έχουν οδηγήσει σε επιτεύγματα όπως η αναγνώριση προσώπων και η εξελιγμένη αναπαραγωγή παιχνιδιών.

Η τεχνητή νοημοσύνη θεωρείται ένας δυναμικός και πολυδιάστατος τομέας με απώτερο στόχο την κατανόηση και αναπαραγωγή όλων των ανθρώπινων διανοητικών δεξιοτήτων, οπότε σε αυτή την περίπτωση αναφέρεται ως Τεχνητή Γενική Νοημοσύνη (Artificial General Intelligence, AGI). Ωστόσο, η επίτευξη αυτού του στόχου σύμφωνα με πολλούς επιστήμονες μπορεί να απέχει αρκετά και ίσως να μην είναι καν εφικτός. Ο υφιστάμενος ορισμός της τεχνητής νοημοσύνης (ΤΝ) περιλαμβάνει κυρίως αυτό που ονομάζεται “στενή” ή “αδύναμη” ΤΝ. Αυτή η τεχνητή νοημοσύνη, η οποία χρησιμοποιείται ευρέως σήμερα, εξειδικεύεται σε συγκεκριμένες εργασίες, όπως η αναγνώριση εικόνας ή ομιλίας, και δεν προσεγγίζει το πλήρες φάσμα των ανθρώπινων γνωστικών ικανοτήτων που αντιπροσωπεύει η τεχνητή γενική νοημοσύνη. Η διάκριση αυτή δεν μειώνει τη σπουδαιότητα των σημερινών εφαρμογών τεχνητής νοημοσύνης, καθώς θέτουν τα δικά τους ουσιαστικά ζητήματα.

Λαμβάνοντας υπόψη δύο βασικούς παράγοντες, φαίνεται πιο ασφαλές να επιλέξουμε έναν ανοικτό ορισμό της ΤΝ. Πρώτον, ο περιορισμός του ορισμού σε μια συγκεκριμένη τεχνολογία, όπως η βαθιά μάθηση, θα παρέβλεπε κρίσιμα ζητήματα που αναδύονται από άλλους τομείς της ΤΝ. Δεύτερον, πολλές κυβερνητικές εφαρμογές της ΤΝ δεν βασίζονται σε προηγμένες τεχνικές της ΤΝ, πλην όμως εξακολουθούν να απαιτούν προσοχή για προφανείς λόγους νομικής και δεοντολογικής φύσης. Ένας ανοικτός ορισμός επιτρέπει ευελιξία και αναγνωρίζει ότι οι μελλοντικές εξελίξεις της ΤΝ μπορεί να προέλθουν από άλλους τομείς αυτής της τεχνολογίας.

Ο Pedro Domingos²⁸, επιστήμονας με ειδίκευση στη μηχανική μάθηση, επισημαίνει ότι η εστίασή μας στην πιθανή μελλοντική εμφάνιση της τεχνητής γενικής νοημοσύνης συχνά επισκιάζει την ήδη εκτεταμένη επιρροή της ειδικής ΤΝ στη ζωή μας. Υποστηρίζει ότι η πρωταρχική ανησυχία μας δεν πρέπει να είναι ότι οι υπολογιστές θα γίνουν πολύ έξυπνοι και θα καταλάβουν τον κόσμο, αλλά ότι είναι πολύ ανόητοι και έχουν ήδη καταλάβει τον κόσμο.

²⁷“The question of whether computers can think is just like the question of whether submarines can swim”, <https://history.computer.org/pioneers/dijkstra.html>

²⁸Domingos, P. (2017), The master algorithm: How the Quest for the ultimate learning machine will remake our world, Penguin Random House.

1.4 Ορισμός της TN στην Ευρωπαϊκή Ένωση

Η Ευρωπαϊκή Επιτροπή, μέσα από τον ενισχυμένο ρόλο που θέλει να αναδείξει για την Ευρωπαϊκή Ένωση στον παγκόσμιο ανταγωνισμό των εξελίξεων της TN, αναζήτησε διάφορους ορισμούς μέσα από ανακοινώσεις, προτάσεις της Ομάδας Εμπειρογνομόνων Υψηλού Επιπέδου για την TN (AI HLEG) και τη διατύπωση που συμπεριέλαβε στην πρόταση της Πράξης για την TN που υπέβαλε στις 21 Απριλίου 2021, προκειμένου να ψηφιστεί ως κανονισμός. Ακολουθεί η παρουσίαση αυτών των ορισμών, που διατυπώθηκαν με πρωτοβουλία της Ευρωπαϊκής Επιτροπής, με χρονολογική σειρά.

1.4.1 Ορισμός TN: Ανακοίνωση της Επιτροπής, Απρίλιος 2018

Ο αρχικός ορισμός της τεχνητής νοημοσύνης όπως προτάθηκε στην Ανακοίνωση της Επιτροπής για την TN²⁹ το 2018, ανέφερε ότι:

“Η τεχνητή νοημοσύνη (TN) αναφέρεται σε συστήματα που χαρακτηρίζονται από ευφυή συμπεριφορά, αναλύοντας το περιβάλλον τους και ενεργώντας —με κάποιο βαθμό αυτονομίας— για την επίτευξη συγκεκριμένων στόχων. Τα συστήματα που λειτουργούν βάσει τεχνητής νοημοσύνης μπορούν να βασίζονται αποκλειστικά σε λογισμικό, ενεργώντας στον εικονικό κόσμο (π.χ. βοηθεί φωνής, λογισμικό ανάλυσης εικόνας, μηχανές αναζήτησης, συστήματα αναγνώρισης ομιλίας και προσώπου) ή η τεχνητή νοημοσύνη μπορεί να ενσωματωθεί σε συσκευές υλισμικού (π.χ. προηγμένα ρομπότ, αυτόνομα αυτοκίνητα, συστήματα μη επανδρωμένων αεροσκαφών ή εφαρμογές του Διαδικτύου των Πραγμάτων).”

1.4.2 Ορισμός TN: Ομάδα εμπειρογνομόνων (HLEG), Απρίλιος 2019

Τον Ιούνιο του 2018 συστάθηκε από την Ευρωπαϊκή Επιτροπή μια ανεξάρτητη ομάδα εμπειρογνομόνων υψηλού επιπέδου για την τεχνητή νοημοσύνη (High Level Expert Group). Στο έγγραφο³⁰ που δημοσίευσε τον Απρίλιο του 2019, με τίτλο “Ορισμός της τεχνητής νοημοσύνης: κύριες δυνατότητες και επιστημονικά πεδία”, πρότεινε τον ακόλουθο επικαιροποιημένο ορισμό:

“Τα συστήματα τεχνητής νοημοσύνης (TN) είναι συστήματα λογισμικού (ή ενδεχομένως και υλισμικού) που σχεδιάζονται από ανθρώπους (οι άνθρωποι σχεδιάζουν άμεσα συστήματα TN, αλλά μπορεί επίσης να χρησιμοποιούν τεχνικές TN για τη βελτιστοποίηση του σχεδιασμού τους) και, βάσει ενός δεδομένου σύνθετου στόχου, ενεργούν στην υλική ή ψηφιακή διάσταση με το να αντιλαμβάνονται το περιβάλλον τους μέσω της απόκτησης δεδομένων, να ερμηνεύουν τα δομημένα ή αδόμητα δεδομένα που έχουν συλλεχθεί, να προβαίνουν σε συλλογισμούς με βάση τις γνώσεις ή να επεξεργάζονται τις πληροφορίες που εξάγονται από αυτά τα δεδομένα και να αποφασίζουν ποια είναι η βέλτιστη ενέργεια (ή οι βέλτιστες ενέργειες) που θα πρέπει να εκτελέσουν για να επιτύχουν τον δεδομένο στόχο. Τα συστήματα TN μπορεί είτε να χρησιμοποιούν συμβολικούς κανόνες είτε να μαθαίνουν ένα αριθμητικό μοντέλο, και μπορεί επίσης να προσαρμόζουν τη συμπεριφορά τους με το να αναλύουν πώς επηρεάζεται το περιβάλλον από τις προηγούμενες ενέργειές τους.”

1.4.3 Ορισμός TN: Πράξη για την TN, Απρίλιος 2021

Στην τελευταία της πρόταση για την Πράξη για την TN που υπέβαλε η Ευρωπαϊκή Επιτροπή στις 21 Απριλίου 2021, προκειμένου να ψηφιστεί ως κανονισμός για τη θέσπιση εναρμονισμένων

²⁹ Ανακοίνωση της Επιτροπής, COM/2018/237 final, Τεχνητή νοημοσύνη για την Ευρώπη, <https://eur-lex.europa.eu/legal-content/EL/TXT/?uri=CELEX:52018DC0237>

³⁰ High-Level Expert Group on Artificial Intelligence, 2019, A definition of AI: Main capabilities and scientific disciplines. European Commission, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56341

κανόνων σχετικά με την τεχνητή νοημοσύνη, αναφέρεται ο εξής ορισμός για το “σύστημα τεχνητής νοημοσύνης” στο άρθρο 3:

“Είναι λογισμικό που αναπτύσσεται με μία ή περισσότερες από τις τεχνικές και προσεγγίσεις που παρατίθενται στο Παράρτημα Ι και μπορεί, για ένα δεδομένο σύνολο στόχων που έχουν καθοριστεί από τον άνθρωπο, να παράγει στοιχεία εξόδου όπως περιεχόμενο, προβλέψεις, συστάσεις ή αποφάσεις που επηρεάζουν τα περιβάλλοντα με τα οποία αλληλεπιδρά.”

Ειδικότερα το παράρτημα Ι υπό τον τίτλο “Τεχνικές και προσεγγίσεις στον τομέα της Τεχνητής Νοημοσύνης” αναφέρει:

- (α) Προσεγγίσεις μηχανικής μάθησης, συμπεριλαμβανομένων της επιβλεπόμενης, της μη επιβλεπόμενης και της ενισχυτικής μάθησης, με τη χρήση ευρέος φάσματος μεθόδων, συμπεριλαμβανομένης της βαθιάς μάθησης·
- (β) Προσεγγίσεις που βασίζονται στη λογική και στις γνώσεις, συμπεριλαμβανομένων της αναπαράστασης γνώσεων, του επαγωγικού (λογικού) προγραμματισμού, των βάσεων γνώσεων, των μηχανών εξαγωγής συμπερασμάτων και παραγωγικών συλλογισμών, των συστημάτων (συμβολικής) συλλογιστικής και των συστημάτων-εμπειρογνομόνων·
- (γ) Στατιστικές προσεγγίσεις, εκτίμηση κατά Bayes, μέθοδοι αναζήτησης και βελτιστοποίησης.

1.4.4 Αναγκαιότητα κοινά αποδεκτού ορισμού για ασφάλεια δικαίου

Ωστόσο ο προαναφερθείς πιο πρόσφατος ορισμός της Ευρωπαϊκής Επιτροπής διαφέρει από τη διατύπωση που υιοθέτησε το Ευρωπαϊκό Κοινοβούλιο (άρθρο 3), κατά την ψηφοφορία στις 14 Ιουνίου 2023³¹. Ενδιάμεσα, τον Δεκέμβριο 2022, το Συμβούλιο της ΕΕ είχε προτείνει τη δική του παραλλαγή ορισμού³². Ωστόσο αναμενόταν να αποκρυσταλλωθεί αυτός ο ορισμός με την ολοκλήρωση της διαδικασίας των τριλόγων, δεδομένων των διαφορετικών προσεγγίσεων μεταξύ Ευρωπαϊκής Επιτροπής, Συμβουλίου της ΕΕ και Ευρωπαϊκού Κοινοβουλίου. Τελικά το δελτίο τύπου του Συμβουλίου της Ευρωπαϊκής Ένωσης που δημοσιεύτηκε στις 9 Δεκεμβρίου 2023³³, σχετικά με την προσωρινή συμφωνία για τους πρώτους κανόνες παγκοσμίως όσον αφορά την ΤΝ, στην οποία κατέληξαν το Συμβούλιο και το Κοινοβούλιο, αναφέρει σχετικά με τον ορισμό και το πεδίο εφαρμογής το εξής:

“Για να διασφαλιστεί ότι ο ορισμός του συστήματος ΤΝ παρέχει επαρκώς σαφή κριτήρια για τη διάκριση της ΤΝ από απλούστερα συστήματα λογισμικού, η συμβιβαστική συμφωνία ευθυγραμμίζει τον ορισμό με την προσέγγιση που προτείνει ο ΟΟΣΑ.

Η παραπάνω εξέλιξη και η υιοθέτηση του ορισμού του ΟΟΣΑ ως τεχνολογικά ουδέτερου είχε διαφανεί από τον Μάρτιο του 2023 σύμφωνα με δημοσιεύματα³⁴, καθώς ένας κοινά αποδεκτός και ευέλικτος ορισμός είναι αναγκαίος για λόγους ασφάλειας δικαίου. Όπως επισημαίνεται στην τροπολογία 18 επί του σημείου 6 της αιτιολογικής σκέψης της πρότασης Κανονισμού³⁵:

“Η έννοια του συστήματος ΤΝ στον παρόντα κανονισμό θα πρέπει να οριστεί με σαφήνεια και να ευθυγραμμιστεί στενά με το έργο των διεθνών οργανισμών που εργάζονται στον τομέα της τεχνητής νοημοσύνης για λόγους ασφάλειας δικαίου, εναρμόνισης και ευρείας αποδοχής, και για να διασφαλιστεί, παράλληλα, ευελιξία για την προσαρμογή στις ταχείες τεχνολογικές εξελίξεις στον τομέα αυτό.”

³¹https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EL.html

³²Πράξη για την τεχνητή νοημοσύνη: Το Συμβούλιο ζητεί την προώθηση ασφαλών συστημάτων ΤΝ που σέβονται τα θεμελιώδη δικαιώματα, Βρυξέλλες, 6 Δεκεμβρίου 2022, <https://t.ly/S99bc/>

³³Δελτίο τύπου του Συμβουλίου της Ευρωπαϊκής Ένωσης, 9 Δεκεμβρίου 2023, <https://t.ly/D8Cdz>

³⁴EU lawmakers set to settle on OECD definition for Artificial Intelligence, 7 Μαρτίου 2023, <https://www.euractiv.com/section/artificial-intelligence/news/eu-lawmakers-set-to-settle-on-oecd-definition-for-artificial-intelligence>

³⁵https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EL.html

Σε αυτό το σημείο έχει ενδιαφέρον να συνοψίσουμε πού βρίσκονταν τα διαφορετικά όργανα της ΕΕ ως προς τον ορισμό της ΤΝ, ενόσω εξελισσόταν η διαδικασία των τριλόγων³⁶. Όλο το διάστημα υπήρχε μεγάλη αβεβαιότητα ως προς το ποια συστήματα και εφαρμογές θα καλύπτονται από τον νέο νόμο, κάτι πολύ σημαντικό για την ασφάλεια δικαίου όπως προαναφέρθηκε. Ενώ το Ευρωπαϊκό Κοινοβούλιο (ΕΚ) είχε προτείνει μια ευρεία λίστα με απαγορευμένα συστήματα τεχνητής νοημοσύνης, συμπεριλαμβανομένου του λογισμικού που επεξεργάζεται εικόνες προσώπων από το διαδίκτυο, το Συμβούλιο της ΕΕ είχε εκφραστεί υπέρ μιας στενότερης λίστας.

Όπως είδαμε ο προτεινόμενος ορισμός της Ευρωπαϊκής Επιτροπής περιλαμβάνει ορισμένες τεχνικές και εφαρμογές που παρατίθενται στο παράρτημα του κανονισμού. Όμως τόσο το Συμβούλιο της ΕΕ όσο και το Ευρωπαϊκό Κοινοβούλιο μετέφεραν τον ορισμό στο κύριο σώμα του νόμου για την τεχνητή νοημοσύνη. Το Ευρωπαϊκό Κοινοβούλιο ευθυγράμμισε τον ορισμό του “συστήματος τεχνητής νοημοσύνης” (του άρθρου 3) με τον ορισμό της τεχνητής νοημοσύνης που αναπτύχθηκε από τον Οργανισμό Οικονομικής Συνεργασίας και Ανάπτυξης (ΟΟΣΑ) κάτι που τελικά επικράτησε³⁷.

1.4.5 Σύνοψη των ορισμών ΤΝ: ΕΕ, ΗΠΑ και Κίνα

Τα κράτη μέλη της ΕΕ που εκπροσωπούνται στο Συμβούλιο της ΕΕ ανησυχούσαν περισσότερο ότι η Πράξη για την ΤΝ είναι πολύ ευρεία με αποτέλεσμα να καλύπτει παραδοσιακές εφαρμογές λογισμικού. Το Συμβούλιο επιχείρησε να περιορίσει τον ορισμό σε συστήματα που χρησιμοποιούν μηχανική μάθηση, λογική ή προσεγγίσεις που βασίζονται στη γνώση. Παρακάτω παραθέτουμε τις διατυπώσεις που προτάθηκαν για τον ορισμό του άρθρου 3.

Ευρωπαϊκή Επιτροπή: “Σύστημα τεχνητής νοημοσύνης σημαίνει λογισμικό που αναπτύσσεται με μία ή περισσότερες από τις τεχνικές και προσεγγίσεις που αναφέρονται στο Παράρτημα Ι³⁸ και μπορεί, για ένα δεδομένο σύνολο στόχων που καθορίζονται από τον άνθρωπο, να παράγει αποτελέσματα όπως περιεχόμενο, προβλέψεις, συστάσεις ή αποφάσεις που επηρεάζουν τα περιβάλλοντα με τα οποία αλληλεπιδρούν.”

Ευρωπαϊκό Κοινοβούλιο (ΕΚ): “Σύστημα τεχνητής νοημοσύνης σημαίνει ένα σύστημα που βασίζεται σε μηχανές που έχει σχεδιαστεί για να λειτουργεί με διαφορετικά επίπεδα αυτονομίας και που μπορεί, για ρητούς ή άρρητους στόχους, να παράγει αποτελέσματα όπως προβλέψεις, συστάσεις ή αποφάσεις, που επηρεάζουν φυσικά ή εικονικά περιβάλλοντα.”

Συμβούλιο ΕΕ: Όπως σημειώθηκε, το Συμβούλιο προσπαθεί να περιορίσει τον ορισμό σε συστήματα που αναπτύσσονται μέσω προσεγγίσεων μηχανικής μάθησης και προσεγγίσεων που βασίζονται στη λογική και τη γνώση³⁹.

Συμπεραίνουμε ότι οι προσεγγίσεις του ΕΚ και της Επιτροπής της ΕΕ παρουσίασαν σημαντική απόκλιση. Στην πρότασή της, η Επιτροπή τονίζει ότι το σύστημα μπορεί “να παράγει αποτελέσματα για ένα δεδομένο σύνολο στόχων που καθορίζονται από τον άνθρωπο”. Αυτή η αναφορά σε στόχους που καθορίζονται από τον άνθρωπο δεν υπάρχει στην πρόταση του ΕΚ.

Η πρόταση της Επιτροπής αναφέρεται στο Παράρτημα Ι, όπου τα “συστήματα τεχνητής νοημοσύνης” ορίζονται λεπτομερέστερα. Αυτή η αναφορά καθιστά δύσκολη την πρόβλεψη του τι θα

³⁶ΝΟΜΟΣ ΤΗΣ ΕΕ ΓΙΑ ΤΗΝ ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ: ΣΥΖΗΤΗΣΗ ΓΙΑ ΤΑ ΘΕΣΜΙΚΑ ΟΡΓΑΝΑ ΟΡΙΣΜΟΣ ΤΗΣ ΤΕΧΝΗΤΗΣ ΝΟΗΜΟΣΥΝΗΣ, LEGAL INSIGHTS GERMANIA, 26 Σεπτεμβρίου 2023, <https://www.morganlewis.com/pubs/2023/09/eu-ai-act-institutions-debate-definition-of-ai>

³⁷ Δελτίο τύπου του Συμβουλίου της Ευρωπαϊκής Ένωσης, 9 Δεκεμβρίου 2023, <https://t.ly/D8Cdz>

³⁸ Παράρτημα Ι, ΤΕΧΝΙΚΕΣ ΚΑΙ ΠΡΟΣΕΓΓΙΣΕΙΣ ΣΤΟΝ ΤΟΜΕΑ ΤΗΣ ΤΕΧΝΗΤΗΣ ΝΟΗΜΟΣΥΝΗΣ που αναφέρονται στο άρθρο 3 σημείο 1: α) Προσεγγίσεις μηχανικής μάθησης, συμπεριλαμβανομένων της επιβλεπόμενης, της μη επιβλεπόμενης και της ενισχυτικής μάθησης, με τη χρήση ευρέος φάσματος μεθόδων, συμπεριλαμβανομένης της βαθιάς μάθησης· β) Προσεγγίσεις που βασίζονται στη λογική και στις γνώσεις, συμπεριλαμβανομένων της αναπαραστάσης γνώσεων, του επαγωγικού (λογικού) προγραμματισμού, των βάσεων γνώσεων, των μηχανών εξαγωγής συμπερασμάτων και παραγωγικών συλλογισμών, των συστημάτων (συμβολικής) συλλογιστικής και των συστημάτων εμπειρογνομόνων· γ) Στατιστικές προσεγγίσεις, εκτίμηση κατά Bayes, μέθοδοι αναζήτησης και βελτιστοποίησης.

³⁹<https://data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/el/pdf>

περιλαμβάνει ο ορισμός. Η έκδοση του Ευρωπαϊκού Κοινοβουλίου περιέχει επίσης κάποια ασάφεια, για παράδειγμα με τον όρο “διαφορετικά επίπεδα αυτονομίας”. Το Ευρωπαϊκό Κοινοβούλιο επίσης είχε εισαγάγει άλλους νέους ορισμούς στην πρότασή του τον Ιούνιο του 2023, συμπεριλαμβανομένου του “συστήματος τεχνητής νοημοσύνης γενικού σκοπού”.

Συγκριτικά, η προσέγγιση των **ΗΠΑ** για τη ρύθμιση της τεχνητής νοημοσύνης ήταν μέχρι πρόσφατα ακόμη πιο ασαφής καθώς ο ορισμός⁴⁰ στο προσχέδιο για τη Διακήρυξη Δικαιωμάτων της ΤΝ ανέφερε: “Αυτό το πλαίσιο ισχύει για (1) αυτοματοποιημένα συστήματα που (2) έχουν τη δυνατότητα να επηρεάσουν ουσιαστικά τα δικαιώματα, τις ευκαιρίες ή την πρόσβαση του αμερικανικού κοινού σε κρίσιμους πόρους ή υπηρεσίες”. Όμως με το Εκτελεστικό Διάταγμα⁴¹ για την ασφαλή, ακίνδυνη και αξιόπιστη ανάπτυξη και χρήση της τεχνητής νοημοσύνης, που εξέδωσε στις 30 Οκτωβρίου 2023 ο Λευκός Οίκος, η ΤΝ περιγράφεται ως ένα σύστημα βασισμένο σε μηχανές που μπορεί, για ένα δεδομένο σύνολο στόχων που καθορίζονται από τον άνθρωπο, να κάνει προβλέψεις, συστάσεις ή αποφάσεις επηρεάζοντας πραγματικά ή εικονικά περιβάλλοντα. Διευκρινίζεται επίσης ότι τα συστήματα τεχνητής νοημοσύνης χρησιμοποιούν εισροές που βασίζονται σε μηχανές και ανθρώπους για να αντιληφθούν πραγματικά και εικονικά περιβάλλοντα, να αναγάγουν αυτή την πρόσληψη σε μοντέλα μέσω ανάλυσης με αυτοματοποιημένο τρόπο και να χρησιμοποιήσουν την εξαγωγή συμπερασμάτων από τα μοντέλα για τη διατύπωση επιλογών πληροφόρησης ή δράσης.

Από την άλλη πλευρά, οι ισχύοντες νόμοι και κανονισμοί της **Κίνας** δεν παρέχουν σαφή ορισμό της ΤΝ⁴². Σύμφωνα με το άρθρο 22 του νόμου “Προσωρινά μέτρα για τη διαχείριση της παραγωγικής ΤΝ”⁴³, η τελευταία αναφέρεται σε μοντέλα και συναφή τεχνολογία που έχουν τη δυνατότητα να παράγουν κείμενο, εικόνες, ήχους, βίντεο ή άλλο περιεχόμενο.

1.5 Ορισμός της ΤΝ από τον ΟΟΣΑ

Καθώς ο ορισμός που τελικά υιοθετήθηκε από την ΕΕ για την ΤΝ ταυτίζεται με τον ορισμό του ΟΟΣΑ, θα επικεντρωθούμε σε αυτόν και θα τον δούμε αναλυτικότερα. Η Ομάδα Εμπειρογνομώνων ΤΝ του ΟΟΣΑ (OECD’s AI Experts Group-AIGO) εξηγεί⁴⁴ ότι ένα σύστημα ΤΝ είναι ένα σύστημα βασισμένο σε μηχανή που μπορεί, για ένα δεδομένο σύνολο στόχων καθορισμένων από τον άνθρωπο, να προβαίνει σε προβλέψεις, συστάσεις ή αποφάσεις και να επηρεάζει πραγματικά ή εικονικά περιβάλλοντα. Χρησιμοποιεί εισροές που βασίζονται σε μηχανή ή/και άνθρωπο για να αντιληφθεί πραγματικά ή/και εικονικά περιβάλλοντα, να αναγάγει αφαιρετικά τις εν λόγω προσλήψεις σε μοντέλα (με αυτοματοποιημένο τρόπο, π.χ. με μηχανική μάθηση ή χειροκίνητα) και να χρησιμοποιήσει την εξαγωγή συμπερασμάτων από τα μοντέλα για να διατυπώσει επιλογές για πληροφορίες ή ενέργειες.

Τα συστήματα ΤΝ σχεδιάζονται για να λειτουργούν με διαφορετικά επίπεδα αυτονομίας. Οι φάσεις του κύκλου ζωής των συστημάτων ΤΝ είναι: α) σχεδιασμός και προγραμματισμός, συλλογή και επεξεργασία δεδομένων, κατασκευή και ερμηνεία μοντέλων, β) επαλήθευση και επικύρωση, γ) ανάπτυξη και δ) λειτουργία και παρακολούθηση. Από την ίδια μελέτη του ΟΟΣΑ αντλούμε και την ακόλουθη απεικόνιση όπου παρουσιάζεται μια εννοιολογική θεώρηση της ΤΝ ως δομή υψηλού επιπέδου ενός γενικού συστήματος ΤΝ. Ένα σύστημα ΤΝ αποτελείται από τρία κύρια στοιχεία:

⁴⁰<https://www.whitehouse.gov/ostp/ai-bill-of-rights>

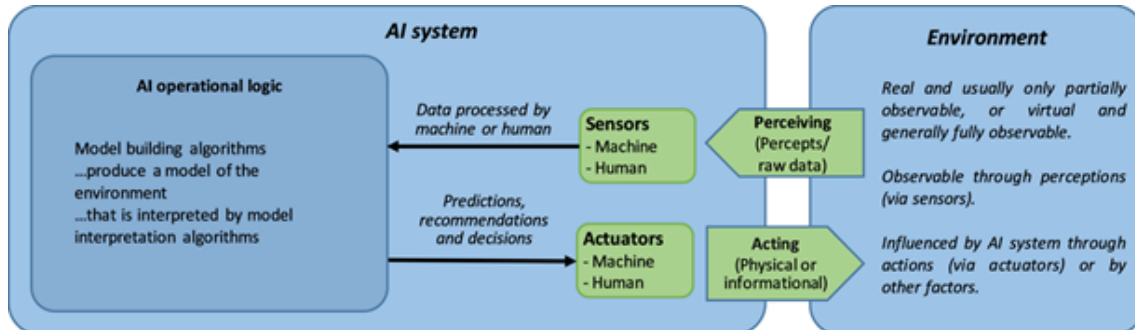
⁴¹Εκτελεστικό Διάταγμα για την ασφαλή, ακίνδυνη και αξιόπιστη ανάπτυξη και χρήση της τεχνητής νοημοσύνης, Λευκός Οίκος, 30 Οκτωβρίου 2023, https://t.ly/p_nHa

⁴²AI Watch: Global regulatory tracker - China, 13 Μαΐου 2024, <https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-china>

⁴³Interim Measures for the Management of Generative Artificial Intelligence Services, <https://www.chinalawtranslate.com/en/generative-ai-interim/>

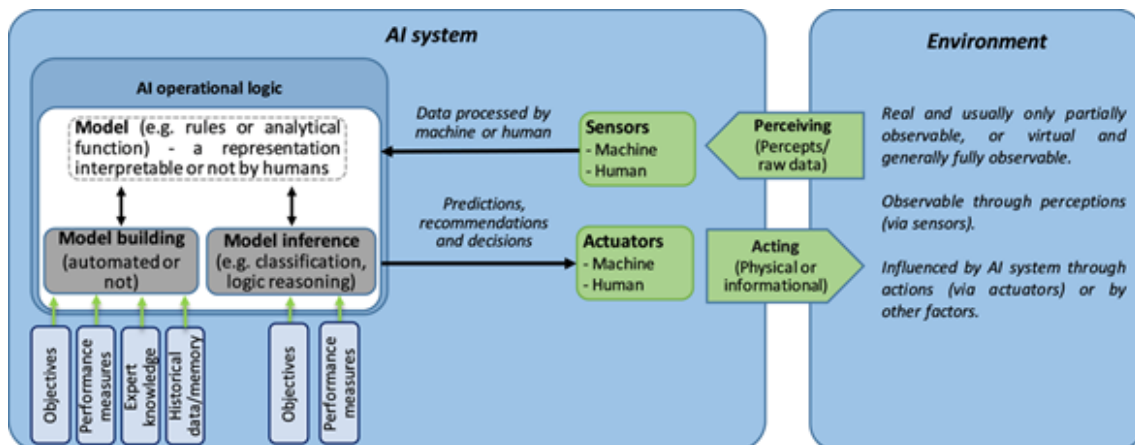
⁴⁴OECD, Artificial Intelligence in Society, 2019, <https://www.oecd.org/publications/artificial-intelligence-in-society-eedfee77-en.htm>

αισθητήρες, λειτουργική λογική και ενεργοποιητές. Οι αισθητήρες συλλέγουν ακατέργαστα δεδομένα από το περιβάλλον, ενώ οι ενεργοποιητές ενεργούν για να αλλάξουν την κατάσταση του περιβάλλοντος. Η βασική ισχύς ενός συστήματος ΤΝ έγκειται στη λειτουργική λογική του. Για ένα δεδομένο σύνολο στόχων και με βάση τα δεδομένα εισόδου από τους αισθητήρες, η λειτουργική λογική παρέχει έξοδο για τους ενεργοποιητές. Αυτοί λαμβάνουν τη μορφή συστάσεων, προβλέψεων ή αποφάσεων που μπορούν να επηρεάσουν την κατάσταση του περιβάλλοντος.



Εικόνα 1.4. A high-level conceptual view of an AI system, OECD, 2019

Περαιτέρω, μια πιο λεπτομερής δομή αποτυπώνει τα κύρια στοιχεία που σχετίζονται με τις μεθόδους των συστημάτων ΤΝ και που παρουσιάζεται παρακάτω. Για να καλυφθούν διαφορετικοί τύποι συστημάτων ΤΝ και διαφορετικά σενάρια, το διάγραμμα διαχωρίζει τη διαδικασία κατασκευής μοντέλων (όπως η ML), από το ίδιο το μοντέλο. Η κατασκευή του μοντέλου διαχωρίζεται επίσης από τη διαδικασία ερμηνείας του μοντέλου, η οποία χρησιμοποιεί το μοντέλο για να παράγει προβλέψεις, συστάσεις και αποφάσεις- οι ενεργοποιητές χρησιμοποιούν αυτές τις εξόδους για να επηρεάσουν το περιβάλλον. Όμως και ο παραπάνω ορισμός επικαιροποιήθηκε πρόσφατα (το Νοέμβριο 2023) με σκοπό να συμπεριληφθεί στην ευρωπαϊκή Πράξη για την ΤΝ, λαμβάνοντας υπόψιν τις πρόσφατες εξελίξεις στην παραγωγική ΤΝ και προκειμένου να ευθυγραμμιστεί καλύτερα με τη μελλοντική πρόοδο.



Εικόνα 1.5. Detailed conceptual view of an AI System, OECD, 2019

Η παρούσα διατύπωση έχει ως εξής⁴⁵:

“Ένα σύστημα τεχνητής νοημοσύνης είναι ένα σύστημα που βασίζεται σε μηχανήματα που, για ρητούς ή άρρητους στόχους, συνάγει, από τα δεδομένα (εισόδους) που λαμβάνει, πώς να πα-

⁴⁵OECD updates definition of Artificial Intelligence ‘to inform EU’s AI Act’, 14 Νοεμβρίου 2023, <https://www.euractiv.com/section/artificial-intelligence/news/oecd-updates-definition-of-artificial-intelligence-to-inform-eus-ai-act/>

ράγει αποτελέσματα όπως προβλέψεις, περιεχόμενο, προτάσεις ή αποφάσεις που [μπορούν] να επηρεάσουν φυσικά ή εικονικά περιβάλλοντα. Τα διαφορετικά συστήματα τεχνητής νοημοσύνης διαφέρουν ως προς τα επίπεδα αυτονομίας και προσαρμοστικότητας μετά την ανάπτυξη”, αναφέρει ο νέος ορισμός.

Στην ιστοσελίδα του ΟΟΣΑ⁴⁶ πληροφορούμαστε αναλυτικά τις αναθεωρήσεις του τρέχοντος ορισμού για ένα “Σύστημα ΤΝ”, όπου επισημαίνονται οι προσθήκες με **έντονους χαρακτήρες** και οι διαγραφές με διαγράμμιση, όπως αποτυπώνεται στην ακόλουθη εικόνα:

How OECD countries updated the definition

Here are the revisions to the current text of the definition of “AI System” in detail, with additions set out in **bold** and subtractions in ~~strikethrough~~:

*An AI system is a machine-based system that ~~can~~, for a given set of human-defined **explicit or implicit** objectives, **infers, from the input it receives, how to generate outputs such as** ~~makes~~ predictions, **content**, recommendations, or decisions **that can influence** ~~ing~~ **physical** ~~real~~ or virtual environments. **Different** AI systems ~~are designed to operate with varying~~ **in their levels of autonomy and adaptiveness after deployment***

Εικόνα 1.6. Updates to the OECD’s definition of an AI system, OECD, Nov. 2023

Αυτές οι αλλαγές αντικατοπτρίζουν τις ακόλουθες παρατηρήσεις:

Περιγραφή των στόχων: Αυτές οι τροποποιήσεις επιδιώκουν να αντικατοπτρίζουν την επισημοποιημένη συναίνεση ότι οι στόχοι ενός συστήματος τεχνητής νοημοσύνης μπορεί να είναι **ρητοί** (π.χ. όπου προγραμματίζονται απευθείας στο σύστημα από άνθρωπο προγραμματιστή) ή **άρρητοι** (π.χ. μέσω ενός συνόλου κανόνων που καθορίζονται από έναν άνθρωπο ή όταν το σύστημα είναι ικανό να μάθει νέους στόχους).

Παραδείγματα συστημάτων με άρρητους στόχους περιλαμβάνουν συστήματα αυτοοδήγησης που είναι προγραμματισμένα να συμμορφώνονται με τους κανόνες οδικής κυκλοφορίας (αλλά δεν “γνωρίζουν” τον σιωπηρό τους στόχο να προστατεύουν ζωές) ή ένα μεγάλο γλωσσικό μοντέλο όπως το ChatGPT και άλλα παρόμοια συστήματα, όπου ο στόχος του συστήματος δεν έχει προγραμματιστεί ρητά αλλά αποκτάται εν μέρει μέσω της διαδικασίας μίμησης μάθησης από κείμενο που δημιουργείται από τον άνθρωπο και εν μέρει από τη διαδικασία ενίσχυσης της μάθησης από την ανθρώπινη ανάδραση (reinforcement learning from human feedback, RLHF).

Είσοδοι: Η προσθήκη του “συνάγει, από τις εισόδους που λαμβάνει” υπογραμμίζει τον σημαντικό ρόλο της εισαγωγής, συμπεριλαμβανομένων κανόνων και δεδομένων, που μπορεί να παρέχονται από ανθρώπους ή μηχανές, σε συστήματα τεχνητής νοημοσύνης που λειτουργούν. Ένα σύστημα τεχνητής νοημοσύνης λέγεται ότι “συμπεραίνει πώς να δημιουργεί εξόδους” όταν λαμβάνει δεδομένα από το περιβάλλον και υπολογίζει μια έξοδο επεξεργαζόμενο την είσοδο μέσω ενός ή περισσότερων μοντέλων και υποκειμένων αλγορίθμων. Για παράδειγμα, ένα οπτικό σύστημα αναγνώρισης αντικειμένων που υλοποιείται από ένα βαθύ νευρωνικό δίκτυο “συνάγει”, δηλαδή συμπεραίνει πώς δημιουργείται η έξοδος του (σε αυτή την περίπτωση, μια ταξινόμηση του αντικειμένου στην εικόνα) περνώντας την είσοδο του (τα pixel της εικόνας) μέσα από το βαθύ νευρωνικό δίκτυο (μια παραμετροποιημένη αλγεβρική έκφραση που αποτελείται από πρόσθεση, πολλαπλασιασμό και ορισμένες μη γραμμικές πράξεις).

⁴⁶OECD updates definition of Artificial Intelligence ‘to inform EU’s AI Act’, 29 Νοεμβρίου 2023, <https://oecd.ai/en/wonk/ai-system-definition-update>

Έξοδοι: Η προσθήκη της λέξης “περιεχόμενο” διευκρινίζει ότι η Σύσταση ισχύει για συστήματα παραγωγικής τεχνητής νοημοσύνης, τα οποία παράγουν “περιεχόμενο” (τεχνικά, ένα υποσύγολο “προβλέψεων, συστάσεων ή αποφάσεων”) όπως κείμενο, βίντεο ή εικόνες.

Περιβάλλον: Η αντικατάσταση του “πραγματικού” με το “φυσικό” αποσαφηνίζει και ευθυγραμμίζεται με άλλες διεθνείς διαδικασίες. Επιπλέον, η αντίθεση του πραγματικού με το εικονικό υποδηλώνει ότι τα εικονικά περιβάλλοντα δεν είναι πραγματικά, κάτι που δεν συμβαίνει: είναι αληθινά καθώς δέχονται πραγματικές ενέργειες από το σύστημα TN και δημιουργούν πραγματικές αισθητηριακές εισόδους στο σύστημα TN.

Προσαρμοστικότητα: Αυτό αντικατοπτρίζει ότι ορισμένα συστήματα τεχνητής νοημοσύνης μπορούν να συνεχίσουν να εξελίσσονται μετά το σχεδιασμό και την ανάπτυξή τους (για παράδειγμα, συστήματα συστάσεων που προσαρμόζονται στις ατομικές προτιμήσεις ή συστήματα αναγνώρισης φωνής που προσαρμόζονται στη φωνή του χρήστη) και είναι ένας πρόσθετος χαρακτηρισμός μιας σημαντικής ομάδας συστημάτων τεχνητής νοημοσύνης.

Ας σημειώσουμε σε αυτό το σημείο ότι την προσέγγιση του ΟΟΣΑ υιοθέτησε και το Συμβούλιο της Ευρώπης στο κείμενο της πρώτης διεθνούς Σύμβασης Πλαίσιο για την τεχνητή νοημοσύνη, τα ανθρώπινα δικαιώματα, τη δημοκρατία και το κράτος δικαίου⁴⁷ και αποδίδει αυτή την επιλογή “στην υψηλή ποιότητα των εργασιών του ΟΟΣΑ και των εμπειρογνομόνων του, αλλά και ενόψει της ανάγκης να ενισχυθεί η διεθνής συνεργασία στο θέμα της τεχνητής νοημοσύνης και να διευκολυνθούν οι προσπάθειες που αποσκοπούν στην εναρμόνιση της διακυβέρνησης της τεχνητής νοημοσύνης σε παγκόσμιο επίπεδο, μεταξύ άλλων με την εναρμόνιση της σχετικής ορολογίας”.

1.6 Συμπέρασμα

Η Τεχνητή Νοημοσύνη ως σύνολο τεχνολογιών και εφαρμογών έχει προκαλέσει ραγδαίες εξελίξεις σε ποικίλες πτυχές της δημόσιας και ιδιωτικής ζωής. Η μέχρι τώρα εμπειρία αλλά και οι προβλέψεις για το μέλλον φανερώνουν τη δυναμική της να επιφέρει ριζικές μεταβολές στον τρόπο με τον οποίο δραστηριοποιούμαστε σε κάθε τομέα, μεταβολές οι οποίες έχουν διττό χαρακτήρα, με θετικές και αρνητικές διαστάσεις. Τους προβληματισμούς επαληθεύει η αναδρομή στο ιστορικό της TN ως επιστήμης, όταν τέθηκαν οι επιστημονικές βάσεις για τη μετέπειτα εξέλιξή της, που ευνοήθηκε και από άλλες συνθήκες όπως η αυξημένη υπολογιστική ισχύς και η αύξηση των δεδομένων.

Γίνεται πλέον αντιληπτό ότι επίμαχο θέμα είναι το μοντέλο διακυβέρνησης της TN προκειμένου να εξισορροπηθούν οι προαναφερθείσες αντικρουόμενες τάσεις, κάτι που απαιτεί ένα κοινό αποδεκτό ορισμό προκειμένου να εδραιωθούν η ασφάλεια δικαίου, η συναντίληψη και ένα κοινό λεξιλόγιο για το σκοπό της ρύθμισής της. Μετά από μακρά πορεία εννοιολογικών αναζητήσεων από επιστημονικούς και πολιτικούς φορείς, έχει επικρατήσει ο ορισμός του ΟΟΣΑ.

⁴⁷Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law, Άρθρο 2, Ορισμός, <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>

Κεφάλαιο 2

Εφαρμογές της ΤΝ στον δημόσιο τομέα και προκλήσεις

- Στο κεφάλαιο αυτό θα αποτυπωθούν οι ποικίλες δυνατότητες της τεχνολογίας ΤΝ και οι προεκτάσεις της προς τομείς που την ανέδειξαν σε ένα πολύτιμο εργαλείο πολιτικής διακυβέρνησης και δημόσιας διοίκησης.
- Ως εκ τούτου θα εκτεθούν κάποια αντιπροσωπευτικά παραδείγματα χρήσεων των συστημάτων τεχνητής νοημοσύνης κατά την άσκηση του έργου του δημόσιου τομέα αλλά και προκλήσεις που οφείλονται στους κινδύνους που ελλοχεύουν σε αυτό το πεδίο.
- Επίσης θα εξεταστούν δειγματοληπτικά κάποιες περιπτώσεις προϋποθέσεων που θέτουν δημόσιοι φορείς ως χρήστες της ΤΝ και ορισμένα ενδεικτικά εργαλεία που έχουν χρησιμοποιηθεί για μετριάσμο των κινδύνων, τα οποία συνδέονται με τη διακυβέρνηση δεδομένων και αλγορίθμων, ως συστατικών στοιχείων της ΤΝ.

2.1 Πρακτική εφαρμογή της ΤΝ στον δημόσιο τομέα

Ενώ το ρυθμιστικό τοπίο που περιβάλλει τη διακυβέρνηση της ΤΝ έχει μελετηθεί αρκετά, ως επί το πλείστον έχει παραβλεφθεί ο ρόλος μιας κυβέρνησης ως χρήστη της ΤΝ. Το άρθρο με τίτλο “An Artificial Intelligence Definition and Classification Framework for Public Sector Applications” παρουσιάζει ένα πλαίσιο ταξινόμησης για τις εφαρμογές ΤΝ που διευκολύνει την προσέγγιση του δημόσιου τομέα στο ρόλο του χρήστη¹. Οι κυβερνήσεις έχουν διττό ρόλο στην υιοθέτηση της ΤΝ, αφενός ως ρυθμιστές της ΤΝ για την προστασία των πολιτών από δυνητικές βλάβες και αφετέρου ως χρήστες της ΤΝ για την καλύτερη παροχή υπηρεσιών και την αποτελεσματική κατανομή των πόρων.

Η ανάθεση της δημόσιας εξουσίας σε συστήματα ΤΝ για την εκτέλεση δημόσιων λειτουργιών παρουσιάζεται συχνά ως επωφελής για τον τελικό χρήστη, επιταχύνοντας τις γραφειοκρατικές διαδικασίες και αυξάνοντας παράλληλα την εμπιστοσύνη και την ικανοποίηση των πολιτών. Μαζί με την ψηφιοποίηση, η χρήση της τεχνητής νοημοσύνης στον δημόσιο τομέα δύναται να αυξήσει την αποδοτικότητα και την αποτελεσματικότητα, απελευθερώνοντας πόρους από διοικητικές εργασίες. Από την άλλη πλευρά, δημιουργούνται ζητήματα προστασίας της ιδιωτικής ζωής και της ασφάλειας των δεδομένων, αντιπροσωπευτικότητας, δικαιοσύνης και λογοδοσίας της ΤΝ. Στο πλαίσιο του δημόσιου τομέα σε ρόλο χρήστη των εφαρμογών ΤΝ, το άρθρο αναδεικνύει ερωτήματα όπως:

- Πότε είναι κρίσιμο να εξετάσουμε τη διαδικασία και το αποτέλεσμα;

¹Omar Ballester, 2021, An Artificial Intelligence Definition and Classification Framework for Public Sector Applications, In DG.O2021: The 22nd Annual International Conference on Digital Government Research (DG.O'21). Association for Computing Machinery, New York, NY, USA, 67–75.<https://doi.org/10.1145/3463677.3463709>

- Υπό ποιες συνθήκες θα πρέπει μας απασχολήσει η λογοδοσία των αλγορίθμων;
- Σε ποιο θεσμικό πλαίσιο μπορούμε να αναπτύξουμε ένα συγκεκριμένο σύστημα;
- Σε ποιο βαθμό ένα σύστημα TN μπορεί να έχει δυσμενείς επιπτώσεις στους πολίτες;

Το άρθρο αναπτύσσει ένα πλαίσιο ταξινόμησης για να βοηθήσει τους διαχειριστές του δημόσιου τομέα να κατανοήσουν και να επιλέξουν εάν και πώς θα εφαρμόσουν τεχνολογίες TN. Το πλαίσιο επικεντρώνεται σε δύο βασικές συνιστώσες: τα καθήκοντα (εντολές/ενέργειες που εκτελούνται από την TN) και τις εισροές (απαιτούμενα δεδομένα). Η έρευνα αποσκοπεί στον προσδιορισμό απτών κριτηρίων για την αξιολόγηση έργων TN στον δημόσιο τομέα. Αναλύοντας 25 πιλοτικά έργα TN σε τομείς όπως η γεωργία, η κοινωνική ασφάλιση και οι μεταφορές, το έγγραφο προσφέρει πληροφορίες για τους κρίσιμους παράγοντες ταξινόμησης των συστημάτων TN εντός του δημοσίου. Προτείνοντας ένα δομημένο σύστημα κατηγοριοποίησης, συμβάλλει σε μια στρατηγική αξιοποίηση δεδομένων, υποστηρίζοντας μια στρατηγική ανάπτυξη της TN.

2.1.1 “Νομικό μαύρο κουτί” έναντι “τεχνικού μαύρου κουτιού”

Το άρθρο επισημαίνει ότι δεν είναι ισότιμες όλες οι διαδικασίες που διευκολύνονται από την τεχνητή νοημοσύνη στον δημόσιο τομέα και επισημαίνει το “τεχνικό μαύρο κουτί” αλλά και το “νομικό μαύρο κουτί” για τις λειτουργίες που δείχνουν επιφανειακά συναφείς ωστόσο έχουν τελείως διαφορετική φύση, για παράδειγμα, τα chatbots που μπορεί να βοηθούν στην εύρεση ηλεκτρονικών εγγράφων αλλά και στην πρόγνωση απάτης. Επιπλέον διευκρινίζει ότι η ίδια τεχνική λύση για δύο διαφορετικά πλαίσια εφαρμογών μπορεί να έχει εντελώς διαφορετικές κανονιστικές επιπτώσεις (π.χ. συστήματα διαχείρισης της κυκλοφορίας με τη βοήθεια της ηλεκτρονικής όρασης από την μία πλευρά και συστήματα αναγνώρισης προσώπου με ηλεκτρονική όραση από την άλλη). Η στάθμιση οφέλους-κινδύνου εξαρτάται από το πλαίσιο εφαρμογής, ακριβώς όπως τα ρομπότ αποθήκης δεν φέρουν το ίδιο βάρος αποφάσεων με τα συστήματα ρομποτικής χειρουργικής.

Συνεπώς το εν λόγω άρθρο εστιάζει την προσοχή του στις συγκεκριμένες εργασίες (εντολές, tasks) που διεκπεραιώνει ένα σύστημα TN και στις εισροές (δεδομένα, data) που απαιτεί, αντί να δίνει έμφαση στην τεχνολογία και τις λειτουργίες των υπό εξέταση συστημάτων TN. Έτσι καταλήγει σε ένα πλαίσιο ταξινόμησης που χρησιμεύει ως εργαλείο του δημόσιου τομέα για την κατανόηση των εφαρμογών TN. Η ανάλυση της βιβλιογραφίας δείχνει ότι υπάρχουν ελλείψεις στη μελέτη των πλαισίων εφαρμογής. Η πλειονότητα των δημοσιεύσεων επικεντρώνεται στις μεγάλες κοινωνικές προκλήσεις της TN: ηθική/δεοντολογία, διακυβέρνηση και νομικές πτυχές, ενώ μια δεύτερη ομάδα ασχολείται με τις τεχνικές πτυχές της υιοθέτησης, την αυτοματοποίηση των εργασιών και των χώρων εργασίας και τα ζητήματα δεδομένων.

2.1.2 Εσωτερική υλοποίηση έναντι εξωτερικής υλοποίησης

Επιπλέον επισημαίνει ότι ενώ ευρύτερα επικρατεί η πεποίθηση ότι τα δεδομένα, η ασφάλεια και οι εφαρμογές TN πρέπει να αντιμετωπίζονται σχεδόν αποκλειστικά εσωτερικά, το άρθρο υποστηρίζει ότι υπό ορισμένες συνθήκες η ανάπτυξη εφαρμογών TN είναι προτιμότερο να ανατίθενται πλήρως σε εξωτερικούς φορείς. Η τεκμηρίωση αυτής της άποψης βασίστηκε σε 25 πιλοτικά έργα με επαρκείς δημόσιες πληροφορίες για την πηγή δεδομένων, τον τύπο της τεχνολογικής λύσης (TN) και της βασικής λειτουργίας της, όπως εμφανίζεται στους Πίνακες 2.1, 2.2 και 2.3. Τα έργα αυτά καλύπτουν πολλούς τομείς των δημόσιων υπηρεσιών, όπως η γεωργία, η τοπική αυτοδιοίκηση, η κοινωνική ασφάλιση ή οι μεταφορές και υλοποιούνται στο πλαίσιο της χρήσης της TN από τον δημόσιο τομέα.

Από τη διερευνητική ανάλυση της βιβλιογραφίας και των πιλοτικών έργων που περιγράφονται στο παραπάνω πίνακα, προτείνονται δύο σημαντικές διαστάσεις του πλαισίου ταξινόμησης: τα καθήκοντα και οι εργασίες του συστήματος TN αφενός, και οι ανάγκες σε δεδομένα για επιτυχή εφαρμογή αφετέρου. Αυτοί οι δύο πυλώνες αποτελούν τον ακρογωνιαίο λίθο της ταξινόμησης, αλλά δεν επαρκούν από μόνοι τους. Πρέπει επιπλέον να λαμβάνονται υπόψιν τα οργανωσιακά

χαρακτηριστικά και οι οργανωτικές δυνατότητες του οργανισμού, καθώς έχουν άμεσο αντίκτυπο στον εντοπισμό σημείων για αυτοματοποίηση, τόσο στις εργασίες που μπορούν να εκτελεστούν όσο και στα ρυθμιστικά, ηθικά και ποιοτικά ζητήματα των δεδομένων.

2.1.3 Οργανωσιακές ικανότητες

Στις οργανωσιακές ικανότητες περιλαμβάνονται τόσο οι προσωπικές όσο και οι τεχνικές ικανότητες. Η τεχνητή νοημοσύνη στον δημόσιο τομέα απαιτεί μια στρατηγική δράσης για τη δημιουργία αξίας. Κατά συνέπεια, η υιοθέτηση λύσεων TN που δεν προσφέρονται έτοιμες από εξωτερικούς προμηθευτές απαιτεί ένα επίπεδο εσωτερικής τεχνογνωσίας για την υποστήριξη της ανάπτυξης. Η ύπαρξη ενός εκπαιδευμένου και ποικιλόμορφου εργατικού δυναμικού, με εξειδίκευση στην TN, διευκολύνει τον εντοπισμό των καθηκόντων που μπορούν να ενισχυθούν από συστήματα TN. Επιπλέον η ύπαρξη ειδικών εντός της ομάδας μειώνει τους κινδύνους που σχετίζονται με την ασφάλεια της TN. Τα τεχνικά χαρακτηριστικά της υπηρεσίας που σκοπεύει να αναπτύξει συστήματα TN είναι εξίσου σημαντικά με το ανθρώπινο κεφάλαιο.

2.1.4 Δεδομένα: εσωτερικά έναντι εξωτερικών

Τα σύγχρονα συστήματα TN που διψούν για δεδομένα είναι τόσο έξυπνα όσο η ποιότητα των δεδομένων που λαμβάνουν. Οι οργανισμοί με χαμηλής ποιότητας και αδόμητα δεδομένα δεν είναι κατάλληλοι για συστήματα TN που αξιοποιούν εσωτερικούς πόρους δεδομένων για εκπαίδευση. Τα δεδομένα που χρειάζονται τα συστήματα τεχνητής νοημοσύνης για την εκπαίδευση είναι θεμελιώδους σημασίας για την εφαρμογή της TN. Οι δημόσιες υπηρεσίες που προσπαθούν να αναπτύξουν λύσεις TN μπορούν να βρεθούν σε δύο σενάρια: τα δεδομένα είναι εσωτερικά του οργανισμού ή τα δεδομένα είναι εξωτερικά.

2.1.4.1 Εσωτερικά δεδομένα

Αναφέρεται σε δεδομένα που υπάρχουν στη δομή του οργανισμού, πριν από την ανάπτυξη του συστήματος TN, αλλά μπορούν να εμπλουτιστούν κατά τη διάρκεια της ανάπτυξης με εξωτερικά δεδομένα. Τα εσωτερικά δεδομένα πρέπει να είναι διαθέσιμα, προσβάσιμα και αναλύσιμα προκειμένου να αξιοποιηθούν συνδυαστικά με ευφυείς αλγόριθμους. Στη δημόσια διοίκηση τα δεδομένα είναι συχνά ιδιαίτερα ευαίσθητα σε θέματα προστασίας της ιδιωτικής ζωής, ιδίως όταν πρόκειται για δεδομένα προσωπικού χαρακτήρα. Σε αυτή την περίπτωση, τα δεδομένα αποτελούν μια σημαντική νομική και τεχνολογική πρόκληση, ιδίως εάν στην ανάπτυξη του συστήματος TN πρέπει να συμμετέχουν εξωτερικοί προμηθευτές ή προγραμματιστές. Τέλος, τα εσωτερικά δεδομένα που χρησιμοποιούνται για την εκπαίδευση ενός συστήματος τεχνητής νοημοσύνης απαιτούν προσεκτικό χειρισμό προκειμένου να αποφευχθούν οι διακρίσεις/προκαταλήψεις που προκαλούνται από τα συστήματα τεχνητής νοημοσύνης. Τα δεδομένα συχνά εμπεριέχουν εγγενείς προκαταλήψεις, κάτι που ενισχύεται από τη στατιστική φύση της TN. Οι οργανισμοί που χρησιμοποιούν εσωτερικά δεδομένα πρέπει να ενσωματώσουν μηχανισμούς πρόληψης προκειμένου να μετριάσουν τους κινδύνους από τυχόν διακρίσεις και άνιση μεταχείριση εκ μέρους των συστημάτων TN.

2.1.4.2 Εξωτερικά δεδομένα

Αναφέρεται σε δεδομένα που υπάρχουν εκτός της δομής του οργανισμού ή που ενσωματώνονται με αποκλειστικό σκοπό την ανάπτυξη και την εκπαίδευση του συστήματος TN. Εάν τα δεδομένα ενσωματώνονται στη συνέχεια, τα εξωτερικά δεδομένα αναφέρονται σε δεδομένα που δεν θα υπήρχαν στον οργανισμό εάν δεν υπήρχε η ανάπτυξη της TN (π.χ. εικόνες όρασης υπολογιστή). Πολλοί από τους κινδύνους που σχετίζονται με την προστασία της ιδιωτικής ζωής και την ασφάλεια μετριάζονται όταν τα δεδομένα είναι εξωτερικά του οργανισμού. Ωστόσο, με τα εξωτερικά

δεδομένα, οι οργανισμοί πρέπει να διασφαλίζουν την ποιότητα των δεδομένων, εάν το σύστημα έχει αναπτυχθεί από τρίτους, ώστε να αποφεύγεται η διακινδύνευση σε θέματα δεοντολογίας και εμπιστοσύνης.

2.1.5 Εργασίες

Γενικά η λεπτομερής δομή και η χαμηλή πολυπλοκότητα διευκολύνουν την αυτοματοποίηση των εργασιών. Ο δημόσιος τομέας, ο οποίος είναι σε μεγάλο βαθμό γραφειοκρατικός και συχνά τυποποιημένος σε διαδικασίες, παρουσιάζει ένα μοναδικό σύνολο εργασιών που είναι επαναλαμβανόμενες από τη φύση τους. Το εν λόγω άρθρο διακρίνει τρεις κατηγορίες τέτοιων εργασιών.

2.1.5.1 Επαναλαμβανόμενες εργασίες

Όπως υποδηλώνει το όνομα, αναφέρεται σε ιδιαίτερα τυποποιημένες απλές διαδικασίες που συχνά απαιτούν επανάληψη ή επανάληψη μιας πολύ παρόμοιας αναζήτησης. Παραδείγματα αυτών είναι ένα μεγάλο μέρος των διοικητικών διαδικασιών ή η σύνταξη απλών νομικών εγγράφων. Τα συστήματα τεχνητής νοημοσύνης που εκτελούν επαναλαμβανόμενες εργασίες θέτουν λιγότερες προκλήσεις από άλλα, καθώς σπάνια εμπλέκονται σε ηθικά διλήμματα όπου μια αυτοματοποιημένη απόφαση θα είχε ενδεχομένως αρνητικές ηθικές συνέπειες. Τα συστήματα τεχνητής νοημοσύνης αυτής της κατηγορίας περιλαμβάνουν και άλλες δημόσιες υπηρεσίες εκτός από τα επαναλαμβανόμενα διοικητικά καθήκοντα, όπως: διαχείριση φωτεινών σηματοδοτών βάσει βίντεο/εικόνας, βιομετρική ταυτοποίηση, έξυπνοι βοηθοί (chatbots) ή ακόμη και υποβοηθούμενη ιατρική ή νομική σύνταξη.

2.1.5.2 Προγνωστικές μη εκτελέσιμες εργασίες

Αναφέρεται στις διαδικασίες που προβαίνουν σε μια σύσταση που δεν έχει άμεσες ή έμμεσες εκτελέσιμες συνέπειες. Συχνά αφορά περιπτώσεις όπου οι προγνωστικές μη εκτελέσιμες αυτοματοποιημένες εργασίες δημιουργούν μια εσωτερική (εντός της δημόσιας υπηρεσίας) υπο-διαδικασία. Τέτοια παραδείγματα περιλαμβάνουν την έξυπνη πρόβλεψη της ζήτησης ενέργειας και νερού, την έξυπνη πρόβλεψη της κυκλοφορίας, ιατρικές συσκευές ΙΟΤ ή ακόμη και έλεγχο αγροτικών καλλιεργειών ή επιδημιών.

2.1.5.3 Προγνωστικές εκτελέσιμες εργασίες

Αναφέρεται στις διαδικασίες που δημιουργούν μια πρόβλεψη που έχει άμεσο ενεργητικό αντίκτυπο σε ένα άτομο ή κοινότητα. Σε αυτή την υποκατηγορία πολλά ζητήματα σχετίζονται με ηθικά διλήμματα, ευθύνη, λογοδοσία και κυρίως με τις διακρίσεις ΤΝ. Επιπλέον, ο κίνδυνος εφησυχασμού σε ένα σύστημα "human-in-the-loop" είναι αναπόφευκτος. Παραδείγματα τέτοιων συστημάτων είναι τα συστήματα υποστήριξης ιατρικών διαγνώσεων, τα εργαλεία ανίχνευσης απάτης, τα λογισμικά αναγνώρισης προσώπου και εικόνας ή η αξιολόγηση κινδύνου για κάποια πληθυσμιακή ομάδα.

2.1.6 Αντίκτυπος της χρήσης ευφύων συστημάτων υποστήριξης

Εκτός από τα προαναφερθέντα κέρδη αποδοτικότητας και αποτελεσματικότητας εντός της δημόσιας υπηρεσίας καθ' εαυτήν, τα συστήματα ΤΝ έχουν ευρύτερο αντίκτυπο. Υπάρχουν επιπτώσεις στη σχέση της με τον διοικούμενο, καθώς μπορεί να μετασχηματίσει είτε θετικά την εμπειρία του ως τελικού χρήστη, παράγοντας ακριβέστερο και ικανοποιητικότερο αποτέλεσμα, είτε αρνητικά, προκαλώντας απώλεια της εμπιστοσύνης προς την παρεχόμενη δημόσια υπηρεσία. Η κοινωνική αποδοχή αυτών των χρήσεων σχετίζεται σε μεγάλο βαθμό με τις δεοντολογικές προκλήσεις

των εφαρμογών TN (ευθύνη, λογοδοσία, διακρίσεις, προκατάληψη, δικαιοσύνη και επίλυση διλημάτων) αλλά και με τις προσδοκίες των χρηστών.

Τέλος, ο αντίκτυπος των συστημάτων TN στις δημόσιες υπηρεσίες δεν αφορά αποκλειστικά τον τελικό χρήστη. Η εισαγωγή των τεχνολογικών εξελίξεων επιφέρει αναστάτωση στην αγορά εργασίας, ιδίως σε θέσεις που υπόκεινται σε αυτοματοποίηση. Η υποκατάσταση και ο μετασχηματισμός του εργατικού δυναμικού είναι από τις πιο αναπόφευκτες προκλήσεις της TN. Επιπλέον, εκτός από τις ανησυχίες των εργαζομένων, υπάρχει το πρόβλημα εφησυχασμού κατά την χρήση ευφών συστημάτων υποστήριξης, όταν οι δημόσιοι λειτουργοί επαναπαύονται και δεν παρεμβαίνουν σε ημιαυτόματες αποφάσεις.

Κατηγορία	Υποκατηγορία	Περιγραφή	Πιθανά Έργα
Οργανωσιακές Ικανότητες	-	<ul style="list-style-type: none"> - Περιλαμβάνουν τόσο προσωπικές όσο και τεχνικές ικανότητες. - Απαιτείται μια στρατηγική πορεία δράσης για τη δημιουργία αξίας. - Ένα καλά εκπαιδευμένο και ποικίλομορφο εργατικό δυναμικό είναι κρίσιμο. - Οι οργανισμοί με χαμηλής ποιότητας και αδόμητα δεδομένα δεν είναι κατάλληλοι για συστήματα TN. 	<ul style="list-style-type: none"> - Εκπαίδευση στελεχών σε TN - Ανάπτυξη εσωτερικών πολιτικών για TN - Δημιουργία ομάδας ειδικών στην TN
Δεδομένα	Εσωτερικά Δεδομένα	<ul style="list-style-type: none"> - Αναφέρεται σε δεδομένα που υπάρχουν στη δομή του οργανισμού πριν από την ανάπτυξη του συστήματος TN. - Πρέπει να είναι διαθέσιμα, προσβάσιμα και αναλύσιμα. - Ο χειρισμός των δεδομένων απαιτεί προσοχή για να αποφευχθούν διακρίσεις από την TN. 	<ul style="list-style-type: none"> - Βελτίωση της ποιότητας των εσωτερικών δεδομένων - Προστασία ευαίσθητων δεδομένων - Εμπλουτισμός δεδομένων με εξωτερικά δεδομένα
Δεδομένα	Εξωτερικά Δεδομένα	<ul style="list-style-type: none"> - Αναφέρεται σε δεδομένα που υπάρχουν εκτός της δομής του οργανισμού ή που ενσωματώνονται για την ανάπτυξη του συστήματος TN. - Απαιτείται διασφάλιση της ποιότητας των δεδομένων, ειδικά όταν αναπτύσσονται από τρίτους. 	<ul style="list-style-type: none"> - Συνεργασία με εξωτερικούς παρόχους δεδομένων - Ανάλυση και διαχείριση εξωτερικών δεδομένων - Προσομοίωση δεδομένων για εκπαίδευση TN
Εργασίες	Επαναλαμβανόμενες Εργασίες	<ul style="list-style-type: none"> - Αναφέρεται σε τυποποιημένες απλές διαδικασίες που απαιτούν επανάληψη. - Συστήματα TN που εκτελούν αυτές τις εργασίες έχουν λιγότερες ηθικές προκλήσεις. 	<ul style="list-style-type: none"> - Διαχείριση φωτεινών σηματοδοτών - Βιομετρική ταυτοποίηση - Έξυπνοι βοηθοί (chatbots)

Κατηγορία	Υποκατηγορία	Περιγραφή	Παραδείγματα
Εργασίες	Προγνωστικές Μη Εκτελέσιμες Εργασίες	<ul style="list-style-type: none"> - Αναφέρεται σε διαδικασίες που δημιουργούν συστάσεις χωρίς άμεσες αναλήψιμες συνέπειες. - Μπορεί να δημιουργήσει εσωτερικές υπο-διαδικασίες στις δημόσιες υπηρεσίες. - Το “τεχνικό μαύρο κουτί” μπορεί να αποτελεί πρόκληση σε αυτές τις περιπτώσεις. 	<ul style="list-style-type: none"> - Πρόβλεψη ζήτησης ενέργειας - Έξυπνη πρόβλεψη κυκλοφορίας - Έλεγχος καλλιεργειών με ΙοΤ.
Εργασίες	Προγνωστικές Εκτελέσιμες Εργασίες	<ul style="list-style-type: none"> - Αναφέρεται σε διαδικασίες που δημιουργούν προβλέψεις με άμεσο αντίκτυπο σε άτομα ή κοινότητες. - Συχνά σχετίζονται με ηθικά διλήμματα, ευθύνη και διακρίσεις. - Το “νομικό μαύρο κουτί” μπορεί να αποτελεί κρίσιμο παράγοντα για τη διαφάνεια και λογοδοσία. 	<ul style="list-style-type: none"> - Υποστήριξη ιατρικών διαγνώσεων - Ανίχνευση απάτης - Αξιολόγηση πληθυσμού σε κίνδυνο.
Επιπτώσεις	-	<ul style="list-style-type: none"> - Τα συστήματα ΤΝ έχουν αντίκτυπο στην εμπειρία του τελικού χρήστη. - Ο αντίκτυπος αυτός μπορεί να είναι θετικός ή αρνητικός, ανάλογα με την εμπιστοσύνη στις παρεχόμενες υπηρεσίες. - Η εισαγωγή της ΤΝ μπορεί να δημιουργήσει αναστάτωση στην αγορά εργασίας, με υποκατάσταση ή μετασχηματισμό των εργαζομένων. 	<ul style="list-style-type: none"> - Αξιολόγηση εμπειρίας τελικού χρήστη - Διαχείριση κοινωνικών αντιδράσεων - Προσαρμογή της αγοράς εργασίας

Πίνακας 2.1. Πλαίσιο ταξινόμησης για τη χρήση εφαρμογών ΤΝ από τον δημόσιο τομέα

2.2 Η ΤΝ στην υπηρεσία της κοινοβουλευτικής λειτουργίας

Όπως εκτέθηκε αναλυτικά, πολλές κυβερνήσεις χρησιμοποιούν πλέον συστηματικά την ΤΝ κατά την υλοποίηση του έργου τους στο πλαίσιο άσκησης της εκτελεστικής τους εξουσίας. Οι αστυνομικές δυνάμεις, οι στρατοί και οι τελωνειακές υπηρεσίες χρησιμοποιούν την τεχνολογία για λόγους ασφαλείας, ενώ τα νοσοκομεία την αναπτύσσουν για την υποστήριξη των διαδικασιών περίθαλψης, τα υπουργεία υποδομών για τη βελτίωση του δημόσιου χώρου και οι τοπικές κυβερνήσεις για έργα έξυπνων πόλεων, για να αναφέρουμε κάποια μόνο παραδείγματα².

Μια άλλη πρωτότυπη εφαρμογή, στον τομέα της νομοθετικής εξουσίας αυτή τη φορά, παρου-

²Sheikh, H., Prins, C., Schrijvers, E., 2023, Introduction. In: Mission AI. Research for Policy. Springer, Cham. https://doi.org/10.1007/978-3-031-21448-6_1

σιάζει η εργασία “Disruptive Technologies for Parliaments: A Literature Review”³ που προτείνει ότι οι δημόσιοι οργανισμοί θα μπορούσαν να χρησιμοποιήσουν τις λεγόμενες “ανατρεπτικές τεχνολογίες” (disruptive technologies) που μπορούν να επιφέρουν ριζικές αλλαγές στη λειτουργία των δημόσιων οργανισμών και, πιο συγκεκριμένα, στα κοινοβούλια. Η συγκεκριμένη δημοσίευση παρέχει μια βιβλιογραφική ανασκόπηση για την χρήση ανατρεπτικών τεχνολογιών στα κοινοβούλια που σκιαγραφεί τη σημασία των ανατρεπτικών τεχνολογιών, όπως την τεχνητή νοημοσύνη (AI), το Διαδίκτυο των πραγμάτων (IoT) και τα συστήματα συστάσεων (RS) για τους δημόσιους οργανισμούς και τον πιθανό αντίκτυπό τους στις διαδικασίες χάραξης πολιτικής, ενώ τονίζει την ανάγκη για προσεγγίσεις με επίκεντρο τον χρήστη. Η δημοσίευση κατηγοριοποιεί διάφορους τύπους τέτοιων τεχνολογικών χρήσεων με στόχο τη δημιουργία έξυπνων οικοσυστημάτων, την ενίσχυση των δημόσιων υπηρεσιών και τη συμβολή στην έξυπνη διακυβέρνηση. Επισημαίνει ότι η εφαρμογή διασπαστικών τεχνολογιών σε κοινοβουλευτικά περιβάλλοντα, απαιτεί ένα πλαίσιο ψηφιακού μετασχηματισμού και μια απτή ψηφιακή στρατηγική και αναγνωρίζει την πολυπλοκότητα της δημόσιας πολιτικής και τις προκλήσεις στην ενσωμάτωση ανατρεπτικών τεχνολογιών, όπως η τεχνητή νοημοσύνη, στις κοινοβουλευτικές διαδικασίες. Διερευνά τις ειδικές εφαρμογές της ΤΝ στα κοινοβούλια, η οποία έχει τη δυνατότητα να συμβάλει σημαντικά στον τομέα της δημοκρατίας με τη βελτίωση της συμμετοχής του κοινού, την ανάλυση χάραξης πολιτικής, τη σύνταξη νομοθεσίας και την κοινοβουλευτική εποπτεία.

Το κείμενο συζητά τον αντίκτυπο των ψηφιακών τεχνολογιών, ιδιαίτερα των ανατρεπτικών τεχνολογιών, υπογραμμίζει τη θετική επιρροή τους στην παροχή δημόσιων υπηρεσιών και δημοκρατικής διακυβέρνησης και τονίζει την ανάγκη για μια προσέγγιση χωρίς αποκλεισμούς, με επίκεντρο τον χρήστη κατά την υιοθέτηση αυτών των τεχνολογιών. Τα ερωτήματα που τίθενται, επικεντρώνονται στην επίτευξη δημοκρατικού μετασχηματισμού, στον εντοπισμό κατάλληλων οργανωτικών μοντέλων για τα κοινοβούλια και στη διερεύνηση της αποτελεσματικής χρήσης των ανατρεπτικών τεχνολογιών.

Ο κοινοβουλευτικός μετασχηματισμός θα συμβάλλει κατά τη χάραξη πολιτικής και θα οδηγήσει σε ένα “έξυπνο κοινοβούλιο”, ενώ η περαιτέρω προσέγγιση με επίκεντρο τον χρήστη θα διασφαλίσει την αποτελεσματική χρήση της ΤΝ στα κοινοβούλια. Αυτό σημαίνει ότι τα κοινοβούλια θα πρέπει να εμπλέκουν τους πολίτες στο σχεδιασμό και την ανάπτυξη των συστημάτων ΤΝ και ότι θα πρέπει να παρέχουν εκπαίδευση και υποστήριξη στους βουλευτές και το προσωπικό σχετικά με την αποτελεσματική χρήση της ΤΝ. Τονίζεται επίσης η σημασία μιας προσέγγισης χωρίς αποκλεισμούς, αναγνωρίζοντας τις διαφορετικές ανάγκες που έχουν αφενός οι υπεύθυνοι λήψης αποφάσεων και αφετέρου οι πολίτες στην υιοθέτηση ανατρεπτικών τεχνολογιών στο κοινοβουλευτικό περιβάλλον.

Οι συγγραφείς πραγματοποίησαν βιβλιογραφική ανασκόπηση 48 ερευνητικών εργασιών και διαπίστωσαν ότι οι ανατρεπτικές τεχνολογίες, στις οποίες συγκαταλέγεται η τεχνητή νοημοσύνη, μπορούν να αυξήσουν την αποτελεσματικότητα και τη διαφάνεια των κοινοβουλίων, καθώς και να βελτιώσουν την ποιότητα των δημόσιων υπηρεσιών, συχνά αλληλοσυμπληρούμενες μεταξύ τους.

Γενικότερα η ΤΝ μπορεί να χρησιμοποιηθεί στα κοινοβούλια με τους εξής τρόπους:

(α) Δημόσια συμμετοχή

Η ΤΝ μπορεί να χρησιμοποιηθεί για τη δημιουργία πιο διαδραστικών και ελκυστικών τρόπων συμμετοχής των πολιτών στη δημοκρατική διαδικασία. Για παράδειγμα, η ΤΝ μπορεί να χρησιμοποιηθεί για την ανάπτυξη chatbots που απαντούν σε ερωτήσεις των πολιτών ή για τη δημιουργία προσομοιώσεων εικονικής πραγματικότητας που επιτρέπουν στους πολίτες να βιώσουν τη νομοθετική διαδικασία.

³Koryzis, Dimitris & Margaris, Dionisis & Vassilakis, Costas & Kotis, Konstantinos & Spiliotopoulos, Dimitris. (2023), Disruptive Technologies for Parliaments: A Literature Review, Future Internet, 15. 66.10.3390/fi15020066, https://www.researchgate.net/publication/368321578_Disruptive_Technologies_for_Parliaments_A_Literature_Review

(β) Ανάλυση πολιτικής

Η τεχνητή νοημοσύνη μπορεί να χρησιμοποιηθεί για την ανάλυση μεγάλου όγκου δεδομένων για τον εντοπισμό τάσεων και μοτίβων που μπορούν να ενημερώσουν για πολιτικές αποφάσεις. Για παράδειγμα, η τεχνητή νοημοσύνη μπορεί να χρησιμοποιηθεί για την ανάλυση δημοσκοπήσεων κοινής γνώμης, δεδομένων των μέσων κοινωνικής δικτύωσης και οικονομικών δεδομένων για τον εντοπισμό τομέων στους οποίους απαιτούνται αλλαγές πολιτικής βάσει καταγεγραμμένων παραδειγμάτων.

(γ) Νομοθετική επεξεργασία

Η τεχνητή νοημοσύνη μπορεί να χρησιμοποιηθεί για την αποτελεσματικότερη και ακριβέστερη σύνταξη νομοθεσίας. Για παράδειγμα, η τεχνητή νοημοσύνη μπορεί να χρησιμοποιηθεί για τον εντοπισμό πιθανών νομικών συγκρούσεων σε σχέδια νομοθεσίας ή για τη δημιουργία εναλλακτικών διατυπώσεων για σχέδια διατάξεων.

(δ) Κοινοβουλευτική εποπτεία

Η ΤΝ μπορεί να χρησιμοποιηθεί για την παρακολούθηση της εφαρμογής των νόμων και για τον εντοπισμό τομέων όπου υπάρχει ανάγκη επιβολής. Για παράδειγμα, η ΤΝ μπορεί να χρησιμοποιηθεί για την ανάλυση κυβερνητικών δεδομένων για τον εντοπισμό περιπτώσεων διαφθοράς ή απάτης. Συμπεραίνουμε λοιπόν ότι η τεχνητή νοημοσύνη έχει τη δυνατότητα να φέρει ρηξικέλευθες αλλαγές στον τρόπο λειτουργίας των κοινοβουλίων και ότι μπορεί να συμβάλει στη βελτίωση της δημοκρατίας και των δημόσιων υπηρεσιών. Ωστόσο πρέπει να χρησιμοποιείται προσεκτικά και υπεύθυνα και τα κοινοβούλια πρέπει να διασφαλίσουν ότι τα συστήματα ΤΝ είναι διαφανή, υπεύθυνα και δίκαια και ότι δεν παραβιάζουν τα δικαιώματα των πολιτών.

2.3 Το παράδοξο της ρύθμισης της ΤΝ και οι εκλογές στο παγκόσμιο τοπίο το 2024

Ενδιαφέρον παρουσιάζει η τοποθέτηση του κοσμήτορα στις παγκόσμιες επιχειρήσεις της Σχολής Δικαίου και Διπλωματίας Fletcher του Πανεπιστημίου Tufts, σε άρθρο⁴ όπου εξετάζει την απειλή της παραπληροφόρησης στο πλαίσιο των προεκλογικών εκστρατειών, εστιάζοντας ιδιαίτερα στη χρήση της τεχνητής νοημοσύνης για την κατασκευή και τη διάδοση ψευδών πολιτικών πληροφοριών. Επισημαίνει παραδείγματα παραπληροφόρησης που διαδόθηκαν από τον πρώην πρόεδρο Donald Trump και τον κυβερνήτη της Φλόριντα Ron DeSantis, καθώς και τη ευρεία πλέον διαθεσιμότητα εργαλείων τεχνητής νοημοσύνης που επιτρέπουν στον καθένα να δημιουργεί και να διανέμει πολιτικό περιεχόμενο. Ο συγγραφέας περιγράφει το φαινόμενο αυτό, με ένα οξύμωρο σχήμα, ως “εκδημοκρατισμό” της παραπληροφόρησης, καθώς αποτελεί ένα μέσο που πλέον βρίσκεται στη διάθεση οποιουδήποτε, και τονίζει ότι η παραπληροφόρηση με τη βοήθεια της ΤΝ μπορεί να αποτελέσει τη σοβαρότερη απειλή για τη λειτουργία των ήδη βαλλόμενων από την παραπληροφόρηση δυτικών δημοκρατιών.

Επισημαίνει όμως ένα παράδοξο στη ρύθμιση της παραπληροφόρησης και διατυπώνει τη γνώμη ότι όσο περισσότερο ρυθμίζεται στη Δύση, τόσο επιδεινώνεται σε παγκόσμιο επίπεδο. Αποδίδει το γεγονός αυτό στο ότι οι κύριοι φορείς της παραπληροφόρησης, οι μεγάλες πλατφόρμες κοινωνικής δικτύωσης, μειώνουν σταθερά το προσωπικό τους για την καταγραφή της παραπληροφόρησης σε παγκόσμιο επίπεδο και στρέφουν τους πόρους τους στην παρακολούθηση τέτοιων φαινομένων

⁴The AI Regulation Paradox - Regulating artificial intelligence to protect U.S. democracy could end up jeopardizing democracy abroad, Bhaskar Chakravorti (dean of global business at Tufts University's Fletcher School of Law and Diplomacy, 4 Αυγούστου 2023, <https://foreignpolicy.com/2023/08/04/ai-regulation-artificial-intelligence-democracy-elections/>

στον Δυτικό κόσμο, ακριβώς για να αυμορφώνονται και να ευθυγραμμίζονται με το διαρκώς αυστηρότερο κανονιστικό πλαίσιο που αναπτύσσεται από τις Δυτικά κράτη. Εκφράζει την άποψη λοιπόν ότι αυτές οι αποδυναμωμένες ομάδες εποπτείας πρέπει να προσηλώνονται σε πιο καίριους στόχους, δηλαδή τη συμμόρφωση με τους νομοθέτες και τις ρυθμιστικές αρχές στις Ηνωμένες Πολιτείες και την ΕΕ.

Το αποτέλεσμα, κατά τη γνώμη του, είναι να μην έχουν απομείνει αρκετοί πόροι για την παρακολούθηση του περιεχομένου στον υπόλοιπο κόσμο και επικαλείται δημοσίευμα της Wall Street Journal⁵ που αποκάλυψε ότι στο Facebook το 2020, το 87 τοις εκατό του χρόνου για τον έλεγχο του περιεχομένου αναλώθηκε σε αναρτήσεις από τις Ηνωμένες Πολιτείες, μια κατανομή δυσανάλογη, δεδομένου ότι το 90 τοις εκατό των χρηστών του Facebook βρίσκονται εκτός των Ηνωμένων Πολιτειών.

Μια άλλη ενδιαφέρουσα διαπίστωση για το παράδοξο της ισχύος της ΤΝ κάνουν οι Ian Bremmer και Mustafa Suleyman στο ομώνυμο άρθρο τους “The AI power paradox”, όπου παρουσιάζουν το πώς η τεχνητή νοημοσύνη είναι διαφορετική από άλλες τεχνολογίες και επιπλέον διαφορετική ως προς την επίδρασή της στην εξουσία⁶. Δεν θέτει απλώς προκλήσεις πολιτικής, αλλά ο ανατρεπτικός της χαρακτήρας καθιστά επίσης την επίλυση αυτών των προκλήσεων προοδευτικά δυσκολότερη, κάτι που αποκαλούν το παράδοξο της ισχύος της ΤΝ. Επισημαίνουν ότι ο ρυθμός προόδου είναι εκπληκτικός σε βαθμό που να καθιστά παρωχημένο τον νόμο του Moore, ο οποίος προέβλεψε με επιτυχία τον διπλασιασμό της υπολογιστικής ισχύος κάθε δύο χρόνια. Σε απόδειξη αυτού αναφέρουν ότι όταν το OpenAI ξεκίνησε το πρώτο μεγάλο γλωσσικό μοντέλο του, γνωστό ως GPT-1, το 2018, είχε 117 εκατομμύρια παραμέτρους. Πέντε χρόνια αργότερα, το μοντέλο τέταρτης γενιάς της εταιρείας, το GPT-4, θεωρείται ότι έχει πάνω από ένα τρισεκατομμύριο. Το ποσό των υπολογισμών που χρησιμοποιούνται για την εκπαίδευση των πιο ισχυρών μοντέλων τεχνητής νοημοσύνης αυξάνεται κατά δέκα φορές ετησίως, τα τελευταία δέκα χρόνια⁷. Με άλλα λόγια, τα πιο προηγμένα μοντέλα τεχνητής νοημοσύνης σήμερα -γνωστά και ως μοντέλα “αιχμής”- χρησιμοποιούν πέντε δισεκατομμύρια φορές την υπολογιστική ισχύ των μοντέλων αιχμής συγκριτικά με πριν από μια δεκαετία. Η επεξεργασία που κάποτε διαρκούσε εβδομάδες τώρα γίνεται σε δευτερόλεπτα.

Και όλα αυτά συμπίπτουν με το 2024, μια χρονιά με εκλογές⁸ σε 50 χώρες του κόσμου κατά την οποία περισσότεροι από 2 δισεκατομμύρια ψηφοφόροι θα κατευθυνθούν στις κάλπες σε χώρες όπως οι Ηνωμένες Πολιτείες, η Ινδία, το Μεξικό και η Νότια Αφρική. Ειδικά τον Ιούνιο 2024 θα διεξαχθούν οι μεγαλύτερες διασυνοριακές εκλογές στον κόσμο, όταν περισσότεροι από 400 εκατομμύρια άνθρωποι αναμένεται να ψηφίσουν στις εκλογές για την ανάδειξη του Ευρωπαϊκού Κοινοβουλίου⁹. Ομοίως ψηφοφορίες υψηλού κινδύνου που έχουν προγραμματιστεί σε πολυπληθή έθνη όπως η Ινδία και η Ινδονησία, σε κράτη με ιδιαίτερα καθεστώτα όπως η Λευκορωσία, το Ιράν και η Ρωσία και οι εκλογές στρατηγικών συμμάχων όπως η Ταϊβάν και Ηνωμένο Βασίλειο, έχουν τη δυνατότητα να επηρεάσουν βαθιά την Ευρωπαϊκή ήπειρο και την παγκόσμια οικονομία. Το σκάνδαλο της Cambridge Analytica κρούει τον κώδωνα του κινδύνου ως προς τον βαθμό που είναι ευάλωτη η πολιτική στην κατάχρηση των κοινωνικών μέσων. Αυτό που συνέβη ήταν ότι συλλέχθηκαν προσωπικά δεδομένα από εκατομμύρια χρήστες του Facebook χωρίς τη συγκατάθεσή τους από τη βρετανική εταιρεία συμβούλων Cambridge Analytica κατά τη δεκαετία του 2010.

⁵Employees raised alarms about how the site (Facebook) is used in developing countries, where its user base is already huge and expanding, <https://www.wsj.com/articles/facebook-drug-cartels-human-traffickers-response-is-weak-documents-11631812953>

⁶BREMMER, Ian; SULEYMAN, Mustafa. The AI power paradox: Can states learn to govern artificial intelligence before it's too late? Foreign Affairs, 2023, vol. 102, number 5, σελ. 26, <https://www.foreignaffairs.com/world/artificial-intelligence-power-paradox>

⁷Όπως παραπάνω, σελ. 29

⁸2024 is a record year for elections, 15 Δεκεμβρίου 2023, <https://www.weforum.org/agenda/2023/12/2024-elections-around-world/>

⁹Five elections set to shape Europe in 2024, the biggest ballot year in history, 27 Δεκεμβρίου 2023, <https://www.euronews.com/my-europe/2023/12/27/five-elections-set-to-shape-europe-in-2024-the-biggest-ballot-year-in-history>

Τα δεδομένα αξιοποιήθηκαν κυρίως για πολιτική διαφήμιση πελατών της εταιρείας, που περιλάμβαναν την εκστρατεία Τραμπ και τους υποστηρικτές του Brexit, γεγονός που έφερε στο προσκήνιο τις προκλήσεις της ψυχολογικής στόχευσης και τον κίνδυνο του δημοσίου συμφέροντος, που παραμένει εξαιρετικά επίκαιρος, για τις οποίες υπήρχαν προειδοποιήσεις από τους επιστήμονες¹⁰.

2.4 Αλγοριθμική μεροληψία και κράτος δικαίου: the Dutch scandal

Σε συνέχεια των παραπάνω θα ακολουθήσει μια ανάλυση χαρακτηριστικού παραδείγματος που προέρχεται από τον πίνακα 2.2 και πιο συγκεκριμένα της περίπτωσης του σχεδίου SyRI που εφαρμόστηκε στην Ολλανδία. Ειδικότερα η ανάλυση των συνολικών πτυχών της εφαρμογής, που προέρχεται από το πόρισμα της Επιτροπής της Βενετίας¹¹ (Επιτροπή του Συμβουλίου της Ευρώπης για τη Δημοκρατία μέσω του Δικαίου), ανταποκρίνεται στους γενικότερους όρους του πλαισίου εφαρμογής της TN από δημόσιες υπηρεσίες, όπως τους αναλύει το άρθρο του Omar Ballester παραπάνω (ενότητα 2.1), το οποίο τονίζει εξίσου τη μεγάλη σημασία της οργανωτικής δομής του φορέα, πέραν των εργασιών προς εκτέλεση και των δεδομένων.

Η διευρυνόμενη χρήση της TN σε συνέχεια όσων αναφέρθηκαν προσφέρει πολλές δυνατότητες για τις κυβερνήσεις σε όλο τον κόσμο. Οι παραδοσιακές μορφές παροχής υπηρεσιών, χάραξης πολιτικής και επιβολής του νόμου μπορούν να αλλάξουν γρήγορα με την εισαγωγή τεχνολογιών τεχνητής νοημοσύνης στις κυβερνητικές πρακτικές και στα οικοσυστήματα του δημόσιου τομέα. Για παράδειγμα, οι κυβερνήσεις μπορούν να χρησιμοποιήσουν τεχνολογίες τεχνητής νοημοσύνης για να βελτιώσουν την ποιότητα των δημόσιων υπηρεσιών, για την ενίσχυση της εμπιστοσύνης των πολιτών και για να αυξηθεί η αποδοτικότητα και η αποτελεσματικότητα στην παροχή υπηρεσιών. Η τεχνητή νοημοσύνη μπορεί επίσης να χρησιμοποιηθεί από τις κυβερνήσεις για τη δημιουργία ακριβέστερων προβλέψεων και για την προσομοίωση πολύπλοκων συστημάτων που επιτρέπουν τον πειραματισμό με διάφορες επιλογές πολιτικής. Μπορεί να συνεισφέρει σε πολλαπλούς λειτουργικούς τομείς της κυβέρνησης, όπως η υποστήριξη αποφάσεων, οι μεταφορές, η δημόσια υγεία και η επιβολή του νόμου.

Ταυτόχρονα, η χρήση AI στην κυβέρνηση δημιουργεί προκλήσεις καθώς μπορεί να αυξήσει την εμπιστοσύνη των πολιτών προς τις κυβερνήσεις, μπορεί όμως και να την μειώσει λόγω κάποιων περιπτώσεων παραβίασης του απορρήτου των πολιτών ή αδικίας κατά τη χρήση της στη δημόσια διακυβέρνηση. Επιπλέον, πρόσθετες προκλήσεις προκύπτουν από την έλλειψη διαφάνειας του φαινομένου του “μαύρου κουτιού”, όταν η τεχνητή νοημοσύνη χρησιμοποιείται στη λήψη αποφάσεων από τις κυβερνήσεις. Αυτοί οι υπαρκτοί κίνδυνοι αυξάνουν το διακύβευμα για τις κυβερνήσεις, καθώς οι αποτυχημένες χρήσεις τεχνητής νοημοσύνης από την κυβέρνηση μπορεί να έχουν έντονες αρνητικές επιπτώσεις τόσο για τις κυβερνήσεις όσο και για την κοινωνία¹².

Για τον λόγο αυτό η κοινωνία των πολιτών και η ακαδημαϊκή κοινότητα εκφράζουν συχνά δικαιολογημένο προβληματισμό ιδιαίτερα με την αφορμή κάποιων αντιπροσωπευτικών παραδειγμάτων κακής εφαρμογής εξαιτίας αλγοριθμικών σφαλμάτων, όπως συνέβη με την Ολλανδική κυβέρνηση¹³. Εν προκειμένω, η εν λόγω αντιπαράθεση έφτασε να λάβει τεράστια νομική διάσταση όταν

¹⁰Σκάνδαλο δεδομένων Facebook-Cambridge Analytica, <https://tinyurl.com/3wcb5tze>

¹¹Netherlands-Opinion on the Legal Protection of Citizens, adopted by Venice Commission at its 128th Plenary Session, Στρασβούργο, 18 Οκτωβρίου 2021, [https://www.venice.coe.int/webforms/documents/?pdf=CDL-AD\(2021\)031-e](https://www.venice.coe.int/webforms/documents/?pdf=CDL-AD(2021)031-e)

¹²Anneke Zuiderwijk, Yu-Che Chen, Fadi Salem, Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda, Government Information Quarterly, Volume 38, Issue 3, 2021, 101577, ISSN 0740-624X, <https://doi.org/10.1016/j.giq.2021.101577>, <https://www.sciencedirect.com/science/article/pii/S0740624X21000137>

¹³Dutch scandal serves as a warning for Europe over risks of using algorithms, 23 Μαρτίου 2022, <https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/>

δικαστική απόφαση¹⁴ απαγόρευσε τη χρήση της τεχνολογίας SyRI (που προορίζεται για τον εντοπισμό της απάτης) μετά το λεγόμενο “ολλανδικό σκάνδαλο των επιδομάτων παιδικής μέριμνας” (Toeslagenaffaire), που προκλήθηκε από τη χρήση αλγορίθμων από την ολλανδική φορολογική διοίκηση για τον εντοπισμό υποτιθέμενων παράνομων απονομών επιδόματος παιδικής μέριμνας. Αυτό κατέληξε στο να κατηγορηθούν άδικα για απάτη χιλιάδες γονείς, να βρεθούν αντιμέτωποι με δυσβάστακτα πρόστιμα, ορισμένοι να οδηγηθούν στην αυτοκτονία και εν τέλει να παραιτηθεί η κυβέρνηση κάτω από το βάρος των αποκαλύψεων¹⁵.

2.4.1 The Dutch scandal: συνοπτικό ιστορικό

Η περίπτωση του επιδόματος παιδικής φροντίδας στις Κάτω Χώρες περιλαμβάνει ένα σύνθετο σύστημα που θεσπίστηκε το 2005, όπου οι γονείς λαμβάνουν επίδομα για τα έξοδα παιδικής φροντίδας με βάση το εισόδημα. Το επίδομα αυτό, γνωστό ως toeslag, είναι μια υπό όρους προκαταβολή που παρέχεται από το κράτος. Το σύστημα αντιμετώπισε δυσκολίες, συμπεριλαμβανομένης της απάτης, που οδήγησε σε μια αυστηρή προσέγγιση “όλα ή τίποτα” στις απαιτήσεις για επιστροφή αχρεωστήτως καταβληθέντων επιδομάτων.

Το 2013, αποκαλύφθηκε ένα σύστημα απάτης μεγάλης κλίμακας, γεγονός που οδήγησε σε μια αυστηρή πολιτική καταπολέμησης της απάτης από τη Φορολογική και Τελωνειακή Διοίκηση. Ένας αλγόριθμος βασισμένος στην τεχνητή νοημοσύνη χρησιμοποιήθηκε για τον εντοπισμό πιθανών περιπτώσεων απάτης, στοχοποιώντας υπέρμετρα αλλοδαπούς αιτούντες. Η προσέγγιση “όλα ή τίποτα” που υιοθετήθηκε είχε ως αποτέλεσμα μαζικές απαιτήσεις επιστροφής χρημάτων και παρά το γεγονός ότι ορισμένοι γονείς κέρδισαν υποθέσεις σε διοικητικό επίπεδο, το ανώτατο δικαστήριο επικύρωσε την εν λόγω άκαμπτη προσέγγιση. Η συγκεντρωτικότητα του συστήματος κατέστησε δύσκολο για τους γονείς να αναζητήσουν εξηγήσεις ή βοήθεια.

Η αποκάλυψη της απάτης οδήγησε στη δημιουργία μιας ομάδας διαχείρισης κατά της απάτης, της CAF, αλλά οι έρευνές της, όπως η CAF 11, ήταν εσφαλμένες, προκαλώντας αδικαιολόγητες παύσεις των επιδομάτων. Η έκθεση του Συνηγόρου του Πολίτη του 2017 επέκρινε τις σκληρές ενέργειες κατά των οικογενειών και συνέστησε αποζημίωση, η οποία όμως δεν προσφέρθηκε εμπράκτως.

Το 2019, το Τμήμα Διοικητικής Δικαιοδοσίας του Συμβουλίου της Επικρατείας άλλαξε την ερμηνεία του, εφαρμόζοντας τις αρχές της χρηστής διοίκησης και της αναλογικότητας. Ο πρωθυπουργός απολογήθηκε το 2021 για την ταλαιπωρία που προκλήθηκε, ανακοινώνοντας μεταρρυθμίσεις. Η κυβέρνηση παραιτήθηκε και διεξήχθησαν εκλογές. Έκθεση της Αρχής Προστασίας Δεδομένων το 2020 αποκάλυψε παραβιάσεις του ΓΚΠΔ στο σύστημα αξιολόγησης κινδύνου, το οποίο στηριζόταν σε μεροληπτικά δεδομένα. Μεταξύ του 2012 και του 2019, χιλιάδες άτομα άδικως κρίθηκαν ένοχα για απάτη.

Η έκθεση της Εξεταστικής Επιτροπής της Βουλής του 2020 κατέληξε στο συμπέρασμα ότι παραβιάστηκε το κράτος δικαίου κατά την εφαρμογή της χορήγησης επιδόματος φύλαξης παιδιών, επικρίνοντας τη φορολογική και τελωνειακή διοίκηση, τον νομοθέτη και τη δικαιοσύνη. Η διοικητική δικαιοσύνη δεν έπραξε το καθήκον της, όσον αφορά τη διασφάλιση των νόμιμων δικαιωμάτων των πολιτών, παραβιάζοντας τις γενικές αρχές της χρηστής διακυβέρνησης μέχρι τον Οκτώβριο του 2019.

¹⁴van Bakkum, M., & Borgesius, F. Z. (2021). Digital welfare fraud detection and the Dutch SyRI judgment. *European Journal of Social Security*, 23(4), 323-340.<https://doi.org/10.1177/13882627211031257>

¹⁵https://en.wikipedia.org/wiki/Dutch_childcare_benefits_scandal

2.4.2 Ανάλυση πραγματικών περιστατικών από την Επιτροπή της Βενετίας

Η Επιτροπή της Βενετίας του Συμβουλίου της Ευρώπης¹⁶ κατέγραψε το παραπάνω ιστορικό λεπτομερώς όπως εκτίθεται παρακάτω. Στις 23 Ιουνίου 2005, το ολλανδικό κοινοβούλιο ενέκρινε τον νόμο για την εναρμόνιση των συστημάτων που σχετίζονται με το εισόδημα (εφεξής “ο νόμος AWIR”), ο οποίος θέσπισε ένα σύνθετο σύστημα επιδόματος φύλαξης παιδιών, στο πλαίσιο του οποίου οι γονείς αγοράζουν συγκεκριμένες υπηρεσίες προσχολικής και εξωσχολικής φροντίδας παιδιών σε μια ρυθμιζόμενη αγορά από ένα εγγεγραμμένο κέντρο παιδικής φροντίδας (kindercentrum - π.χ. παιδικός σταθμός) ή από έναν παιδαγωγό (gastouder). Στο πλαίσιο αυτού του συστήματος, οι γονείς αποζημιώνονται για μέρος του κόστους, ανάλογα με το εισόδημά τους, ως επίδομα το οποίο εξαρτάται από την απόδειξη ότι το εισόδημα του ατόμου είναι κάτω από ένα ορισμένο επίπεδο και το οποίο καταβάλλεται μόνο κατόπιν αιτήματος. Αυτό το “επίδομα” (toeslag) είναι μια προσωρινή και υπό όρους “προκαταβολή” που καταβάλλεται από το κράτος στους αιτούντες εκ των προτέρων. Το κράτος καλύπτει μόνο ένα μέρος του κόστους της φροντίδας των παιδιών. Οι γονείς πρέπει πάντα να πληρώνουν οι ίδιοι το υπόλοιπο. Η προσωπική τους συνεισφορά είναι υποχρεωτική και πρέπει να είναι σε θέση να αποδείξουν ότι την παρείχαν με σχετικές τραπεζικές βεβαιώσεις.

Με την ψήφιση του AWIR τον Σεπτέμβριο του 2005, το νέο σύστημα συνεπαγόταν επίσης τη μεταφορά της λειτουργίας καταβολής του επιδόματος από τις κοινωνικές υπηρεσίες στη Φορολογική και Τελωνειακή Διοίκηση του Υπουργείου Οικονομικών (Belastingdienst/ Toeslagen). Χρειάστηκε να φτάσει το 2011 για να αναπτύξει η φορολογική και τελωνειακή διοίκηση την ικανότητά της και να εισαγάγει ένα σύστημα ΤΠΕ που επέτρεπε τη διενέργεια ελέγχων σε μεγαλύτερη κλίμακα. Στο πλαίσιο αυτού του συστήματος ΤΠΕ, οι εφαρμογές που θεωρούνταν ύποπτες επιλέγονταν βάσει ενός μοντέλου ταξινόμησης κινδύνου, βασισμένου στην Τεχνητή Νοημοσύνη (TN), το οποίο ήταν ένας αλγόριθμος που “αυτοεκπαιδευόταν” από παραδείγματα σωστών και λανθασμένων εφαρμογών. Ένας από τους πολλούς δείκτες που χρησιμοποιούνταν για τον εντοπισμό περιπτώσεων απάτης ήταν η ιθαγένεια και οι αιτούντες με ξένη καταγωγή επιλέγονταν από το σύστημα για λεπτομερή έλεγχο των αιτήσεών τους.

Το 2013, αποκαλύφθηκε ότι σε μια μεγάλης κλίμακας απάτη είχε τεθεί σε εφαρμογή ένα σχέδιο για τη συστηματική εξαπάτηση του ολλανδικού κράτους από τις πληρωμές κοινωνικής βοήθειας επί χρόνια. Οι εγκέφαλοι του συστήματος είχαν στρατολογήσει άτομα για να ταξιδέψουν στις Κάτω Χώρες, όπου εγγράφηκαν ως κάτοικοι και στη συνέχεια υπέβαλαν αίτηση για επιδότηση στέγασης και υγειονομικής περίθαλψης. Στη συνέχεια επέστρεφαν στην πατρίδα τους, έπαιρναν το επίδομα από μηχανήματα αυτόματων συναλλαγών και μοιράζονταν τα έσοδα με τους πρωτεργάτες. Το σύστημα ήταν σχεδιασμένο έτσι ώστε το επίδομα να καταβάλλεται κατόπιν αίτησης των γονέων, η οποία περιείχε μια εκτίμηση των εξόδων φροντίδας των παιδιών για το επόμενο έτος. Λόγω της ελάχιστης αρχικής επαλήθευσης, σχεδόν όλοι όσοι έκαναν αίτηση έπαιρναν “προκαταβολές” για το επίδομα. Στη συνέχεια γίνονταν επαληθεύσεις.

Μέχρι την 1η Ιανουαρίου 2021, το άρθρο 26 του AWIR, με τίτλο “Ανάκτηση οφειλής από ενδιαφερόμενο μέρος”, είχε ως εξής “Εάν η αναθεώρηση ενός επιδόματος ή η αναθεώρηση μιας προκαταβολής οδηγήσει σε ποσό προς ανάκτηση ή εάν ο διακανονισμός μιας προκαταβολής με ένα επίδομα οδηγήσει σε αυτό, ο ενδιαφερόμενος οφείλει ολόκληρο το ποσό της ανάκτησης”. Η διάταξη αυτή ερμηνεύτηκε από τη Φορολογική και Τελωνειακή Διοίκηση ως η λεγόμενη “προσέγγιση όλα ή τίποτα”, έτσι ώστε ακόμη και αν ένας γονέας είχε ενεργήσει καλόπιστα, αλλά ούτε ο γονέας ούτε ο φροντιστής δεν μπορούσαν να προσκομίσουν αποδείξεις για τις ώρες που χρησι-

¹⁶ Η Επιτροπή της Βενετίας (Επιτροπή του Συμβουλίου της Ευρώπης για τη Δημοκρατία μέσω του Δικαίου) αποτελεί υψηλού κύρους τεχνοκρατικό και συμβουλευτικό όργανο του Συμβουλίου της Ευρώπης για συνταγματικά θέματα με αποστολή να παρέχει νομικές συμβουλές στα κράτη μέλη της και να βοηθά τα κράτη που επιθυμούν να ευθυγραμμίσουν τις νομικές και θεσμικές δομές τους με τα ευρωπαϊκά πρότυπα και τη διεθνή εμπειρία στους τομείς της δημοκρατίας, των ανθρωπίνων δικαιωμάτων και του κράτους δικαίου. https://www.venice.coe.int/WebForms/pages/?p=01_Presentation&lang=EN

μποιήθηκαν ή τη γονική οικονομική συνεισφορά, ο γονέας έπρεπε να επιστρέψει ολόκληρο το ποσό για ολόκληρο το έτος. Η προσέγγιση “όλα ή τίποτα” οδήγησε σε υποχρέωση επιστροφής υψηλών ποσών, κάτι που συνέβη με το 15% των γονέων. Τα αιτήματα επιστροφής είχαν αναδρομική ισχύ για τα προηγούμενα πέντε έτη. Οποιαδήποτε μικρή διαφορά μεταξύ των πληρωμών που είχαν ανακοινωθεί εκ των προτέρων για το έτος και εκείνων που πράγματι πραγματοποιήθηκαν κατά τη διάρκεια του έτους, ερμηνεύτηκε ότι προέκυψε από δόλο και κατέληξε στην υποχρέωση πλήρους επιστροφής του επιδόματος (η λεγόμενη προσέγγιση “όλα ή τίποτα”). Αυτή η αυστηρή προσέγγιση δεν μετριάστηκε από κανένα έλεγχο αναλογικότητας ή ρήτρας δυσχέρειας εκπλήρωσης.

Η προσέγγιση “όλα ή τίποτα” οδήγησε σε μαζικές απαιτήσεις επιστροφής χρημάτων από τους γονείς. Σε ορισμένες περιπτώσεις, το ετήσιο ποσό που είχε καταβληθεί ήταν περίπου 30.000 ευρώ και έπρεπε να επιστραφεί ολόκληρο το ποσό, για αρκετά χρόνια αναδρομικά, παρόλο που το πραγματικά οφειλόμενο ποσό ήταν πολύ μικρότερο. Πολλοί γονείς προσέφυγαν στα διοικητικά δικαστήρια και ορισμένοι από αυτούς κέρδισαν τις υποθέσεις τους σε πρώτο βαθμό. Σε αυτές τις περιπτώσεις η Φορολογική και Τελωνειακή Διοίκηση προσέφυγε στο Συμβούλιο της Επικρατείας, το οποίο όμως επιβεβαίωσε την προσέγγιση “όλα ή τίποτα”. Οι γονείς ήταν πολύ δύσκολο να λάβουν εξηγήσεις για την προσέγγιση αυτή, καθώς το σύστημα ήταν εξαιρετικά συγκεντρωτικό. Οι τοπικές κοινωνικές υπηρεσίες δεν ήταν αρμόδιες καθώς η διαχείριση είχε ανατεθεί κεντρικά στη φορολογική και τελωνειακή διοίκηση και οι καταγγελίες μπορούσαν να γίνουν μόνο μέσω ενός τηλεφωνικού κέντρου. Φαίνεται ότι από το 2013 δόθηκαν εσωτερικές οδηγίες να μην παρασχεθεί καμία βοήθεια στους ενδιαφερόμενους γονείς.

Υπό την πίεση της κοινής γνώμης και των πολιτικών, η Φορολογική και Τελωνειακή Διοίκηση υιοθέτησε μια αυστηρή μέθοδο καταπολέμησης της απάτης και δημιούργησε μια ειδική ομάδα διαχείρισης της απάτης στο Υπουργείο Οικονομικών, την “Combiteam Aanpak Facilitators” (CAF). Οι υπάλληλοι από κάθε τμήμα των εμπλεκόμενων υπουργείων μπορούσαν να διαβιβάζουν σήματα για πιθανούς συνεργούς στην CAF. Η ομάδα CAF ξεκίνησε τη λεγόμενη έρευνα “CAF 11” κατά ενός κέντρου παιδικών σταθμών, με αποτέλεσμα πολυάριθμες επισκέψεις στο κέντρο και διοικητικούς ελέγχους. Σύμφωνα με την έκθεση της Εξεταστικής Επιτροπής της Βουλής του 2020, το επίδομα για ολόκληρες ομάδες γονέων στην υπόθεση CAF 11 διακόπηκε παράνομα, δεν υπήρξε δίκαιη αντιμετώπιση και ο χρόνος που απαιτήθηκε για την εκδίκαση των ενστάσεων ήταν υπερβολικά μεγάλος. Τελικά, δεν διαπιστώθηκε καμία απάτη. Αποδείχθηκε ότι το CAF είχε εργαστεί με την παραδοχή ότι το 80% των δικαιούχων ήταν “κακοί” και το 20% “καλοί”, παρά το γεγονός ότι μια τέτοια προσέγγιση σήμαινε ότι θα επιβάλλονταν κυρώσεις ακόμη και σε άτομα που δεν είχαν διαπράξει κανένα αδίκημα. Αυτό ονομάστηκε “προσέγγιση 80/20” ή ομαδική προσέγγιση και προκάλεσε σχετική έκθεση του Συνηγόρου του Πολίτη του 2017 και, παρά τις συστάσεις του, δεν προσφέρθηκε εμπράκτως αποζημίωση στους γονείς των οποίων το επίδομα είχε άδικα διακοπεί και οι ίδιοι οι γονείς έπρεπε να υποβάλουν αίτηση αποζημίωσης.

Τέλος, στις 23 Οκτωβρίου 2019, το Τμήμα Διοικητικής Δικαιοδοσίας του Συμβουλίου της Επικρατείας άλλαξε ριζικά την ερμηνεία του AWIR και εξέδωσε δύο αποφάσεις που εφάρμοζαν τις γενικές αρχές της χρηστής διοίκησης και την αρχή της αναλογικότητας. Στις 15 Ιανουαρίου 2021, ο Πρωθυπουργός απέστειλε επιστολή στον Πρόεδρο της Βουλής των Αντιπροσώπων στην οποία ζητούσε συγγνώμη εκ μέρους της κυβέρνησης για τις πρωτοφανείς δυσκολίες που έπρεπε να υποστούν οι γονείς και τα παιδιά τους. Ανακοίνωσε μια σειρά μεταρρυθμίσεων που αφορούν το σύστημα παροχών αλλά και γενικότερα μέτρα για τη βελτίωση του τρόπου με τον οποίο τα προειδοποιητικά σήματα μπορούν να ληφθούν πιο σοβαρά υπόψη, για τη βελτίωση της πρόσβασης στη διοίκηση, τη βελτίωση της νομοθεσίας, την αποφυγή διακρίσεων, την παροχή πληροφοριών στο Κοινοβούλιο. Την ίδια ημέρα η κυβέρνηση παραιτήθηκε. Στη συνέχεια, διεξήχθησαν εκλογές. Στις 17 Ιουλίου 2020, η Αρχή Προστασίας Δεδομένων Προσωπικού Χαρακτήρα δημοσίευσε έκθεση στην οποία επισημαίνεται ότι το σύστημα αξιολόγησης κινδύνου της Φορολογικής και Τελωνειακής Διοίκησης, το οποίο βασίζεται στην Τεχνητή Νοημοσύνη, δεν είχε τηρήσει τον Γενικό Κανό-

νισμό της ΕΕ για την Προστασία Δεδομένων (ΓΚΠΔ) και είχε βασιστεί σε μεροληπτικά δεδομένα σχετικά με την ιθαγένεια.

Σύμφωνα με την έκθεση της Εξεταστικής Επιτροπής της Βουλής του 2020, αργότερα αποκαλύφθηκε ότι μεταξύ 2012 και 2019, περίπου 25.000 έως 35.000 άτομα είχαν κριθεί ένοχα για δόλο ή βαριά αμέλεια, αλλά φάνηκε ότι στο 94% των περιπτώσεων αυτών, η απόδοση δόλου ή βαριάς αμέλειας ήταν αδικαιολόγητη, "επειδή ο λόγος δεν είχε καταγραφεί σωστά, επειδή δεν υπήρχαν σαφείς αποδείξεις για δόλο ή βαριά αμέλεια ή επειδή οι λόγοι για τον χαρακτηρισμό αυτό δεν είχαν δοθεί στους εν λόγω γονείς.

2.4.3 Συμπέρασμα της Επιτροπής της Βενετίας

Η Επιτροπή της Βενετίας του Συμβουλίου της Ευρώπης κατέγραψε το παραπάνω ιστορικό λεπτομερώς, εξέτασε σε βάθος όλες τις παραμέτρους και απεφάνθη στο πλαίσιο του ρόλου της. Στη σχετική έκθεση που συνέταξε¹⁷ διατύπωσε αναλυτικές προτάσεις, από τις οποίες ξεχωρίζουμε:

- Νομοθετική εξουσία:
 1. Η νομοθεσία για χρήση καινοτόμων τεχνολογιών να περιλαμβάνει διατάξεις που υπενθυμίζουν τις γενικές βασικές αρχές της χρηστής διοίκησης.
- Εκτελεστική εξουσία:
 1. Για τα άτομα, η πρόσβαση στις σχετικές πληροφορίες θα πρέπει να γίνει ευκολότερη, οι διαδικασίες καταγγελίας θα πρέπει να γίνουν απλές και ανεπίσημες και θα πρέπει να προσφέρεται βοήθεια για τον τρόπο υποβολής καταγγελίας.
 2. η εκτελεστική εξουσία, ο Επίτροπος Προστασίας Δεδομένων και άλλοι αρμόδιοι φορείς θα πρέπει να παρακολουθούν στενά τις εξελίξεις στην Τεχνητή Νοημοσύνη και οι νέες εξελίξεις θα πρέπει να λαμβάνονται υπόψη κατά το σχεδιασμό των μελλοντικών συστημάτων ΤΝ και κατά την αναθεώρηση των υφιστάμενων
 3. οι τομεακές πληροφορίες θα πρέπει να διαμοιράζονται ευρέως εντός της διοίκησης, ώστε να είναι δυνατή η σχετική συμβολή και από άλλους τομείς της διοίκησης.
- Δικαστική εξουσία
 1. Θα μπορούσαν να δημιουργηθούν δίαυλοι για τη δικαστική εξουσία ώστε να εφιστά την προσοχή των άλλων κλάδων στη νομοθεσία που δημιουργεί συστημικά προβλήματα στην πράξη.

2.5 Εθνικές στρατηγικές και εργαλεία αλγοριθμικής λογοδοσίας

Η Τεχνητή Νοημοσύνη ανήκει στα διεπιστημονικά ερευνητικά πεδία που ξεπερνούν τον ακαδημαϊκό χώρο. Στην παρούσα φάση της ΤΝ, το μεγαλύτερο μέρος των εξελίξεων και των εφαρμογών της αφορά μια κατηγορία αλγορίθμων που είναι γνωστή ως μηχανική μάθηση, συμπεριλαμβανομένης της βαθιάς μάθησης. Το σημερινό κύμα της ΤΝ είναι αλγόριθμοι αυτοεκμάθησης που

¹⁷CDL-AD(2021)031-e, Netherlands - Opinion on the Legal Protection of Citizens, adopted Venice Commission at its 128th Plenary Session, Στρασβούργο, 18 Οκτωβρίου 2021, [https://www.venice.coe.int/webforms/documents/?pdf=CDL-AD\(2021\)031-e](https://www.venice.coe.int/webforms/documents/?pdf=CDL-AD(2021)031-e)

βασίζονται σε τεράστιες ποσότητες δεδομένων, οδηγώντας σε καινοτομίες ΤΝ σε όλους τους τομείς. Οι διαφαινόμενες δυνατότητες έχουν δελεάσει τόσο τον δημόσιο όσο και τον ιδιωτικό τομέα με τις εφαρμογές τους να μετασχηματίζουν την κοινωνία¹⁸.

Καθώς ο αριθμός των εφαρμογών αυξάνεται, οι κυβερνήσεις και η δημόσια διοίκηση δέχονται μεγαλύτερη πίεση για την αντιμετώπιση των κινδύνων και των προκλήσεων. Από τη μία πλευρά, οι απειλές για το εργατικό δυναμικό, τις οργανωτικές δομές και την οικονομία δεν έχουν ακόμη κατανοηθεί πλήρως. Εκτός αυτού, η εφαρμογή της τεχνητής νοημοσύνης αναμφισβήτητα εγείρει ηθικά ζητήματα στον τομέα της ιδιωτικής ζωής, της προστασίας των δεδομένων και της διαφάνειας λόγω των εκτεταμένων συνεπειών της. Οι κυβερνήσεις, λοιπόν, έχουν διττό και αντικρουόμενο ρόλο στην υιοθέτηση της ΤΝ: από τη μία πλευρά ως ρυθμιστές της ΤΝ για την προστασία των πολιτών από πιθανές βλάβες και από την άλλη ως χρήστες της ΤΝ για την καλύτερη παροχή υπηρεσιών και την αποτελεσματική κατανομή των πόρων.

Πράγματι, οι εργασίες σχετικά με τον αντίκτυπο των εφαρμογών τεχνητής νοημοσύνης έχουν δείξει ότι ένα ευρύ φάσμα δημόσιων λειτουργιών μπορεί να ενισχυθεί με τη χρήση τους. Αυτό περιλαμβάνει, μεταξύ άλλων, την υγεία, την εκπαίδευση, τις μεταφορές και την ενέργεια, την κυβέρνηση και την άμυνα. Όλα αυτά ενεργοποιούνται από μια μεγάλη ποικιλία τεχνολογιών, μεθόδων και δεδομένων, παρουσιάζοντας διάφορες ευκαιρίες και προκλήσεις ιδιαίτερα στο πλαίσιο της λήψης αποφάσεων από τον δημόσιο τομέα. Η αυξανόμενη ανάγκη για αξιοποίηση των εφαρμογών της ΤΝ τα τελευταία χρόνια οδήγησε πολλές χώρες στη δημοσίευση εθνικών στρατηγικών για την τεχνητή νοημοσύνη.

Μια έρευνα αυτών των κειμένων παρουσιάζεται στη μελέτη “The global landscape of AI ethics guidelines” του 2019¹⁹. Η μελέτη επικεντρώθηκε στο να αποτυπώσει το παγκόσμιο τοπίο ως προς τη δημοσίευση τέτοιων κειμένων. Διαπίστωσε ότι ενώ τα τελευταία χρόνια, ιδιωτικές εταιρείες, ερευνητικά ιδρύματα και οργανισμοί του δημόσιου τομέα έχουν εκδώσει αρχές και κατευθυντήριες γραμμές για την δεοντολογική τεχνητή νοημοσύνη (ΤΝ), ωστόσο υπάρχει συζήτηση τόσο για το τι συνιστά “δεοντολογική τεχνητή νοημοσύνη” όσο και για το ποιες ηθικές απαιτήσεις, τεχνικά πρότυπα και βέλτιστες πρακτικές απαιτούνται για την υλοποίησή της. Τα αποτελέσματα αυτής της χαρτογράφησης αποκαλύπτουν μια παγκόσμια σύγκλιση γύρω από πέντε αρχές: διαφάνεια, δικαιοσύνη και δίκαιη μεταχείριση, έλλειψη δόλου, υπευθυνότητα και προστασία της ιδιωτικής ζωής (transparency, justice and fairness, non-maleficence, responsibility and privacy).

Εντούτοις είναι σημαντικό να αναφερθούν σε αυτό το σημείο οι ευρύτερες έννοιες της αξιοπιστίας και ανθρωποκεντρικής ΤΝ, που αναλύονται σε επόμενο κεφάλαιο καθώς αποτελούν κομβικές αξίες, ειδικά στα πρόσφατα αποκρυσταλλωμένα ρυθμιστικά πλαίσια, κατά τον έλεγχο συμμόρφωσης (compliance) και συνεπώς αξιολόγησης των συστημάτων ΤΝ, καθώς και να τονιστεί η έννοια της ανθρώπινης εποπτείας κατά τη χρήση τους.

Ωστόσο υπάρχουν ουσιαστικές αποκλίσεις στον τρόπο ερμηνείας αυτών των αρχών, τους λόγους για τους οποίους θεωρούνται σημαντικές, το θέμα, τον τομέα ή τους φορείς που αφορούν και τον τρόπο εφαρμογής τους. Επίσης τα ευρήματα της εν λόγω μελέτης υπογραμμίζουν τη σημασία του να ενσωματωθούν οι κατευθυντήριες γραμμές σε εφαρμόσιμα κείμενα στρατηγικής. Για αναλυτικότερη αποτύπωση των εθνικών πρωτοβουλιών ήδη από το 2019 σε όλο το γεωγραφικό τους εύρος μπορεί κανείς να ανατρέξει στο ίδιο το άρθρο όπου συγκεντρώνονται οι κατευθυντήριες γραμμές δεοντολογίας ανά χώρα και ανά ήπειρο.

Επίσης δεν πρέπει να διαφύγει της προσοχής μας το σχετικό με την περιγραφόμενη γεωγραφική αποτύπωση συμπέρασμα των συγγραφέων. Επισημαίνουν μια υποεκπροσώπηση γεωγραφικών περιοχών όπως η Αφρική, η Νότια και Κεντρική Αμερική και η Κεντρική Ασία που υποδηλώ-

¹⁸Omar Ballester, 2021, An Artificial Intelligence Definition and Classification Framework for Public Sector Applications, In DG.O2021: The 22nd Annual International Conference on Digital Government Research (DG.O'21). Association for Computing Machinery, New York, NY, USA, 67□75.<https://doi.org/10.1145/3463677.3463709>

¹⁹Jobin, Anna & Ienca, Marcello & Vayena, Effy, 2019, The global landscape of AI ethics guidelines, Nature Machine Intelligence, DOI 10.1038/s42256-019-0088-2, <http://dx.doi.org/10.1038/s42256-019-0088-2>

νει ότι παγκοσμίως δεν υπάρχει ισότιμη συμμετοχή του “Παγκόσμιου Νότου” στη συζήτηση για τη δεοντολογία της ΤΝ. Οι πιο οικονομικά ανεπτυγμένες χώρες διαμορφώνουν αυτή τη συζήτηση περισσότερο από άλλες, γεγονός που εγείρει ανησυχίες σχετικά με την παραμέληση της τοπικών συνθηκών, του πολιτισμικού πλουραλισμού και των απαιτήσεων για παγκόσμια δικαιοσύνη.

Αναλυτική και επικαιροποιημένη παρουσίαση αυτής της γεωγραφικής κατανομής κάνει το εργαλείο οπτικοποίησης δεδομένων σχετικά με τις πρωτοβουλίες για την Τεχνητή Νοημοσύνη, που έχει εκπονήσει το Συμβούλιο της Ευρώπης (με δεδομένα μέχρι το 2022). Η συγκεκριμένη οπτικοποίηση δεδομένων²⁰ βασίζεται σε συλλογή δεδομένων κατά την περίοδο 2016-2022 και συγκεντρώνει περισσότερα από 450 έγγραφα που σχετίζονται με την τεχνητή νοημοσύνη, τα οποία προέρχονται από εθνικές αρχές, τον ιδιωτικό τομέα, διεθνείς οργανισμούς ή πρωτοβουλίες άλλων ενδιαφερομένων φορέων. Το συμπέρασμα που αναφέρει το εν λόγω εργαλείο οπτικοποίησης είναι ότι εξετάζοντας την οπτικοποίηση που αντιπροσωπεύει τους διεθνείς οργανισμούς, γίνεται σαφές ότι το Συμβούλιο της Ευρώπης, μαζί με την Ευρωπαϊκή Ένωση, τον ΟΟΣΑ και την UNESCO, είναι από τους πιο παραγωγικούς οργανισμούς στον τομέα.

Ενδιαφέρον παρουσιάζει και η ενδιάμεση των παραπάνω, από χρονολογικής άποψης, μελέτη (2021) των Ada Lovelace Institute²¹, AI Now Institute²² και Open Government Partnership²³ με τίτλο “Algorithmic accountability for the public sector”²⁴ η οποία επικεντρώνεται στη χρήση των αλγορίθμων για την αυτοματοποίηση της λήψης αποφάσεων στις δημόσιες υπηρεσίες, στη βλάβη που μπορεί να προκαλέσουν και στην πιθανή αδιαφάνεια που περιβάλλει την εφαρμογή τους. Πολλά αλγοριθμικά συστήματα στερούνται πολιτικής λογοδοσίας καθώς αναπαράγουν ή και ενισχύουν υπάρχουσες προκαταλήψεις εναντίον ομάδων ανθρώπων που έχουν ιστορικά υποστεί διακρίσεις, είτε διευκολύνουν νέες μορφές παραβίασης της ιδιωτικής ζωής. Η εν λόγω έρευνα αξιολογεί αυτό το πρώτο κύμα λογοδοσίας για τους αλγόριθμους κατά το οποίο οι υπεύθυνοι χάραξης πολιτικής στράφηκαν σε ρυθμιστικά εργαλεία για “αλγοριθμική λογοδοσία” όπως η εκτίμηση αλγοριθμικού αντικτύπου, τα μητρώα αλγορίθμων και ελεγκτικοί μηχανισμοί για την εφαρμογή αλγορίθμων (algorithm auditing).

2.6 Εκτίμηση Αλγοριθμικού Αντικτύπου της Ομοσπονδιακής Κυβέρνησης του Καναδά

Ο ιδιαίτερος αυτός ρόλος καθιστά ακόμη σπουδαιότερη τη σημασία της ορθής και οριοθετημένης χρήσης της τεχνητής νοημοσύνης όταν αυτή γίνεται εφαρμοστικό εργαλείο αυτοματοποιημένης λήψης απόφασης κατά την άσκηση της δημόσιας διοίκησης²⁵. Για τον λόγο αυτό εξετάζεται το πώς στις περιπτώσεις αυτές μεγάλο βάρος δίνεται σε διάφορες απαιτήσεις διαφάνειας, όπως η ολοκλήρωση και δημοσίευση μιας Εκτίμησης Αλγοριθμικών Επιπτώσεων, η παροχή ειδοποιη-

²⁰Visualization of AI initiatives, Council of Europe, 2016-2022, <https://www.coe.int/en/web/artificial-intelligence/national-initiatives>

²¹The Ada Lovelace Institute was established by the Nuffield Foundation in early 2018, in collaboration with the Alan Turing Institute, the Royal Society, the British Academy, the Royal Statistical Society, the Wellcome Trust, Luminate, techUK and the Nuffield Council on Bioethics, <https://www.adalovelaceinstitute.org/about/>

²²Founded in 2017, the AI Now Institute produces diagnosis and policy research on artificial intelligence. AI Now’s leadership were invited to advise the US Federal Trade Commission on artificial intelligence in 2021, <https://ainowinstitute.org/about>

²³The Open Government Partnership (OGP) includes 75 countries and 104 local governments, representing more than two billion people, and thousands of civil society organizations and promotes transparent, participatory, inclusive and accountable governance, <https://www.opengovpartnership.org/about/>

²⁴Ada Lovelace Institute, AI Now Institute and Open Government Partnership. 2021. Algorithmic Accountability for the Public Sector. <https://www.opengovpartnership.org/documents/algorithmic-accountability-public-sector/>

²⁵Allison Merchant, Three Recommendations for More Inclusive and Equitable AI in the Public Sector, 24 Ιανουαρίου 2023, Open Government Partnership, <https://www.opengovpartnership.org/stories/three-recommendations-for-more-inclusive-and-equitable-ai-in-the-public-sector>

σεων και εξηγήσεων στους πολίτες πριν και μετά τη λήψη αποφάσεων, η εξασφάλιση πρόσβασης στα στοιχεία ενός συστήματος (π.χ. για ελέγχους ή έρευνες), η δημοσιοποίηση του πηγαιού κώδικα, όπου χρειάζεται, και η τεκμηρίωση των αυτοματοποιημένων αποφάσεων που προκύπτουν από μη ντετερμινιστικούς αλγόριθμους²⁶.

Διατρέχοντας τις οργανωμένες πρωτοβουλίες για τη ρύθμιση της ΤΝ που σημειώθηκαν στην παγκόσμια κοινότητα, με ποικίλους βαθμούς υποχρέωσης συμμόρφωσης, ως απάντηση στις προκλήσεις που θέτει η ραγδαία ανάπτυξή της, επισημάνουμε ότι κυβερνητικοί και διακυβερνητικοί φορείς επιδεικνύουν αυξημένο ενδιαφέρον για τη ρύθμιση της ΤΝ αλλά και για τη χρήση της. Εκπονούν και δημοσιεύουν στρατηγικές διακυβέρνησης των συστημάτων ΤΝ και των συστατικών της και την αξιολογούν σε τομείς άσκησης κυβερνητικής πολιτικής. Για το λόγο αυτό συγχρόνως προβλέπουν, όπως ο ελληνικός νόμος 4961/22, άρθρο 5²⁷, ο οποίος θα παρουσιαστεί παρακάτω αναλυτικότερα, ή εφαρμόζουν ήδη, όπως το Καναδικό κράτος, την Εκτίμηση Αλγοριθμικού Αντικτύπου (Algorithmic Impact Assessment, εφεξής ΑΙΑ) ως εργαλείο μετριάσμού των πιθανών κινδύνων.

Περιεκτική παρουσίαση του εργαλείου για την Εκτίμηση Αλγοριθμικού Αντικτύπου της Γραμματείας του Υπουργείου Οικονομικών του Καναδά έγινε από το κυβερνητικό στέλεχος Omar Bitar στο πλαίσιο της 3ης Ολομέλειας της Επιτροπής για την Τεχνητή Νοημοσύνη του Συμβουλίου της Ευρώπης με θέμα “Διαχείριση των κινδύνων της ΤΝ: εθνικές και διεθνείς πρωτοβουλίες” στις 11 Ιανουαρίου 2023 στο Στρασβούργο²⁸, την οποία αντλήσαμε και παραθέτουμε παρακάτω:

“Το 2019, η ομοσπονδιακή κυβέρνηση του Καναδά έγινε μία από τις πρώτες κυβερνήσεις παγκοσμίως που εξέδωσε Οδηγία η οποία ρυθμίζει τη χρήση αυτοματοποιημένων συστημάτων λήψης αποφάσεων²⁹, συμπεριλαμβανομένης της τεχνητής νοημοσύνης, στον δημόσιο τομέα. Η οδηγία του Συμβουλίου του Υπουργείου Οικονομικών για την αυτοματοποιημένη λήψη αποφάσεων είναι ένα υποχρεωτικό μέσο πολιτικής που επιδιώκει να διασφαλίσει τη διαφάνεια, τη λογοδοσία και τη δίκαιη μεταχείριση στην αυτοματοποιημένη λήψη αποφάσεων στον ομοσπονδιακό δημόσιο τομέα.

Η οδηγία υποστηρίζεται από την Εκτίμηση Αλγοριθμικού Αντικτύπου³⁰ (Algorithmic Impact Assessment, ΑΙΑ) ένα διαδικτυακό ερωτηματολόγιο που επιτρέπει στους χρήστες (κυβερνητικές υπηρεσίες) να κατανοήσουν, να αξιολογήσουν και να διαχειριστούν τους κινδύνους ενός έργου αυτοματοποίησης και να καθορίσουν τα εφαρμοστέα μέτρα βάσει της οδηγίας. Η ΑΙΑ αναπτύχθηκε σε ανοικτό περιβάλλον και είναι διαθέσιμη για κοινή χρήση και επαναχρησιμοποίηση με ανοικτή άδεια χρήσης. Η οδηγία θεσπίζει μια σειρά μέτρων διαφάνειας, διασφάλισης ποιότητας και προσφυγής για έργα που αφορούν ένα αυτοματοποιημένο σύστημα που χρησιμοποιείται για τη λήψη ή την υποστήριξη μιας διοικητικής απόφασης, μιας απόφασης που επηρεάζει ατομικά δικαιώματα, συμφέροντα ή προνόμια.

Η οδηγία απαιτεί από τις υπηρεσίες που υπάγονται σε αυτήν να συμπληρώνουν και να δημο-

²⁶A. Feder Cooper, Jonathan Frankle, and Christopher De Sa. 2022. Non-Determinism and the Lawlessness of Machine Learning Code. In Proceedings of the 2022 Symposium on Computer Science and Law (CSLAW '22). Association for Computing Machinery, New York, NY, USA, 1–8, <https://dl.acm.org/doi/abs/10.1145/3511265.3550446>

²⁷N. 4961/22, άρθ. 5: Θεσπίζεται σχετική υποχρέωση εκπόνησης αλγοριθμικής εκτίμησης αντικτύπου, λαμβανομένων υπόψιν, ιδίως, του επιδιωκόμενου σκοπού, των δυνατοτήτων, τεχνικών χαρακτηριστικών και παραμέτρων λειτουργίας του συστήματος, της φύσης των αποφάσεων, των κατηγοριών δεδομένων, των ενδεχόμενων κινδύνων για τα δικαιώματα, τις ελευθερίες και τα έννομα συμφέροντα των προσώπων που αφορά ή επηρεάζει η λήψη της εκάστοτε απόφασης, καθώς και του προσδοκώμενου οφέλους σε συνάρτηση με τους κινδύνους και τις επιπτώσεις σε ομάδες ή κατηγορίες του πληθυσμού.

²⁸Webinar: Managing AI risks: national and International initiatives, 11 Ιανουαρίου 2023, Στρασβούργο, <https://www.coe.int/en/web/artificial-intelligence/-/webinar-managing-ai-risks-national-and-international-initiatives-5>

²⁹Directive on Automated Decision-Making, <https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592>

³⁰Εργαλείο Αλγοριθμικής Εκτίμησης Επιπτώσεων του Καναδά, <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html>

σιεύουν μια ΑΙΑ. Η ΑΙΑ περιέχει ερωτήσεις που υποστηρίζουν τον εντοπισμό, την αξιολόγηση και τον μετριασμό των κινδύνων που προκύπτουν από τη χρήση ΤΝ και άλλων αυτοματοποιημένων συστημάτων στη λήψη αποφάσεων. Για παράδειγμα, προτρέπει τους χρήστες να εξετάσουν τις πιθανές επιπτώσεις ενός έργου αυτοματοποίησης στα ατομικά δικαιώματα, την υγεία και την ευημερία, τα οικονομικά συμφέροντα και τη βιωσιμότητα των οικοσυστημάτων. Ρωτά επίσης τους χρήστες αν έχουν νομική εξουσιοδότηση για την πλήρη ή μερική αυτοματοποίηση μιας απόφασης και επισημαίνει τους πιθανούς κινδύνους για την προστασία της ιδιωτικής ζωής και τις επιπτώσεις σε ευάλωτα άτομα. Περιλαμβάνει επίσης ερωτήσεις σχετικά με μέτρα μετριασμού των κινδύνων, όπως δοκιμές μεροληψίας, ελέγχους ποιότητας δεδομένων, διαβουλεύσεις με εξωτερικούς και εσωτερικούς ενδιαφερόμενους φορείς, διαδρομές ελέγχου και προστασία της ιδιωτικής ζωής για προσωπικές πληροφορίες.

Οι απαντήσεις στις ερωτήσεις της ΑΙΑ για τους κινδύνους και τον μετριασμό των επιπτώσεων παράγουν μια συνολική βαθμολογία που αντιστοιχεί από το 1 έως το 4, δηλαδή από μικρή ή καμία επίπτωση έως πολύ υψηλή επίπτωση. Το επίπεδο επιπτώσεων ενός έργου καθορίζει τα εφαρμοστέα μέτρα στο πλαίσιο ορισμένων απαιτήσεων της οδηγίας, τα οποία διαφέρουν ανάλογα με το επίπεδο επιπτώσεων.

Ορισμένες ΑΙΑ έχουν συγκεντρώσει την προσοχή των μέσων ενημέρωσης και έχουν προωθήσει μια δημόσια συζήτηση σχετικά με τα πλεονεκτήματα και τα όρια της χρήσης ΤΝ στην ομοσπονδιακή κυβέρνηση. Αυτή η διαφάνεια, με τη σειρά της, ενισχύει τη λογοδοσία για τις αυτοματοποιημένες αποφάσεις σε υπηρεσίες και άλλες κυβερνητικές δραστηριότητες. Η εποπτεία της οδηγίας και της ΑΙΑ έχει οδηγήσει στον εντοπισμό αρκετών προκλήσεων που σχεδιάζεται να αντιμετωπιστούν. Σε αυτές περιλαμβάνεται η ποιότητα των ΑΙΑ, η οποία διαφέρει σημαντικά από έργο σε έργο. Μια άλλη πρόκληση που αξίζει να σημειωθεί έχει να κάνει με την ανάγκη να διασφαλιστεί ότι οι ΑΙΑ επανεξετάζονται τακτικά καθ' όλη τη διάρκεια του κύκλου ζωής ενός έργου αυτοματισμού. Μια τρίτη πρόκληση είναι ότι ενώ η ΑΙΑ επιτρέπει στους πολίτες ή τις επιχειρήσεις, που υπόκεινται σε αυτοματοποιημένες αποφάσεις, να κατανοήσουν καλύτερα πώς η κυβέρνηση χρησιμοποιεί τα συστήματα ΤΝ, δεν παρέχουν επαρκείς πληροφορίες που να αποδεικνύουν πώς αντιμετωπίζουν τους κινδύνους που έχουν εντοπιστεί. Αυτό αποτελεί μέρος μιας ευρύτερης ανάγκης για μεγαλύτερη διαφάνεια όσον αφορά τη συμμόρφωση με τα μέτρα που απαιτούνται βάσει της οδηγίας.”

Ακολούθως παρατίθενται κάποιες αντιπροσωπευτικές ερωτήσεις της ΑΙΑ, από την ιστοσελίδα ανοικτών δεδομένων της Κυβέρνησης του Καναδά³¹:

1. ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΣ ΠΑΡΑΓΟΝΤΑΣ / ΘΕΤΙΚΟΣ ΑΝΤΙΚΤΥΠΟΣ

Τι παρακινεί την ομάδα σας να εισαγάγει τη μηχανική μάθηση σε αυτή τη διαδικασία λήψης αποφάσεων; (Σημειώστε όλα όσα ισχύουν)

Υπάρχουσα συσσώρευση εργασιών ή υποθέσεων.

Βελτίωση της συνολικής ποιότητας των αποφάσεων.

Χαμηλότερο κόστος συναλλαγής ενός υφιστάμενου προγράμματος.

Το σύστημα εκτελεί εργασίες που ο άνθρωπος δεν μπορεί να ολοκληρώσει σε εύλογο χρονικό διάστημα.

Χρήση καινοτόμων προσεγγίσεων.

Άλλο (διευκρινίστε)

2. ΣΧΕΤΙΚΑ ΜΕ ΤΟ ΣΥΣΤΗΜΑ

Ελέγξτε ποιες από τις παρακάτω δυνατότητες ισχύουν για το σύστημά σας.

Αναγνώριση εικόνων και αντικειμένων: Ανάλυση πολύ μεγάλων συνόλων δεδομένων για την αυτοματοποίηση της αναγνώρισης, της ταξινόμησης και του πλαισίου που σχετίζεται

³¹<https://open.canada.ca/data/en/info/5423054a-093c-4239-85be-fa0b36ae0b2e>

με μια εικόνα ή ένα αντικείμενο.

Ανάλυση κειμένου και ομιλίας: Ανάλυση μεγάλων συνόλων δεδομένων για την αναγνώριση, επεξεργασία και επισήμανση κειμένου, ομιλίας, φωνής και την υποβολή συστάσεων με βάση την επισήμανση.

Αξιολόγηση κινδύνου: Ανάλυση πολύ μεγάλων συνόλων δεδομένων για τον εντοπισμό μοτίβων και τη σύσταση τρόπων δράσης και, σε ορισμένες περιπτώσεις, την ενεργοποίηση συγκεκριμένων ενεργειών.

Δημιουργία περιεχομένου: Ανάλυση μεγάλων συνόλων δεδομένων για την κατηγοριοποίηση, επεξεργασία, ταξινόμηση, εξατομίκευση και παροχή συγκεκριμένου περιεχομένου για συγκεκριμένα πλαίσια.

Βελτιστοποίηση διαδικασιών αυτοματοποίηση ροής εργασιών: Ανάλυση μεγάλων συνόλων δεδομένων για τον εντοπισμό ανωμαλιών, ομαδοποίηση μοτίβων, πρόβλεψη αποτελεσμάτων ή τρόπων βελτιστοποίησης και αυτοματοποίηση συγκεκριμένων ροών εργασίας.

3. ΣΧΕΤΙΚΑ ΜΕ ΤΗΝ ΑΠΟΦΑΣΗ

Αφορά η απόφαση κάποια από τις παρακάτω κατηγορίες (*επιλέξτε όλες τις κατηγορίες που ισχύουν):

Υπηρεσίες που σχετίζονται με την υγεία.

Οικονομικά συμφέροντα (επιχορηγήσεις και εισφορές, φορολογικές παροχές, είσπραξη οφειλών).

Κοινωνική βοήθεια (πχ επιδόματα αναπηρίας).

Πρόσβαση και κινητικότητα (έλεγχος ασφαλείας, διασυννοριακές μετακινήσεις).

Αδειοδότηση και έκδοση αδειών.

4. ΕΚΤΙΜΗΣΗ ΕΠΙΠΤΩΣΕΩΝ

Περιλαμβάνει η σύσταση ή η απόφαση που λαμβάνει το σύστημα στοιχεία διακριτικής ευχέρειας;

Περιγράψτε τι αποτελεί διακριτική ευχέρεια στην απόφαση.

Χρησιμοποιείται το σύστημα από διαφορετικό τμήμα του οργανισμού από εκείνο που το ανέπτυξε;

Είναι αναστρέψιμες οι επιπτώσεις που προκύπτουν από την απόφαση;

Πόσο καιρό θα διαρκέσουν οι επιπτώσεις της απόφασης; Παρακαλείστε να περιγράψετε τις επιπτώσεις που προκύπτουν από την απόφαση (σύμφωνα με την επιλογή ανωτέρω).

Οι επιπτώσεις που θα έχει η απόφαση δικαιώματα ή τις ελευθερίες ατόμων ή των κοινοτήτων, θα είναι πιθανότατα... Παρακαλείστε να περιγράψετε γιατί οι επιπτώσεις που προκύπτουν από την απόφαση είναι ... (σύμφωνα με την επιλογή ανωτέρω).

Οι επιπτώσεις που θα έχει η απόφαση στην υγεία και την ευημερία των ατόμων ή των κοινοτήτων, θα είναι πιθανότατα... Παρακαλείστε να περιγράψετε γιατί οι επιπτώσεις που προκύπτουν από την απόφαση είναι ... (σύμφωνα με την επιλογή ανωτέρω)

Οι επιπτώσεις που θα έχει η απόφαση στα οικονομικά συμφέροντα ατόμων ή των κοινοτήτων, θα είναι πιθανότατα... Παρακαλείστε να περιγράψετε γιατί οι επιπτώσεις που προκύπτουν από την απόφαση είναι ... (σύμφωνα με την επιλογή ανωτέρω)

Οι επιπτώσεις που θα έχει η απόφαση διαρκή βιωσιμότητα ενός περιβαλλοντικού οικοσυστήματος, θα είναι πιθανότατα... Παρακαλείστε να περιγράψετε γιατί οι επιπτώσεις που προκύπτουν από την απόφαση είναι ... (σύμφωνα με την επιλογή ανωτέρω)

5. ΓΙΑ ΤΑ ΔΕΔΟΜΕΝΑ

A. Πηγή δεδομένων

Θα χρησιμοποιεί το αυτοματοποιημένο σύστημα λήψης αποφάσεων προσωπικές πληροφορίες ως δεδομένα εισόδου;

Ποια είναι η υψηλότερη διαβάθμιση ασφαλείας των δεδομένων εισόδου που χρησιμοποιούνται από το σύστημα; (Επιλέξτε ένα)

Ποιος ελέγχει τα δεδομένα;

Θα χρησιμοποιηθούν δεδομένα του συστήματος από πολλές διαφορετικές πηγές;

Το σύστημα θα απαιτεί δεδομένα εισόδου από συσκευή συνδεδεμένη στο Διαδίκτυο ή στην τηλεφωνία; (π.χ. αισθητήρας)

Θα διασυνδεθεί το σύστημα με άλλα συστήματα ΤΠ;

Ποιος συνέλεξε τα δεδομένα που χρησιμοποιήθηκαν για την εκπαίδευση του συστήματος;

Ποιος συνέλεξε τα δεδομένα εισόδου που χρησιμοποιούνται από το σύστημα;

B. Τύπος δεδομένων

Θα απαιτεί το σύστημα την ανάλυση μη δομημένων δεδομένων για την απόδοση μιας σύστασης ή μιας απόφασης;

Ποιους τύπους μη δομημένων δεδομένων (Σημειώστε όλα όσα ισχύουν);

6. ΜΕΤΡΑ ΜΕΙΩΣΗΣ ΚΑΙ ΜΕΤΡΙΑΣΜΟΥ ΚΙΝΔΥΝΟΥ

A. Ποιότητα δεδομένων

Έχετε τεκμηριώσει τις διαδικασίες που εφαρμόζετε για τον έλεγχο των συνόλων δεδομένων έναντι μεροληψιών και άλλων απροσδόκητων αποτελεσμάτων; Αυτό θα μπορούσε να περιλαμβάνει εμπειρία στην εφαρμογή πλαισίων, μεθόδων, κατευθυντήριων γραμμών ή άλλων εργαλείων αξιολόγησης.

Υπάρχει επί του παρόντος διαδικασία για την τεκμηρίωση του τρόπου με τον οποίο επιλύθηκαν ζητήματα ποιότητας δεδομένων κατά τη διαδικασία σχεδιασμού;

Έχετε προβεί σε ανάλυση των δεδομένων με βάση το φύλο Plus;

Έχετε αναθέσει στο ίδρυμά σας την ευθύνη για το σχεδιασμό, την ανάπτυξη, τη συντήρηση και τη βελτίωση του συστήματος;

Έχετε τεκμηριώσει διαδικασίες για τη διαχείριση του κινδύνου να χρησιμοποιηθούν ξεπερασμένα ή αναξιόπιστα δεδομένα για τη λήψη αυτοματοποιημένης απόφασης;

Θα αναρτηθούν τα δεδομένα στο Open Gov;

B. Διαδικαστική δικαιοσύνη (Εποπτεία και παρέμβαση)

Παρέχει το σύστημα διαδρομή ελέγχου που καταγράφει τις συστάσεις ή τις αποφάσεις;

Διατηρείτε αρχείο καταγραφής με όλες τις αλλαγές που γίνονται στο μοντέλο και στο σύστημα;

Περιγράφει η διαδρομή ελέγχου με σαφήνεια όλα τα σημεία απόφασης που λαμβάνονται από το σύστημα;

Δείχνει η διαδρομή ελέγχου ποιος είναι ο εξουσιοδοτημένος λήπτης αποφάσεων;

Το σύστημα αιτιολογεί τις αποφάσεις ή τις συστάσεις του όταν απαιτείται;

Υπάρχει διαδικασία χορήγησης, παρακολούθησης και ανάκλησης της άδειας πρόσβασης στο σύστημα;

Υπάρχει μηχανισμός για τη συλλογή ανατροφοδότησης από τους χρήστες του συστήματος;

Προβλέπεται ή θεσπίζεται διαδικασία προσφυγής για τους πελάτες που επιθυμούν να αμφισβητήσουν την απόφαση;

Επιτρέπει το σύστημα την παράκαμψη των αποφάσεων του συστήματος από τον άνθρωπο;

Υπάρχει μια διαδικασία για την καταγραφή των περιπτώσεων κατά τις οποίες πραγματοποιήθηκαν παρακάμψεις;

Περιλαμβάνει η διαδρομή ελέγχου διαδικασίες ελέγχου αλλαγών για την καταγραφή τροποποιήσεων στη λειτουργία ή την απόδοση του συστήματος;

2.7 Συμπέρασμα

Η αυξανόμενη ανάγκη για αξιοποίηση των εφαρμογών της ΤΝ τα τελευταία χρόνια οδήγησε πολλές χώρες στη δημοσίευση εθνικών στρατηγικών για την τεχνητή νοημοσύνη. Πλέον το παγκόσμιο τοπίο βρίθει αρχών και κατευθυντηρίων γραμμών για την ηθική τεχνητή νοημοσύνη (ΤΝ), που προβληματίζουν για το τι συνιστά “ηθική τεχνητή νοημοσύνη” και τι απαιτείται πρακτικά για την υλοποίησή της. Ωστόσο κατατείνουν σε μια παγκόσμια σύγκλιση γύρω από ηθικές αρχές όπως διαφάνεια, δικαιοσύνη και δίκαιη μεταχείριση, έλλειψη δόλου, υπευθυνότητα και προστασία της ιδιωτικής ζωής αλλά και σε έννοιες όπως αξιόπιστη και ανθρωποκεντρική ΤΝ που χαρακτηρίζεται από την ανθρώπινη εποπτεία. Παράλληλα παρατηρείται μια υποεκπροσώπηση του “Παγκόσμιου Νότου” σε αυτές τις ρυθμιστικές πρωτοβουλίες.

Όσον αφορά το κρίσιμο ζήτημα της λογοδοσίας των αλγορίθμων, παρατηρείται μια στροφή σε ρυθμιστικά εργαλεία όπως Εκτίμηση Αλγοριθμικού Αντικτύπου, Μητρώα Αλγορίθμων και ελεγκτικούς μηχανισμούς για την εφαρμογή αλγορίθμων (algorithm auditing). Πέραν της ρύθμισης αναδεικνύεται ο σημαντικός ρόλος της κυβέρνησης ως χρήστη της ΤΝ που εγείρει ποικίλα ζητήματα κυρίως στο πεδίο της αυτοματοποιημένης λήψης απόφασης, επί της οποίας αναλύθηκε το χαρακτηριστικό αντιπαράδειγμα του προνοιακού επιδόματος της Ολλανδίας. Ιδιαίτερο ενδιαφέρον παρουσιάζει η δυνατότητα προαγωγής της κοινοβουλευτικής λειτουργίας με την υποστήριξη εργαλείων ΤΝ.

Παράλληλα ελλοχεύει ο κίνδυνος όλη αυτή η υπερδραστηριότητα στη διατύπωση ηθικών αρχών να οδηγήσει σε μια “αγορά αρχών και αξιών” όπου ιδιωτικοί και δημόσιοι φορείς μπορούν να “αγοράσουν” το είδος της ηθικής που τους ταιριάζει περισσότερο για να δικαιολογήσουν τις τακτικές τους.

Τέλος το παράδοξο στη ρύθμιση της παραπληροφόρησης δείχνει να είναι ότι όσο περισσότερο ρυθμίζεται στη Δύση, τόσο παραμελείται και επιδεινώνεται σε παγκόσμιο επίπεδο. Αυτό μπορεί να αποβεί επικίνδυνο ειδικά κατά το 2024, μια χρονιά με εκλογές σε 50 χώρες του κόσμου καθώς και στην Ευρωπαϊκή Ένωση.

Table 1: AI-based pilot projects in public services

Project	Country	Sector	Description	Data	Technology type	Knowledge Source
ADELE	Switzerland	Statistics Office / Land use	Use of artificial intelligence (AI) technologies to eventually automate (even partially) the visual interpretation of aerial images in order to detect and classify changes in land use	Internal aerial images + classified metadata (auxiliary).	DEEP L - Image recognition	External Mandate. Collaboration with University
NOGAuto	Switzerland	Statistics Office / Economic Activity	Automation of the coding of the economic activity of enterprises using Machine Learning methods	Internal data already available	ML Supervised + NLP	Internal resources
ML_SoSi	Switzerland	Statistics Office / Social Security	Clustering of typical prospective trajectories patterns concerning the receipt of benefits in the social security system as well as employment and the estimation of cluster affiliation	Internal data already available, from different sources	ML Unsupervised	Internal Resources
Controles Plausibilite	Switzerland	Statistics Office	To extend and speed up data validation in the FSO by means of machine learning algorithms and at the same time to improve data quality. Task automation of a tedious manual/rule-based labor	Internal data already available, from different sources	ML Supervised	Internal Resources
BESTA	Switzerland	Statistics Office / Social Security	To test new methods for estimating Job Statistics with better predictors	Internal data already available, from different sources	Regression	Internal Resources
Chatbot St. Gallen	Switzerland	City Administration / Tourism	A local administration online assistant.	Pre-trained chatbot with external data	NLP - Deep - API	External Mandate. Byerley
Free Flow	Italy	Road Transportation	Automatic toll system: elimination of tolls and image based detection of travel, vehicle category and cost	Image-recognition software externally pre-trained	Computer vision	
Satikas	Estonia	Economic affairs	Automatic detection of grasslands state (mowing or not): requirement for subsidies.	Satellite data with reference historical data	DEEP L - Image recognition	Multi-stakeholder partnership. University of Tartu, State and Private
Copernicus - related projects. Multiple applications	EU	Agriculture / Economic affairs / CAP compliance / Farming Monitoring	Automatic detection of land usage for agriculture. Automatisation of CAP-compliance for subsidies and supervision.	Satellite data with reference historical data	DEEP L - Image recognition	Multi-stakeholder

Εικόνα 2.1. (1/3) Πιλοτικά έργα Τεχνητής Νοημοσύνης για δημόσιες υπηρεσίες, Omar Ballester, 2021, An Artificial Intelligence Definition and Classification Framework for Public Sector Applications, ACM

Project	Country	Sector	Description	Data	Technology type	Knowledge Source
Text from image	Belgium	Social Security / Health	Tool to automatically detect lot numbers from medication boxes. Task automation of a manually anotated job	External data.	DEEP L - Computer vision	Private vendor
Day-care prediction	Belgium	Social Security / Health	Recommender system for predicting day-care services requiring inspection	Own data	ML - regression	Internal resources and other agencies
Robotic Process Automation	Trelleborg, Copenhagen,	Local GVT / Social Security	Automatic Decision-Making regarding social welfare applications. Compliance, calculations, and determination	Own process with consultant	Task Automation -	External Vendor UIPatch
ChatBot UNA	Latvia	Economic affairs / Registry (enterprise)	FAQ answering. Responding to recurring questions for simple, standard tasks	External + internal	Chatbot / Recommender system / NLP	External Vendor - Tilde
Tengai	Sweden	Local GVT	"unbiased" Recruitment process automation: performing automatic robot interviews and delivering unbiased candidate assessments.	External data	Speech recognition, Cognitive robotics, Deep + Computer vision	External Vendor - Tengai
SyRi	Netherlands	Law Enforcement / Economic affairs / Social Security	Social Security fraud detection. Cancelled in 2020 for non-compliance with ECHR	Internal data: tax, health insurance, residence, education, etc. . . fraud in welfare benefits	Predictive / Fraud detection	Central government, used by local administration
Unemployed profile	Poland	Social Security / Economic affairs	Automated profiling system for unemployment benefits. Classification into one of three categories that determines programs that recipient is eligible for. Decision support (recommender). Dismantled as of 2019	Internal data	Rule based system? Classifier? Unknown	Mintory of labour and social policy
VeriPol	Spain	Law Enforcement	Fraud detection in police reports, particularly for low-level crimes.	Internal Police Data (reports)	Binary Classifier combination of NLP + ML	External academic partnerships.
Roxanne	Switzerland / EU	Law Enforcement	Decision-support system for Law Enforcement Agencies accross europe. Goal to enhance investigation capabilities especially for large criminal cases.	Some Internal + External training sources (openly available)	Network analysis + Speech recognition + visual analysis	University-led partnership with industry and LEA
Copkit	Spain / EU	Law Enforcement	Intelligent early warning / Early action system for tackling the use of new ICT technologies in organised crime	Combination of Internal (intel and criminal) and external	NLP - ML	Public company led consortium (academia + LEA + industry)

Εικόνα 2.2. (2/3) Πιλοτικά έργα Τεχνητής Νοημοσύνης για δημόσιες υπηρεσίες, Omar Ballester, 2021, An Artificial Intelligence Definition and Classification Framework for Public Sector Applications, ACM

Project	Country	Sector	Description	Data	Technology type	Knowledge Source
Deep AR	EU	Law Enforcement	Enhance police operations with the use of Augmented Reality tools combined with ML predictive	?	ML + Vision	Academia + LEAs + Private
INSPECTr	EU	Law Enforcement	Digital Forensics platform for knowledge discovery support. shared intelligent platform and a novel process for gathering, analysing, prioritising and presenting key data to help in the prediction, detection and management of crime. UNIL is partner	Web data and dark web	Cognitive ML	Academia + LEAs + Private
SBB	Switzerland	Transportation / Mobility	Multiple projects: Automation, Predictive Algorithms, Security, IoT and Smart mobility, Crowd Management, Machine Perception, Maintenance, Scheduling (flatland)	Internal data, sensor data. Enterprise-like management.	Reinforcement Learning, Computer Vision, ML, Expert Systems, NLP, Image Recognition	Internal?
Language Programme Railnet Europe MM4SIGHT	EU Denmark	Transportation / Mobility Migration / Social Security / International Cooperation	Provide Translation support for international rail interoperability at the highest security standards Decision support system to find connection between drivers of migration and migration itself. Predictive modelling of migration flows	External data Past Migration Data	Translation NLP Speech ML + Statistical models	Industry vendor + Gov Agency
PRECOBS	Switzerland + Germany (Bayern, Zurich, Aargau)	Law Enforcement	Forecasting of near-repeat-crimes for burglary prevention with human-in-the-loop. Aargau makes prediction public.	Police Report Data (from previous cases)	ML + Expert System	Industry vendor + Gov Agency

Εικόνα 2.3. (3/3) Πιλοτικά έργα Τεχνητής Νοημοσύνης για δημόσιες υπηρεσίες, Omar Ballester, 2021, An Artificial Intelligence Definition and Classification Framework for Public Sector Applications, ACM

Κεφάλαιο 3

Γιατί και πώς ρυθμίζεται θεσμικά η ΤΝ

- Στο παρόν κεφάλαιο θα αναλυθούν οι τρόποι ρύθμισης για δεοντολογική και σύννομη χρήση των συστημάτων τεχνητής νοημοσύνης από τον δημόσιο τομέα, δηλαδή από διεθνείς οργανισμούς, την Ευρωπαϊκή Ένωση, τις κυβερνήσεις αλλά και τη δημόσια σφαίρα γενικότερα, όπως μη κυβερνητικούς οργανισμούς, την κοινωνία των πολιτών και εκπροσώπους της ακαδημαϊκής και ερευνητικής κοινότητας.
- Θα εξετάσουμε πώς η αναζήτηση κατάλληλου μοντέλου διακυβέρνησης διαμορφώνεται υπό το πρίσμα της αυξανόμενης επίγνωσης ότι η χρήση της ΤΝ εκφεύγει των παραδοσιακών ορίων δικαιοδοσιών και συνόρων, δημόσιας και ιδιωτικής σφαίρας, φυσικού και ψηφιακού κόσμου.
- Επίσης Θα παρακολουθήσουμε πώς η αναζήτηση κατάλληλου μοντέλου διακυβέρνησης καθορίζεται από τους γεωπολιτικούς ανταγωνισμούς, ενώ συγχρόνως διαπλέκεται με περαιτέρω ζητήματα όπως η κυβερνοασφάλεια και η διακυβέρνηση των δεδομένων.

3.1 Παγκόσμια επισκόπηση πρωτοβουλιών ρύθμισης της ΤΝ

Εάν αναλογιστούμε τι υπηρεσίες θέλουμε να λαμβάνουμε από τη χρήση της τεχνητής νοημοσύνης, είτε ατομικά είτε συλλογικά, οι πιθανές απαντήσεις βασίζονται στις αξίες που έχει αναδείξει πανανθρώπινα ο πολιτισμός, ανεξαρτήτως γεωπολιτικών προτεραιοτήτων, κυβερνητικών επιλογών και πολιτικών καθεστώτων. Από όλες τις διεθνείς διαβουλεύσεις, τις κατευθυντήριες γραμμές, τα έγγραφα εθνικών στρατηγικών, τις διακηρύξεις ηθικών αρχών και δεοντολογίας, συμπεραίνουμε ασφαλέστερα αυτό που δεν θέλουμε να συμβεί από τη χρήση της τεχνητής νοημοσύνης, ως κοινωνίες και άτομα. Δεν θέλουμε μια τεχνητή νοημοσύνη που θα παρακάμπτει την ανθρώπινη σκέψη, θα ποδηγετεί την ελεύθερη βούληση, θα χειραγωγεί, θα κατευθύνει σε υπαγορευμένες επιλογές, θα παραποιεί και θα διαστρεβλώνει την πραγματικότητα, θα οδηγεί σε διακρίσεις, μεροληψία, προκαταλήψεις, ανισότητα, αδικία, έκνομες πρακτικές, παραβίαση δικαιωμάτων, βλάβη της ανθρώπινης ζωής και του περιβάλλοντος.

Καμία παγκοσμίως από τις πολιτικές δυνάμεις που εμπλέκονται στη ρύθμιση της ΤΝ δεν διακηρύσσει κάτι διαφορετικό ως προς τα παραπάνω. Ας επισημανθεί, κάτι όχι ευρύτερα γνωστό στο μέτρο των στερεοτύπων που μας αναλογούν ως Δύσης έναντι του “αλλότριου”, ότι ακόμη και η Κίνα¹ στους κανόνες δεοντολογίας που δημοσίευσε για την ΤΝ αναφέρεται σε έννοιες που μας είναι απολύτως οικείες. Συγκεκριμένα οι “Κανόνες δεοντολογίας για την Τεχνητή Νοημοσύνη

¹China’s New AI Governance Initiatives Shouldn’t Be Ignored, M. Sheehan, 4 Ιανουαρίου 2022, <https://carnegieendowment.org/2022/01/04/china-s-new-ai-governance-initiatives-shouldn-t-be-ignored-pub-86127>

Νέας Γενιάς”², με στόχο την ενσωμάτωση της δεοντολογίας σε όλο τον κύκλο ζωής της Τεχνητής Νοημοσύνης, αντιμετωπίζουν ζητήματα όπως η προστασία της ιδιωτικής ζωής, η προκατάληψη, οι διακρίσεις και η δικαιοσύνη και θέτουν έξι βασικές ηθικές απαιτήσεις: ανθρώπινη ευημερία, δικαιοσύνη, προστασία της ιδιωτικής ζωής, δυνατότητα ελέγχου, ενίσχυση της λογοδοσίας και καλλιέργεια της ηθικής.

Με τη σειρά τους, οι ΗΠΑ εξέδωσαν, τον Οκτώβριο του 2023, το Εκτελεστικό Διάταγμα του Λευκού Οίκου για την ασφαλή, ακίνδυνη και αξιόπιστη ανάπτυξη και χρήση της Τεχνητής Νοημοσύνης, το οποίο ωστόσο βασίζεται σε προγενέστερες ενέργειες της κυβέρνησης Μπάιντεν, όπως ο κατάλογος των Εθελοντικών Δεσμεύσεων³, τις οποίες ανέλαβαν πολλές μεγάλες εταιρείες τεχνολογίας τον Ιούλιο του 2023, και στο Προσχέδιο για μια Διακήρυξη Δικαιωμάτων για την ΤΝ⁴ που δημοσιεύτηκε τον Οκτώβριο του 2022.

Η εμφάνιση του GPT-4 και άλλων παρόμοιων συστημάτων έστρεψε θεαματικά το ενδιαφέρον της διεθνούς κοινότητας στην αναγκαιότητα της θεσμικής ρύθμισης ειδικότερα της Παραγωγικής Τεχνητής Νοημοσύνης καθώς κατέστησε αναγκαία την τροποποίηση του σχεδίου, που ήταν υπό συζήτηση, και την πρόβλεψη ειδικών ρυθμίσεων για την παραγωγική τεχνητή νοημοσύνη. Αυτό κατέδειξε πόσο σημαντική ήταν η πρωτοβουλία που είχαν ήδη εκδηλώσει πολλά κράτη, διεθνείς οργανισμοί και διακρατικοί φορείς στο πεδίο της ρύθμισης της ΤΝ γενικότερα. Κάποια παραδείγματα τέτοιων πρωτοβουλιών ρύθμισης είναι τα παρακάτω:

- η Πράξη της Ευρωπαϊκής Ένωσης για την Τεχνητή Νοημοσύνη⁵,
- οι Συστάσεις της UNESCO για τη Δεοντολογία της Τεχνητής Νοημοσύνης⁶,
- η Διεθνής Σύμβαση του Συμβουλίου της Ευρώπης για την τεχνητή νοημοσύνη, τα ανθρώπινα δικαιώματα, τη δημοκρατία και το κράτος δικαίου,⁷ και η τελική της μορφή που τελικά υιοθετήθηκε στις 17 Μαΐου 2024⁸
- το Εκτελεστικό Διάταγμα για την ασφαλή, ακίνδυνη και αξιόπιστη ανάπτυξη και χρήση της τεχνητής νοημοσύνης που εξέδωσε στις 30 Οκτωβρίου 2023 ο Λευκός Οίκος⁹,
- οι Εθνικές Στρατηγικές για την Τεχνητή Νοημοσύνη διαφόρων χωρών¹⁰,
- οι Αρχές Διακυβέρνησης της Τεχνητής Νοημοσύνης της Κίνας¹¹,
- ο Οδικός Χάρτης για Αξιόπιστη Τεχνητή Νοημοσύνη του Διατλαντικού Συμβουλίου Εμπορίου και Τεχνολογίας¹², ενός φόρουμ για τον συντονισμό των προσεγγίσεων της ΕΕ-ΗΠΑ στην αντιμετώπιση βασικών εμπορικών και τεχνολογικών ζητημάτων,
- οι Αρχές της Τεχνητής Νοημοσύνης του Οργανισμού Οικονομικής Συνεργασίας και Ανάπτυξης (ΟΟΣΑ)¹³,

²Ethical Norms for New Generation Artificial Intelligence Released, 21 Οκτωβρίου 2021, <https://cset.georgetown.edu/publication/ethical-norms-for-new-generation-artificial-intelligence-released/>

³21 Ιουλίου 2023, <https://www.whitehouse.gov/briefing-room/statements-releases/2023/07/21/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-leading-artificial-intelligence-c>

⁴4 Οκτωβρίου 2022, <https://www.whitehouse.gov/ostp/ai-bill-of-rights/>

⁵Πρόταση κανονισμού “Πράξη για την ΤΝ”, COM/2021/206 final, Βρυξέλλες, 21 Απριλίου 2021, <https://eur-lex.europa.eu/legal-content/EL/TXT/?uri=CELEX:52021PC0206>

⁶23 Νοεμβρίου 2021, <https://unesdoc.unesco.org/ark:/48223/pf0000381137>

⁷18 Δεκεμβρίου 2023, <https://rm.coe.int/cai-2023-28-draft-framework-convention/1680ade043>

⁸Council of Europe adopts first international treaty on artificial intelligence, 17 Μαΐου 2024, <https://www.coe.int/en/web/portal/-/council-of-europe-adopts-first-international-treaty-on-artificial-intelligence>

⁹Εκτελεστικό Διάταγμα για την ασφαλή, ακίνδυνη και αξιόπιστη ανάπτυξη και χρήση της τεχνητής νοημοσύνης, Λευκός Οίκος, 30 Οκτωβρίου 2023, <https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/>

¹⁰National AI policies & strategies, <https://oecd.ai/en/dashboards/overview>

¹¹17 Ιουνίου 2019, <https://www.loc.gov/item/global-legal-monitor/2019-09-09/china-ai-governance-principles-released/>

¹²1 Δεκεμβρίου 2022, <https://digital-strategy.ec.europa.eu/en/library/ttc-joint-roadmap-trustworthy-ai-and-risk-management>

¹³22 Μαΐου 2019 και επικαιροποίηση τον Μάιο 2024, <https://oecd.ai/en/ai-principles>

- η Συμφωνία των ηγετών της G7 για διεθνείς κατευθυντήριες αρχές και εθελοντικό κώδικα δεοντολογίας για τους προγραμματιστές ΤΝ στο πλαίσιο της Διαδικασίας της Χιροσίμα για την ΤΝ¹⁴,
- η Διακήρυξη του Bletchley¹⁵ για ανάπτυξη της ΤΝ με ασφαλή και υπεύθυνο τρόπο που υπογράφηκε στη Σύνοδο Κορυφής για την Ασφάλεια της Τεχνητής Νοημοσύνης¹⁶, 1-2 Νοεμβρίου 2023 στο Ηνωμένο Βασίλειο.

“Η Ευρωπαϊκή Ένωση ολοκληρώνει την Πράξη για την Τεχνητή Νοημοσύνη, έναν κανονισμό που αποσκοπεί στη ρύθμιση των πιο “επικίνδυνων” χρήσεων των συστημάτων Τεχνητής Νοημοσύνης. Η πρόταση κανονισμού (Πράξη για την ΤΝ), που δημοσιεύθηκε το 2021, θα ρυθμίζει την ΤΝ σε ποικίλους τομείς όπως η υγειονομική περίθαλψη και η εκπαίδευση.

Πλεονεκτήματα: Το νομοσχέδιο θα καταστήσει υπόλογους όσους επιβουλεύονται τη σύννομη (κατά τις διατάξεις του εν λόγω νομοθετήματος) χρήση της ΤΝ και θα αποτρέψει επαπειλούμενες βλάβες εξαιτίας μιας τέτοιας χρήσης, επιβάλλοντας πρόστιμα και εμποδίζοντας την πώληση και τη χρήση τεχνολογίας ΤΝ εντός ΕΕ όταν αυτή δεν συμμορφώνεται με τους κανόνες της ΕΕ. Το νομοσχέδιο αναμένεται να ρυθμίσει επίσης την παραγωγική ΤΝ (Generative AI) και να επιβάλει περιορισμούς στα συστήματα ΤΝ που θεωρεί απαγορευμένες πρακτικές της ΤΝ διότι δημιουργούν μη αποδεκτό κίνδυνο, όπως η αναγνώριση προσώπου. Δεδομένου ότι πρόκειται για τον μοναδικό ολοκληρωμένο κανονισμό για την τεχνητή νοημοσύνη που υπάρχει, η ΕΕ δείχνει ότι θα έχει το πλεονέκτημα του πρωτοπόρου. Υπάρχει μεγάλη πιθανότητα το καθεστώς της ΕΕ να καταλήξει να είναι η de facto ρύθμιση της ΤΝ παγκοσμίως, επειδή οι εταιρείες σε χώρες εκτός ΕΕ που θέλουν να δραστηριοποιηθούν σε αυτή την αγορά θα πρέπει να προσαρμόσουν τις πρακτικές τους για να συμμορφωθούν με τον νόμο.

Μειονεκτήματα: Πολλά στοιχεία του νομοσχεδίου, όπως η απαγόρευση της αναγνώρισης προσώπου και οι προσεγγίσεις για τη ρύθμιση της παραγωγικής ΤΝ, είναι εξαιρετικά αμφιλεγόμενα και η ΕΕ θα αντιμετωπίσει έντονες πιέσεις από τις εταιρείες τεχνολογίας για την αποδυνάμωσή τους. Θα χρειαστούν αρκετός χρόνος προτού εγκριθεί το νομοσχέδιο από το νομοθετικό σύστημα της ΕΕ και τεθεί σε ισχύ.

Βαθμολόγηση επιδραστικότητας: 5/5”

¹⁴30 Οκτωβρίου 2023, <https://digital-strategy.ec.europa.eu/el/library/hiroshima-process-international-code-conduct-advanced-ai-systems>

¹⁵SUMMIT 2023, AI Safety. The Bletchley Declaration by Countries Attending the AI Safety Summit, <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>

¹⁶<https://www.gov.uk/government/topical-events/ai-safety-summit-2023/about>

Στο πνεύμα αυτό μάς εισάγει η Τεχνολογική Επιθεώρηση του MIT (Technology Review of the Massachusetts Institute of Technology) μέσα από σχετικό άρθρο όπου συνόψισε έξι διαφορετικά κείμενα για τη ρύθμιση της τεχνητής νοημοσύνης.¹⁷ Αφού εξέθεσε τα πλεονεκτήματα και τα μειονεκτήματά τους, τα βαθμολόγησε αναλόγως του πόσο επιδραστικά μπορούν να είναι. Μεταξύ αυτών των έξι κειμένων η νομοθετική πρόταση της ΕΕ “Πράξη για την Τεχνητή Νοημοσύνη” είναι η μόνη που έλαβε άριστη βαθμολογία 5/5 με το εξής σκεπτικό, όπως αυτό αναλύεται στο συγκεκριμένο άρθρο, απόσπασμα του οποίου παρατίθεται στο παραπάνω πλαίσιο.

3.2 Γιατί και σε ποιους τομείς είναι αναγκαία η ρύθμιση

Ακολούθως θα δούμε γιατί και σε ποιους τομείς θεωρείται αναγκαία η ρύθμιση της ΤΝ από τον δημόσιο τομέα, παραθέτοντας τη συρραφή αντιπροσωπευτικών επιχειρημάτων που συνηγορούν σε αυτό και προέρχονται από τα προαναφερθέντα ενδεικτικά ρυθμιστικά κείμενα.

Η **Πράξη της Ευρωπαϊκής Ένωσης για την Τεχνητή Νοημοσύνη**¹⁸ στην αιτιολογική έκθεσή της αναφέρει ότι η τεχνητή νοημοσύνη είναι ένα ταχέως εξελισσόμενο σύνολο τεχνολογιών που μπορούν να αποφέρουν οικονομικά και κοινωνικά οφέλη και να παράσχουν ανταγωνιστικά πλεονεκτήματα στην ευρωπαϊκή οικονομία μέσω της εξατομίκευσης της παροχής υπηρεσιών και της βελτίωσης της πρόβλεψης και της κατανομής των πόρων σε τομείς υψηλού αντικτύπου όπως η κλιματική αλλαγή, το περιβάλλον, η υγεία, ο δημόσιος τομέας, τα οικονομικά, η κινητικότητα, οι εσωτερικές υποθέσεις και η γεωργία. Ωστόσο, τα ίδια στοιχεία και οι τεχνικές, που διαμορφώνουν τα κοινωνικοοικονομικά οφέλη της ΤΝ, μπορούν επίσης να επιφέρουν νέους κινδύνους ή αρνητικές συνέπειες για τα άτομα ή την κοινωνία. Καταλήγει δε ότι είναι προς το συμφέρον της Ένωσης να διατηρηθεί η τεχνολογική πρωτοπορία της ΕΕ και να διασφαλιστεί ότι οι Ευρωπαίοι μπορούν να επωφεληθούν από τις νέες τεχνολογίες που αναπτύσσονται και λειτουργούν σύμφωνα με τις αξίες, τα θεμελιώδη δικαιώματα και τις αρχές της Ένωσης.

Η **UNESCO** στην εισαγωγική ιστοσελίδα της επισημαίνει¹⁹ ότι η ραγδαία άνοδος της τεχνητής νοημοσύνης έχει δημιουργήσει πολλές ευκαιρίες σε παγκόσμιο επίπεδο, από τη διευκόλυνση των ιατρικών διαγνώσεων έως την αύξηση της εργασιακής απόδοσης μέσω αυτοματοποιημένων εργασιών, ωστόσο εγείρει μεγάλο προβληματισμό σε θέματα ηθικής και δεοντολογίας. Αυτό προκύπτει από τη δυνατότητα που έχουν τα συστήματα τεχνητής νοημοσύνης να αναπαράγουν προκαταλήψεις, να επιβαρύνουν το περιβάλλον, να απειλούν θεμελιώδη ανθρώπινα δικαιώματα και γενικότερα να δημιουργούν κίνδυνο επιδείνωσης των υφιστάμενων ανισοτήτων, με αποτέλεσμα να πλήττονται περαιτέρω οι ήδη περιθωριοποιημένες ομάδες.

Το **Συμβούλιο της Ευρώπης** στο προοίμιο του σχεδίου της Σύμβασης Πλαίσιο για την ΤΝ²⁰ (Σχέδιο Σύμβασης για την ΤΝ, Ιανουάριος 2023) τονίζει ότι αποστολή του είναι να διασφαλίζει ότι τα ανθρώπινα δικαιώματα, η δημοκρατία και το κράτος δικαίου προστατεύονται και προωθούνται στο ψηφιακό περιβάλλον. Ως εκ τούτου επικεντρώνεται στη χρήση της ΤΝ με τρόπο που δεν θα παρεμβαίνει αδικαιολόγητα στην άσκηση των ανθρωπίνων δικαιωμάτων και των θεμελιωδών ελευθεριών, δε θα υπονομεύει τη δημοκρατία και δεν θα παραβιάζει το κράτος δικαίου. Τονίζει επίσης την ανάγκη θέσπισης ενός κοινού νομικού πλαισίου που θα θέτει τις θεμελιώδεις αρχές που πρέπει να διέπουν τον σχεδιασμό, την ανάπτυξη και την εφαρμογή συστημάτων τεχνητής νοημοσύνης, το οποίο θα διαφυλάσσει αποτελεσματικά τις κοινές αξίες και ταυτόχρονα θα ευνοεί

¹⁷MIT Technology Review, Our quick guide to the 6 ways we can regulate AI, Melissa Heikkiläarchive, 22 Μαΐου 2023, <https://www.technologyreview.com/2023/05/22/1073482/our-quick-guide-to-the-6-ways-we-can-regulate-ai/>

¹⁸Πρόταση κανονισμού “Πράξη για την ΤΝ”, COM/2021/206 final, Βρυξέλλες, 21 Απριλίου 2021, <https://eur-lex.europa.eu/legal-content/EL/TXT/?uri=CELEX:52021PC0206>

¹⁹<https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>

²⁰6 Ιανουαρίου 2023, <https://rm.coe.int/cai-2023-01-revised-zero-draft-framework-convention-public/1680aa193f>

την καινοτομία.

Περαιτέρω στο ενοποιημένο σχέδιο της υπό επεξεργασία Σύμβασης Πλαίσιο για την ΤΝ²¹ (Ιούλιος 2023) το Συμβούλιο της Ευρώπης εκφράζει την ανησυχία ότι ορισμένα συστήματα τεχνητής νοημοσύνης μπορεί να υπονομεύσουν την ανθρώπινη αξιοπρέπεια και την ατομική αυτονομία, τα ανθρώπινα δικαιώματα και τις θεμελιώδεις ελευθερίες, τη δημοκρατία και το κράτος δικαίου και να επιφέρουν διακρίσεις σε βάρος των γυναικών, των παιδιών και άλλων ομάδων. Επιπλέον εκφράζει ανησυχία για την τεκμηριωμένη και συνεχιζόμενη χρήση συστημάτων τεχνητής νοημοσύνης από ορισμένα κράτη για κατασταλτικούς σκοπούς, συχνά με την αξιοποίηση εργαλείων του ιδιωτικού τομέα, κατά παράβαση του διεθνούς δικαίου των ανθρωπίνων δικαιωμάτων, μέσω πρακτικών αυθαίρετης ή παράνομης παρακολούθησης και λογοκρισίας που διαβρώνουν την ιδιωτική ζωή και την αυτονομία.

Όσον αφορά το **Εκτελεστικό Διάταγμα του Λευκού Οίκου για την ασφαλή, ακίνδυνη και αξιόπιστη ανάπτυξη και χρήση της Τεχνητής Νοημοσύνης** που δημοσιεύθηκε στις 30 Οκτωβρίου 2023²², πρόκειται για ένα νομικό κείμενο το οποίο ωθεί τις ΗΠΑ προς μια πιο ολοκληρωμένη διακυβέρνηση της ΤΝ. Το νέο αυτό Εκτελεστικό Διάταγμα επιπλέον ακολουθεί δύο άλλα προηγούμενα εκτελεστικά διατάγματα με επίκεντρο την ΤΝ: ένα σχετικά με τη χρήση της ΤΝ από την ίδια την ομοσπονδιακή κυβέρνηση και ένα άλλο με στόχο την ενίσχυση των ομοσπονδιακών προσλήψεων στον τομέα της ΤΝ.

Σε αντίθεση με αυτά τα προηγούμενα διατάγματα, το από 30 Οκτωβρίου 2023 εκτελεστικό διάταγμα πηγαίνει μακρύτερα από το να διατυπώνει γενικές αρχές και κατευθυντήριες γραμμές και θέτει απαιτήσεις για συγκεκριμένη δράση από τις εταιρείες τεχνολογίας και τις ομοσπονδιακές υπηρεσίες²³. Για παράδειγμα, το νέο εκτελεστικό διάταγμα επιβάλλει στους προγραμματιστές ΤΝ να μοιράζονται δεδομένα ασφαλείας, πληροφορίες εκπαίδευσης και εκθέσεις με την κυβέρνηση των ΗΠΑ πριν από τη δημόσια διάθεση μελλοντικών μεγάλων μοντέλων ΤΝ ή ενημερωμένων εκδόσεων τέτοιων μοντέλων. Συγκεκριμένα, η απαίτηση ισχύει για μοντέλα που περιέχουν “δεκάδες δισεκατομμύρια παραμέτρους”, τα οποία εκπαιδεύτηκαν σε δεδομένα μεγάλης κλίμακας και θα μπορούσαν να θέσουν σε κίνδυνο την εθνική ασφάλεια, την οικονομία, τη δημόσια υγεία ή την ασφάλεια.

Η πολιτική του Λευκού Οίκου απαιτεί επίσης τη δημιουργία ομοσπονδιακών προτύπων και δοκιμών που θα αναπτυχθούν από υπηρεσίες όπως το Υπουργείο Εσωτερικής Ασφάλειας και το Υπουργείο Ενέργειας για να διασφαλιστεί καλύτερα ότι η τεχνητή νοημοσύνη δεν απειλεί την εθνική ασφάλεια. Τα εν λόγω πρότυπα θα αναπτυχθούν εν μέρει από το Εθνικό Ινστιτούτο Προτύπων και Τεχνολογίας, το οποίο δημοσίευσε τον Ιανουάριο το δικό του πλαίσιο για τη διαχείριση των κινδύνων της τεχνητής νοημοσύνης. Η διαδικασία ανάπτυξης θα περιλαμβάνει “red-teaming”, όπου καλοπροαίρετοι χάκερς θα συνεργάζονται με τους δημιουργούς του μοντέλου για να αναλύσουν προληπτικά τα τρωτά σημεία.

Πέρα από αυτές τις εντολές, το εκτελεστικό διάταγμα δημιουργεί κυρίως ομάδες εργασίας και συμβουλευτικές επιτροπές, προτρέπει σε πρωτοβουλίες υποβολής εκθέσεων και κατευθύνει τις ομοσπονδιακές υπηρεσίες να εκδώσουν κατευθυντήριες γραμμές για την ΤΝ εντός του επόμενου έτους. Το διάταγμα καλύπτει οκτώ πεδία που περιγράφονται σε ενημερωτικό δελτίο: εθνική ασφάλεια, προστασία ιδιωτικότητας, ισότητα και πολιτικά δικαιώματα, προστασία των καταναλωτών, εργασιακά ζητήματα, καινοτομία και ανταγωνιστικότητα των ΗΠΑ, διεθνής συνεργασία για την πολιτική της ΤΝ καθώς και δεξιότητες στην ΤΝ εντός της ομοσπονδιακής κυβέρνησης. Εντός

²¹ 7 Ιουλίου 2023, <https://rm.coe.int/cai-2023-18-consolidated-working-draft-framework-convention/1680abde66>

²² Εκτελεστικό Διάταγμα για την ασφαλή, ακίνδυνη και αξιόπιστη ανάπτυξη και χρήση της τεχνητής νοημοσύνης, Λευκός Οίκος, 30 Οκτωβρίου 2023, <https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/>

²³ 30 Οκτωβρίου 2023, <https://www.scientificamerican.com/article/bidens-executive-order-on-ai-is-a-good-start-experts-say-but-not-enough>

αυτών των οριζόντιων κατηγοριών υπάρχουν ενότητες σχετικά με την αξιολόγηση και την προώθηση της ηθικής χρήσης της ΤΝ στην εκπαίδευση, την υγειονομική περίθαλψη και την ποινική δικαιοσύνη.

Το διατλαντικό **Συμβούλιο Εμπορίου και Τεχνολογίας ΗΠΑ-ΕΕ** ιδρύθηκε κατά τη διάρκεια της συνόδου κορυφής ΕΕ-ΗΠΑ στις 15 Ιουνίου 2021 στις Βρυξέλλες και αποσκοπεί στην προώθηση του ψηφιακού μετασχηματισμού και τη συνεργασία στις νέες τεχνολογίες, με βάση τις κοινές δημοκρατικές αξίες και τα ανθρώπινα δικαιώματα²⁴. Χρησιμεύει ως φόρουμ για τον συντονισμό τους σε βασικά παγκόσμια εμπορικά, οικονομικά και τεχνολογικά ζητήματα με βάση τις κοινές τους δημοκρατικές αξίες. Το Συμβούλιο με κοινή δήλωση τον Δεκέμβριο του 2022 κατόρθωσε Κοινό Οδικό Χάρτη²⁵ που αποσκοπεί στην ανάπτυξη εργαλείων, μεθοδολογιών και προσεγγίσεων για τη διαχείριση κινδύνων και στην προώθηση διεθνών προσπαθειών τυποποίησης και δημιουργίας τεχνικών προτύπων για την ΤΝ.

Ο οδικός χάρτης λαμβάνει πρακτικά μέτρα για την προώθηση της αξιοπιστίας ΤΝ και τη διατήρηση της κοινής δέσμευσης στις συστάσεις του Οργανισμού Οικονομικής Συνεργασίας και Ανάπτυξης (ΟΟΣΑ) για την ΤΝ. Το Συμβούλιο συνέρχεται σε τακτικές συνεδριάσεις και μάλιστα κατά την τελευταία, τον Μάιο του 2023 παρουσίασε τα πεπραγμένα σχετικά με την ΤΝ²⁶ μεταξύ των οποίων μια λίστα με 65 βασικούς όρους τεχνητής νοημοσύνης που είναι απαραίτητοι για την κατανόηση των προσεγγίσεων της τεχνητής νοημοσύνης που βασίζονται στον κίνδυνο, μαζί με τις ερμηνείες που τους αποδίδονται στις ΗΠΑ και στην ΕΕ²⁷.

Εν κατακλείδι, κοινός παρονομαστής πολλών από τα παραπάνω ρυθμιστικά κείμενα είναι το σχόλιο που δανειζόμαστε από το Digital Bridge του Politico στις 2 Νοεμβρίου 2023²⁸, σχετικά με το Εκτελεστικό Διάταγμα του Λευκού Οίκου: “είναι λεπτομερές και εκτεταμένο έγγραφο αλλά του λείπει ένας βασικός παράγοντας: τα μέσα επιβολής αυτών των απαιτήσεων”.

3.3 Η Ευρώπη πρωτοστατεί στη ρύθμιση με γεωπολιτικούς στόχους

3.3.1 Το σκεπτικό που οδήγησε την Ευρώπη στη ρύθμιση της ΤΝ

Αποτελεί κοινή παραδοχή ότι η τεχνητή νοημοσύνη έχει ήδη προσφέρει πολύτιμα εργαλεία που χρησιμοποιούνται καθημερινά σε όλο τον κόσμο. Η συνεχής ανάπτυξή της, με γνώμονα κοινά αποδεκτές αρχές, θα προσφέρει μοναδικές ευκαιρίες βοήθειας και ενδυνάμωσης των ατόμων και των κοινωνιών στο μέλλον²⁹ και αυτή τη διαπίστωση έχει πολλές φορές διατρανώσει η ΕΕ σε ποικίλα κείμενά της.

Στην αιτιολογική έκθεση της Πράξης για την ΤΝ διαβάζουμε ότι πρόκειται για ένα ταχέως εξελισσόμενο σύνολο τεχνολογιών που μπορούν να αποφέρουν ευρείας κλίμακας οικονομικά και κοινωνικά οφέλη σε ολόκληρο το φάσμα των κλάδων και των κοινωνικών δραστηριοτήτων. Μέσω της βελτίωσης της πρόβλεψης, της βελτιστοποίησης των δραστηριοτήτων και της κατανομής των πόρων, καθώς και της εξατομίκευσης της παροχής υπηρεσιών, η χρήση της τεχνητής νοημοσύνης μπορεί να στηρίξει κοινωνικά και περιβαλλοντικά επωφελή αποτελέσματα και να παράσχει βασικά ανταγωνιστικά πλεονεκτήματα στις εταιρείες και στην ευρωπαϊκή οικονομία. Η δράση αυτή είναι ιδιαίτερα αναγκαία σε τομείς υψηλού αντικτύπου, μεταξύ των οποίων συγκαταλέγονται η κλιματική αλλαγή, το περιβάλλον και η υγεία, ο δημόσιος τομέας, τα οικονομικά, η κινητικότητα,

²⁴<https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/stronger-europe-world/eu-us-trade-and-technology-council>

²⁵<https://digital-strategy.ec.europa.eu/en/library/ttc-joint-roadmap-trustworthy-ai-and-risk-management>

²⁶<https://www.whitehouse.gov/briefing-room/statements-releases/2023/05/31/u-s-eu-joint-statement-of-the-trade-and-technology-council-2/>

²⁷<https://www.nist.gov/artificial-intelligence/ai-policy-contributions>

²⁸Mark Scott, The week when AI and geopolitics collided, 2 Νοεμβρίου 2023, <https://www.politico.eu/newsletter/digital-bridge/the-week-where-ai-and-geopolitics-collided/>

²⁹<https://futureoflife.org/open-letter/ai-principles/>

οι εσωτερικές υποθέσεις και η γεωργία. Ωστόσο, τα ίδια στοιχεία και τεχνικές που διαμορφώνουν τα κοινωνικοοικονομικά οφέλη της ΤΝ μπορούν επίσης να επιφέρουν νέους κινδύνους ή αρνητικές συνέπειες για τα άτομα ή την κοινωνία³⁰.

Η Λευκή Βίβλος για την Τεχνητή Νοημοσύνη και την ευρωπαϊκή προσέγγιση της αριστείας και της εμπιστοσύνης³¹ καθόρισε επιλογές πολιτικής σχετικά με τον τρόπο επίτευξης του διττού στόχου που συνίσταται αφενός στην προώθηση της αποδοχής της ΤΝ και της αξιοποίησης των οικονομικών πλεονεκτημάτων που προσφέρει, και αφετέρου στην αντιμετώπιση των κινδύνων που συνδέονται με ορισμένες χρήσεις της εν λόγω τεχνολογίας. Η ΤΝ θα πρέπει να αποτελεί εργαλείο για τους ανθρώπους και να λειτουργεί ως δύναμη για το καλό της κοινωνίας, με απώτερο στόχο την αύξηση της ανθρώπινης ευημερίας. Οι κανόνες για την ΤΝ που διατίθεται στην αγορά της Ένωσης ή επηρεάζει με άλλον τρόπο τους πολίτες της Ένωσης θα πρέπει, συνεπώς, να είναι ανθρωποκεντρικοί, ώστε οι πολίτες να μπορούν να έχουν τη βεβαιότητα ότι η τεχνολογία χρησιμοποιείται κατά τρόπο ασφαλή και συμβατό με τη νομοθεσία.

Μετά τη δημοσίευση της Λευκής Βίβλου, η Επιτροπή ξεκίνησε ευρεία διαβούλευση με τα ενδιαφερόμενα μέρη, την οποία παρακολούθησαν με μεγάλο ενδιαφέρον πολυάριθμα ενδιαφερόμενα μέρη που υποστήριξαν σημαντικά τη ρυθμιστική παρέμβαση για την αντιμετώπιση των προκλήσεων και των ανησυχιών που εγείρει η αυξανόμενη χρήση της ΤΝ. Η Πράξη της ΕΕ για την Τεχνητή Νοημοσύνη αποσκοπεί στην υλοποίηση του δεύτερου στόχου για την ανάπτυξη ενός οικοσυστήματος εμπιστοσύνης, προτείνοντας ένα νομικό πλαίσιο για αξιόπιστη ΤΝ. Βασίζεται στις αξίες και στα θεμελιώδη δικαιώματα της ΕΕ και έχει ως στόχο να ενισχύσει την εμπιστοσύνη των πολιτών και άλλων χρηστών ώστε να υιοθετήσουν λύσεις που βασίζονται στην ΤΝ, ενθαρρύνοντας παράλληλα τις επιχειρήσεις να αναπτύξουν τέτοιες λύσεις.

3.3.2 Επιδίωξη για πρωτοπορία της Ευρώπης

Λαμβανομένων υπόψη της ταχύτητας των τεχνολογικών αλλαγών και των διαρκών προκλήσεων που θέτουν στο προσκήνιο, η ΕΕ διακηρύσσει ότι καταβάλλει προσπάθειες για μια ισορροπημένη προσέγγιση και αναγνωρίζει πως είναι προς το συμφέρον της Ένωσης να διατηρηθεί η τεχνολογική πρωτοπορία της ΕΕ και να διασφαλιστεί ότι οι Ευρωπαίοι μπορούν να επωφεληθούν από τις νέες τεχνολογίες που αναπτύσσονται και λειτουργούν σύμφωνα με τις αξίες, τα θεμελιώδη δικαιώματα και τις αρχές της Ένωσης. Ωστόσο, οι ρυθμιστικές πρωτοβουλίες ασθμαίνουν καθώς προσπαθούν να συγχρονιστούν με την αλματώδη πρόοδο που συντελείται καθημερινά χάρη στις γοργές τεχνολογικές εξελίξεις. Το φαινόμενο των καθυστερήσεων εντείνεται δεδομένου ότι ακολουθείται η “συνήθης νομοθετική διαδικασία” λήψης αποφάσεων της ΕΕ για τη ρύθμιση της ΤΝ, στο πλαίσιο της οποίας το Ευρωπαϊκό Κοινοβούλιο και το Συμβούλιο της Ευρωπαϊκής Ένωσης ενεργούν από κοινού ως “συννομοθέτες”³² (παλαιότερα γνωστή ως “διαδικασία συναπόφασης”).

Η Επιτροπή πρότεινε το Σχέδιο Κανονισμού για την Τεχνητή Νοημοσύνη τον Απρίλιο 2021, καθιερώνοντας μια προσέγγιση για τη ρύθμιση των εφαρμογών της τεχνητής νοημοσύνης με βάση τον κίνδυνο και αναθέτοντας στους συννομοθέτες, δηλαδή το Κοινοβούλιο και το Συμβούλιο, την ψήφιση του πρώτου οριζόντιου κανονισμού για την τεχνητή νοημοσύνη στον κόσμο. Σκοπός είναι, πέραν του να καταστεί η Ευρώπη παγκόσμιος κόμβος αξιόπιστης τεχνητής νοημοσύνης που θα

³⁰ ΑΙΤΙΟΛΟΓΙΚΗ ΕΚΘΕΣΗ, 1. ΠΛΑΙΣΙΟ ΤΗΣ ΠΡΟΤΑΣΗΣ, 1.1. Αιτιολόγηση και στόχοι της πρότασης, Πρόταση, ΚΑΝΟΝΙΣΜΟΣ ΤΟΥ ΕΥΡΩΠΑΪΚΟΥ ΚΟΙΝΟΒΟΥΛΙΟΥ ΚΑΙ ΤΟΥ ΣΥΜΒΟΥΛΙΟΥ ΓΙΑ ΤΗ ΘΕΣΠΙΣΗ ΕΝΑΡΜΟΝΙΣΜΕΝΩΝ ΚΑΝΟΝΩΝ ΣΧΕΤΙΚΑ ΜΕ ΤΗΝ ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ “Πράξη για την ΤΝ” ΚΑΙ ΓΙΑ ΤΗΝ ΤΡΟΠΟΠΟΙΗΣΗ ΟΡΙΣΜΕΝΩΝ ΝΟΜΟΘΕΤΙΚΩΝ ΠΡΑΞΕΩΝ ΤΗΣ ΕΝΩΣΗΣ, COM(2021) 206 final, Βρυξέλλες, 21 Απριλίου 2021, <https://eur-lex.europa.eu/legal-content/EL/TXT/?uri=CELEX:52021PC0206>

³¹ ΛΕΥΚΗ ΒΙΒΛΟΣ Τεχνητή Νοημοσύνη - Ευρωπαϊκή προσέγγιση της αριστείας και της εμπιστοσύνης, COM(2020)65 final, Βρυξέλλες, 19 Φεβρουαρίου 2020, <https://op.europa.eu/el/publication-detail/-/publication/ac957f13-53c6-11ea-aece-01aa75ed71a1>

³² https://eur-lex.europa.eu/legal-content/EL/TXT/?uri=LEGISSUM:ordinary_legislative_procedure

εγγυάται την υγεία, την ασφάλεια και τα θεμελιώδη δικαιώματα των πολιτών της, να καθιερωθεί ως πρωτοπόρος και ηγέτιδα δύναμη στη ρύθμιση και την ανάπτυξη υπηρεσιών και λογισμικού τεχνητής νοημοσύνης, διασφαλίζοντας συγχρόνως ότι η ΤΝ θα έχει θετικό αντίκτυπο για τους Ευρωπαίους πολίτες και θα εναρμονίζεται με τις ευρωπαϊκές αξίες και τα θεμελιώδη δικαιώματα³³.

Κατά την υποβολή της πρότασης της Πράξης για την Τεχνητή Νοημοσύνη, η Margrethe Vestager, εκτελεστική αντιπρόεδρος για “Μια Ευρώπη έτοιμη για την ψηφιακή εποχή”, δήλωσε σχετικά: “Με αυτούς τους κανόνες-ορόσημο, η ΕΕ πρωτοστατεί στην ανάπτυξη νέων παγκόσμιων προτύπων ώστε να διασφαλιστεί ότι μπορούμε να έχουμε εμπιστοσύνη στην τεχνητή νοημοσύνη. Καθορίζοντας τα πρότυπα, μπορούμε να προετοιμάσουμε το έδαφος για μια τεχνολογία βασιζόμενη σε δεοντολογικές αρχές σε παγκόσμιο επίπεδο και να διασφαλίσουμε ότι η ΕΕ παραμένει ανταγωνιστική καθ’ όλη τη διαδικασία. Οι κανόνες μας, διαχρονικοί και φιλικόι προς την καινοτομία, θα παρεμβαίνουν όπου είναι απολύτως αναγκαίο: όταν διακυβεύονται η ασφάλεια και τα θεμελιώδη δικαιώματα των πολιτών της ΕΕ.” Με την ίδια αφορμή ο Επίτροπος Εσωτερικής Αγοράς, Thierry Breton δήλωσε: “Στόχος των σημερινών προτάσεων είναι να ενισχυθεί η θέση της Ευρώπης ως παγκόσμιος κόμβος αριστείας στον τομέα της ΤΝ από το εργαστήριο στην αγορά, να διασφαλιστεί ότι η ΤΝ στην Ευρώπη σέβεται τις αξίες και τους κανόνες μας και να αξιοποιηθεί το δυναμικό της ΤΝ για βιομηχανική χρήση.”

3.4 Γεωπολιτική και τεχνολογία

Το διακύβευμα της πρωτοπορίας στη ρύθμιση της ΤΝ παρουσιάζεται πολύ γλαφυρά στη Γεωστρατηγική Έκθεση για το 2022 που εκπόνησε η Ernst & Young³⁴ όπου συνόψισε το παγκόσμιο περιβάλλον πολιτικών κινδύνων (political risks) που προέβλεπε η Ομάδα Γεωστρατηγικών και Επιχειρηματικών Μελετών για το 2022. Όρισε κατ’ αρχάς τον πολιτικό κίνδυνο ως την πιθανότητα οι πολιτικές αποφάσεις, τα γεγονότα ή οι συνθήκες σε γεωπολιτικό, εθνικό, ρυθμιστικό ή κοινωνικό επίπεδο να επηρεάσουν την απόδοση μιας εταιρείας, μιας αγοράς ή μιας οικονομίας, προκαλώντας είτε προκλήσεις είτε ευκαιρίες. Η έκθεση παρουσίασε τις 10 σημαντικότερες εξελίξεις πολιτικού κινδύνου τόσο ως προς την υψηλή πιθανότητα να συμβούν όσο και ως προς τον υψηλό αντίκτυπο που θα είχαν. Μεταξύ αυτών, οι εξής δύο έκαναν ιδιαίτερη μνεία στη Τεχνητή Νοημοσύνη, υπό τους τίτλους:

1. “Τεχνολογικός εθνικισμός”
2. “το Φαινόμενο των Βρυξελλών”.

Σε αυτή τη βάση θα δούμε αμέσως παρακάτω πώς προέβλεπαν οι μελετητές της Ernst & Young τις γεωπολιτικές εξελίξεις που θα ήταν συνδεδεμένες με την ΤΝ, στο κατώφλι του 2022, μια χρονιά που αποδείχτηκε σταθμός για την επιδραστικότητα της ΤΝ στο ευρύ κοινό λόγω της εμφάνισης του GPT-4³⁵ στις 30 Νοεμβρίου 2022.

³³Μια Ευρώπη έτοιμη για την ψηφιακή εποχή: η Επιτροπή προτείνει νέους κανόνες και δράσεις για την αριστεία και την εμπιστοσύνη στην τεχνητή νοημοσύνη, Βρυξέλλες, 21 Απριλίου 2021, https://ec.europa.eu/commission/presscorner/detail/el/IP_21_1682

³⁴Ernst & Young, 2022 Geostrategic Outlook, How to thrive in a turbulent 2022, December 2021, https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/geostrategy/ey-2022-geostrategic-outlook.pdf

³⁵<https://openai.com/blog/chatgpt>

3.4.1 “Εθνικισμός” μέσω της τεχνολογίας

Οι περιορισμοί στον εφοδιασμό και οι αλλαγές πολιτικής στους ημιαγωγούς, ο μεγαλύτερος ρυθμιστικός και πολιτικός έλεγχος επί των μεγάλων τεχνολογικών εταιρειών και η αυξανόμενη χρήση βιομηχανικών πολιτικών οδήγησαν στη μεγέθυνση του γεωπολιτικού ρόλου της τεχνολογίας και των δεδομένων. Τεχνολογίες όπως η τεχνητή νοημοσύνη, η κβαντική πληροφορική, τα δίκτυα 5G, οι μπαταρίες υψηλής χωρητικότητας και η διαστημική τεχνολογία αναδεικνύονται ολοένα και περισσότερο στην πρώτη γραμμή του γεωστρατηγικού ανταγωνισμού, όπου η επιτυχία ενός έθνους καθορίζεται από το πόσο αυτό καινοτομεί και επιβάλλεται με την τεχνολογική του πρόοδο. Ένας τρόπος με τον οποίο αυτός ο ιδίotypος “εθνικισμός” μέσω της τεχνολογίας εκδηλώνεται είναι μέσω αυστηρότερων ρυθμίσεων διακυβέρνησης των ψηφιακών τεχνολογιών, με την ΕΕ να είναι πρωτοπόρος σε αυτόν τον τομέα. Ομοίως, η ρυθμιστική ατζέντα της ΕΕ για κανόνες γύρω από τη δεοντολογία και τις ηθικές χρήσεις της τεχνητής νοημοσύνης θα μπορούσε να θεωρηθεί και ως μια έκφραση του “τεχνολογικού εθνικισμού”.

Παράλληλα ο ανταγωνισμός για το πλεονέκτημα της ηγέτιδας δύναμης στον καθορισμό προτύπων γίνεται όλο και πιο έντονος. Η ΕΕ έχει μια αργή νομοθετική διαδικασία, ιδίως σε σύγκριση με την Κίνα³⁶, η οποία έχει αποδυθεί στον αγώνα για τη θέσπιση τεχνολογικών προτύπων μέσω της πρωτοβουλίας China Standards 2035³⁷. Το νέο Διατλαντικό Συμβούλιο Εμπορίου και Τεχνολογίας³⁸ (TTC, Transatlantic Trade and Technology Council) είναι πιθανό να αποτελέσει ένα σημαντικό φόρουμ για διεθνείς διαπραγματεύσεις, αλλά τα ανταγωνιστικά συμφέροντα πιθανόν να αυξήσουν τον κίνδυνο ασύμβατων φυγόκεντρων προτύπων και να καταστήσουν τις φιλόδοξες συμφωνίες δύσκολες. Ωστόσο, η ΕΕ προχώρησε στη δημιουργία γραφείου στο Σαν Φρανσίσκο για την ενίσχυση της ψηφιακής διπλωματίας της τον Σεπτέμβριο του 2022³⁹.

Επιπλέον είναι σαφέστατο ότι η πρωτοβουλία της Ευρωπαϊκής Ένωσης για την Πράξη για την Τεχνητή Νοημοσύνη έχει καταστεί επίκαιρη - δεδομένου ότι η τάση έχει κορυφωθεί, με αφορμή τις εξελίξεις στο πεδίο των ισχυρών εργαλείων παραγωγικής τεχνητής νοημοσύνης, όπως το DALL-E 2⁴⁰, το Chat GPT-4⁴¹, το Bard⁴² και άλλα παρόμοια συστήματα.

Ωστόσο, όσον αφορά την ψηφιακή οικονομία, η Κίνα κάνει μεγάλα βήματα στην επέκταση της επιρροής της με άλλα μέσα. Οι κινεζικές εταιρείες, εκ των οποίων όλες διατηρούν δεσμούς με το Κινεζικό Κομμουνιστικό Κόμμα, έχουν προμηθεύσει κρίσιμες τεχνολογικές υποδομές σε χώρες σε όλο τον κόσμο, εξάγοντας τα αυταρχικά τους πρότυπα. Η Κίνα έχει επίσης προμηθεύσει τεχνολογία επιτήρησης τεχνητής νοημοσύνης σε πολυάριθμες κυβερνήσεις που είναι πρόθυμες να χρησιμοποιήσουν την κινεζική τεχνολογία για περιορισμό των ελευθεριών. Συνεπώς παρατηρούμε τη διάσπαση της παγκόσμιας ψηφιακής οικονομίας, με πολλές κινεζικές εταιρείες να δραστηριοποιούνται στην εγχώρια αγορά τους, ενώ οι αμερικανικοί τεχνολογικοί κολοσσοί απομακρύνονται από την Κίνα. Αυτό έχει ως αποτέλεσμα το ότι πολλές κινεζικές εταιρείες τεχνολογίας δεν εμπίπτουν καν στην αρμοδιότητα των ρυθμιστικών αρχών των ΗΠΑ και της ΕΕ - εκτός από εταιρείες όπως η Huawei και η TikTok που επιδιώκουν να εισέλθουν βαθύτερα στις αγορές της ΕΕ και των ΗΠΑ. Οι εταιρείες αυτές αντιμετωπίζουν ένα όλο και πιο εχθρικό ρυθμιστικό περιβάλλον στις ΗΠΑ και την ΕΕ.

³⁶China Is Quietly Rewriting Rules That Run the World, 5 Νοεμβρίου 2022, <https://www.bloomberg.com/news/newsletters/2022-11-05/china-rewrites-the-rules-that-run-the-world-new-economy-saturday>

³⁷Εθνικό Πρόγραμμα-Πλαίσιο Τυποποίησης, Πεκίνο, 10 Οκτωβρίου 2021, http://www.gov.cn/zhengce/2021-10/10/content_5641727.htm

³⁸Digital in the EU-US Trade and Technology Council, <https://digital-strategy.ec.europa.eu/en/policies/trade-and-technology-council>

³⁹EU opens new office in San Francisco to reinforce its digital diplomacy, 5 Σεπτεμβρίου 2022, <https://digital-strategy.ec.europa.eu/en/news/eu-opens-new-office-san-francisco-reinforce-its-digital-diplomacy>

⁴⁰<https://openai.com/dall-e-2>

⁴¹<https://openai.com/gpt-4>

⁴²<https://bard.google.com/>

3.4.2 Brussels effect (Φαινόμενο των Βρυξελλών)

Ομοίως η έκθεση αυτή επισημαίνει ότι η επιβολή ρυθμίσεων και η από κοινού θέσπιση προτύπων είχαν ανέκαθεν ζωτική σημασία για την ευρωπαϊκή ολοκλήρωση και την ενιαία αγορά. Όχι μόνο επηρέασαν τις ευρωπαϊκές επιχειρήσεις και τους καταναλωτές αλλά έδωσαν στην ΕΕ την ισχύ να καθιερώσει ορισμένα παγκόσμια πρότυπα. Αυτό ονομάστηκε Φαινόμενο των Βρυξελλών⁴³ και το πιο αντιπροσωπευτικό παράδειγμα είναι ο τρόπος με τον οποίο ο Ευρωπαϊκός Γενικός Κανονισμός για την Προστασία Δεδομένων (ΓΚΠΔ) είχε διεθνές αποτέλεσμα και επιπλέον ενέπνευσε άλλες κυβερνήσεις σε όλο τον κόσμο να θεσπίσουν παρόμοιες πολιτικές. Η σημασία του Φαινομένου των Βρυξελλών αναμένεται να αυξάνεται διαρκώς καθώς η ΕΕ ξεκινά ένα φιλόδοξο ρυθμιστικό πρόγραμμα και γίνεται πιο διεκδικητική ως προς τον διεθνή της ρόλο στο να θέτει παγκόσμια πρότυπα. Μάλιστα το Φαινόμενο των Βρυξελλών έχει δύο σκέλη:

- (α) De Facto: Οι εταιρείες ακολουθούν καθολικά τις απαιτήσεις της ΕΕ για την τυποποίηση προϊόντων ή υπηρεσιών ώστε να μπορούν να κινούνται σε όλες τις αγορές χωρίς ειδική προσαρμογή και συνεπώς χωρίς επιπλέον έξοδα.
- (β) De Jure: Άλλες χώρες, εκτός ΕΕ, εκδίδουν επίσημη νομοθεσία που ευθυγραμμίζεται με τον κανονισμό της ΕΕ για να αποφευχθούν συγκρούσεις με την πρόσφατα τυποποιημένη διαδικασία.

Για την καλύτερη παρουσίαση του Brussels Effect δεν έχουμε παρά να προστρέξουμε στην ίδια την Anu Bradford που επινόησε τον όρο. Μέσα από μια συνέντευξή της⁴⁴ πληροφορούμαστε ότι το Brussels Effect αναφέρεται στη μονομερή ικανότητα της Ευρωπαϊκής Ένωσης να ρυθμίζει τις παγκόσμιες αγορές και καταδεικνύει την ικανότητα της ΕΕ να διαμορφώνει το διεθνές ρυθμιστικό περιβάλλον χωρίς να χρειάζεται να καταφεύγει σε εξαναγκασμό ή συνεργασία. Γεννήθηκε από την ανάγκη της ΕΕ να οικοδομήσει και να προστατεύσει τη στρατηγική της αυτονομία και να αναλάβει έναν πιο διεκδικητικό γεωπολιτικό ρόλο.

Καθώς η ΕΕ είναι μία από τις μεγαλύτερες και πιο εύπορες καταναλωτικές αγορές του κόσμου, οι πολυεθνικές εταιρείες αποδέχονται τη συμμόρφωση με τους κανονισμούς της ΕΕ ως τίμημα για τη δραστηριοποίησή τους στην Ευρώπη. Ωστόσο, οι εταιρείες προτιμούν την ομοιομορφία και, ως εκ τούτου, συχνά επεκτείνουν εθελοντικά αυτούς τους κανόνες της ΕΕ στις παγκόσμιες δραστηριότητές τους για να αποφύγουν το κόστος συμμόρφωσης με πολλαπλά ρυθμιστικά καθεστώτα. Έτσι το Brussels Effect διαμορφώνει πολλούς τομείς πολιτικής, συμπεριλαμβανομένης της ψηφιακής οικονομίας.

Αντίστοιχη διευρυμένη νομική επιδραστικότητα παρουσιάζουν οι Πράξεις Digital Services Act (DSA) και Digital Markets Act (DMA)⁴⁵, οι οποίες έχουν ως στόχο να θέσουν περιορισμούς στις μεγάλες εταιρείες τεχνολογίας που δραστηριοποιούνται στην Ευρώπη και θα αναλυθούν διεξοδικότερα παρακάτω μαζί με το, άρρηκτα συνδεδεμένο με την ανάπτυξη της ΤΝ, ρυθμιστικό τοπίο για τα Δεδομένα.

ΣΥΓΚΡΙΣΗ ΤΟΥ DMA ΚΑΙ ΤΟΥ DSA	ΤΙ ΘΑ ΡΥΘΜΙΖΟΥΝ;	ΠΟΙΟΣ ΥΠΟΚΕΙΤΑΙ ΣΤΟΥΣ ΚΑΝΟΝΙΣΜΟΥΣ ΑΥΤΟΥΣ;	ΠΟΤΕ ΘΑ ΕΦΑΡΜΟΣΤΟΥΝ ΟΙ ΚΑΝΟΝΙΣΜΟΙ;
ΝΟΜΟΣ ΓΙΑ ΤΙΣ ΨΗΦΙΑΚΕΣ ΑΓΟΡΕΣ (DMA)	Ο DMA ασχολείται με θέματα ανταγωνισμού και αντιμονοπωλιακά ζητήματα. Θα επεκτείνει τα υπάρχοντα μέτρα δημιουργώντας προληπτικούς κανόνες που απαγορεύουν συγκεκριμένες συμπεριφορές.	Πάροχοι υπηρεσιών βασικής πλατφόρμας (CPS) Πρακτικά σημαίνει ότι αφορά μόνο έναν μικρό αριθμό μεγάλων τεχνολογικών εταιρειών. Στοχεύει εταιρείες με ετήσιο κύκλο εργασιών άνω των 7,5 δισεκατομμυρίων ευρώ για κάθε μία από τις τρεις τελευταίες οικονομικές χρήσεις ή με μέση κεφαλαιοποίηση αγοράς ή αντίστοιχη εύλογη αξία τουλάχιστον 75 δισεκατομμύρια ευρώ την τελευταία οικονομική χρήση.	Θα τεθεί σε ισχύ από την 1η Νοεμβρίου 2022 και θα αρχίσει να εφαρμόζεται από την 2α Μαΐου 2023.
ΝΟΜΟΣ ΓΙΑ ΤΙΣ ΨΗΦΙΑΚΕΣ ΥΠΗΡΕΣΙΕΣ (DSA)	Ο DSA ασχολείται με επιβλαβή και παράνομα αγαθά, υπηρεσίες και περιεχόμενο στο διαδίκτυο. Θα αντικαταστήσει την Οδηγία για το Ηλεκτρονικό Εμπόριο (2000).	Ενδιάμεσες υπηρεσίες - Πάροχοι βασικών ενδιάμεσων υπηρεσιών - Πάροχοι υπηρεσιών φιλοξενίας - Διαδικτυακές πλατφόρμες, συμπεριλαμβανομένων των αγορών - Πολύ μεγάλες διαδικτυακές πλατφόρμες (VLOPs), δηλαδή εκείνες με τουλάχιστον 45 εκατομμύρια μέσους μηνιαίους ενεργούς χρήστες.	Θα είναι άμεσα εφαρμοσίμος σε ολόκληρη την ΕΕ 20 ημέρες μετά τη δημοσίευση στην Επίσημη Εφημερίδα μετά την τελική υιοθέτηση του κειμένου, ή από την 1η Ιανουαρίου 2024, όποιο είναι μεταγενέστερο.

Πίνακας 3.1. DSA & DMA: Βασικές πτυχές και διαφορές, Sérvulo

3.4.3 Digital Services Act και Digital Markets Act

Όπως προαναφέρθηκε, οι Πράξεις Digital Services Act (DSA) και Digital Markets Act (DMA)⁴⁶ παρουσιάζουν εκτεταμένη, γεωγραφικά και δικαιοδοτικά, νομική επιδραστικότητα αντίστοιχη με την ευρωπαϊκή νομοθεσία για τα Προσωπικά Δεδομένα. Αυτές έχουν ως στόχο να θέσουν περιο-

⁴³Anu Bradford, The Brussels effect, Nw. UL Rev., 2012, 107: 1, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2770634

⁴⁴Mercy A. Kuo, The Brussels Effect and China: Shaping Tech Standards - Insights from Anu Bradford, The Diplomat, 7 January 2021, <https://thediplomat.com/2021/01/the-brussels-effect-and-china-shaping-tech-standards/>

⁴⁵Η Πράξη για τις Ψηφιακές Υπηρεσίες δημοσιεύθηκε στην Επίσημη Εφημερίδα στις 27 Οκτωβρίου 2022, τέθηκε σε ισχύ στις 16 Νοεμβρίου 2022 και είναι εφαρμοστέα άμεσα σε ολόκληρη την ΕΕ από τις 17 Φεβρουαρίου 2024, <https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package>

⁴⁶Η Πράξη για τις Ψηφιακές Υπηρεσίες δημοσιεύθηκε στην Επίσημη Εφημερίδα στις 27 Οκτωβρίου 2022, τέθηκε σε ισχύ στις 16 Νοεμβρίου 2022 και είναι εφαρμοστέα άμεσα σε ολόκληρη την ΕΕ από τις 17 Φεβρουαρίου 2024, <https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package>

ρισμούς στις μεγάλες εταιρείες τεχνολογίας που δραστηριοποιούνται στην Ευρώπη⁴⁷. Χαρακτηριστική ήταν η ανακοίνωση από τον Επίτροπο Thierry Breton στις 25 Απριλίου 2023⁴⁸ ότι οι εξής 19 “πολύ μεγάλες διαδικτυακές πλατφόρμες” (VLOP) και μηχανές αναζήτησης (με περισσότερους από 45 εκατομμύρια μηνιαίους ενεργούς χρήστες στην ΕΕ) οφείλουν να συμμορφωθούν πλήρως με τις ειδικές υποχρεώσεις που τους επιβάλλει η Πράξη για τις Ψηφιακές Υπηρεσίες (DSA) με ρητή προθεσμία συμμόρφωσης τις 25 Αυγούστου 2023: AliExpress, Amazon Store, AppStore, Bing, Booking, Facebook, Google Maps, Google Play, Google Search, Google Shopping, Instagram, LinkedIn, Pinterest, Snapchat, TikTok, Twitter, Wikipedia, YouTube, Zalando.

Συνεπώς, ανεξαρτήτως της εθνικής δικαιοδοσίας στην οποία υπάγονται οι διαδικτυακές πλατφόρμες, η Digital Services Act επιβάλλει κανονισμούς σε αυτές για τη λογοδοσία και τον περιορισμό των παράνομων περιεχομένου και αγαθών στο διαδίκτυο, ενώ η Digital Markets Act θα ορίζει τους λεγόμενους gatekeepers, μεγάλους παρόχους διαδικτυακών υπηρεσιών που μονοπωλιακά ελέγχουν τμήματα της αγοράς, και θα επιβάλλει κανόνες για τη διασφάλιση δίκαιου ανταγωνισμού για την προσέγγιση των καταναλωτών. Ομοίως η Πράξη για την Τεχνητή Νοημοσύνη που συζητήθηκε καθ’ όλη τη διάρκεια του 2022 και του 2023, αναμένεται να θέσει διεθνή πρότυπα για την υπεύθυνη τεχνητή νοημοσύνη μόλις τεθεί σε ισχύ.

3.4.4 Ευρωπαϊκή στρατηγική για τα δεδομένα (A European strategy for data)

Όπως προκύπτει από τους ορισμούς της τεχνητής νοημοσύνης, τα δεδομένα αποτελούν δομικά στοιχεία για τη λειτουργία της. Παρατηρούμε επίσης ότι ολοένα και περισσότερες δημόσιες οντότητες σε όλο τον κόσμο χρησιμοποιούν την ΤΝ για να βελτιώσουν τις πολιτικές και τη διαδικασία λήψης αποφάσεων ακριβώς με βάση τα δεδομένα που διαθέτουν. Για τον λόγο αυτό στο πλαίσιο της επισκόπησης του θεσμικού πλαισίου για τη ρύθμιση της τεχνητής νοημοσύνης, που βρίσκεται διεθνώς υπό εξέλιξη, θα παρουσιάσουμε τις λύσεις που προτείνονται από την ΕΕ για μια Στρατηγική των Δεδομένων⁴⁹.

Η αιτιολογική έκθεση της πρότασης της Πράξης για την ΤΝ αναφέρει ότι η προώθηση της καινοτομίας με βάση την ΤΝ συνδέεται στενά με την Πράξη για τη Διακυβέρνηση Δεδομένων (Data Governance Act)⁵⁰, την Οδηγία για τα Ανοικτά Δεδομένα (Directive on open data and the re-use of public sector information)⁵¹ και άλλες πρωτοβουλίες στο πλαίσιο της προαναφερθείσας ενωσιακής Στρατηγικής για τα Δεδομένα, που θα καθιερώσει αξιόπιστους μηχανισμούς και υπηρεσίες για την περαιτέρω χρήση, την κοινοχρησία και τη διασύνδεση δεδομένων, που είναι απαραίτητα για την ανάπτυξη υψηλής ποιότητας μοντέλων ΤΝ που βασίζονται σε δεδομένα.

Πρόκειται για μία μετάβαση σε μέτρα που ρυθμίζουν τη χρήση και την πρόσβαση σε δεδομένα εντός της Ευρωπαϊκής Ένωσης σε όλους τους οικονομικούς τομείς και τα οποία δεν είναι “δεδομένα προσωπικού χαρακτήρα” όπως τα εννοεί ο ΓΚΠΔ αλλά κάτι διαφορετικό. Οι ρυθμίσεις και τα μέτρα για τα δεδομένα, έτσι όπως περιγράφησαν αμέσως πιο πάνω, πρέπει να συνδυαστούν περαιτέρω με την Πράξη της ΕΕ για τα Δεδομένα, που έχει στόχο να θεσπίσει κοινά πρότυπα για την περαιτέρω χρήση των δεδομένων σε όλους τους τομείς και να αντιμετωπίσει τις αθέμιτες εμπορικές πρακτικές μαζί με την έλλειψη διαθεσιμότητας δεδομένων.

⁴⁷The Digital Markets Act will enter into force today: an overview, 1 Νοεμβρίου 2022, <https://www.servulo.com/en/knowledge/The-Digital-Markets-Act-will-enter-into-force-today-an-overview/8097/>

⁴⁸More responsibility, less opacity: what it means to be a “Very Large Online Platform”, Βρυξέλλες, 25 Απριλίου 2023, https://ec.europa.eu/commission/presscorner/detail/en/STATEMENT_23_2452

⁴⁹Ευρωπαϊκή στρατηγική για τα δεδομένα, Βρυξέλλες, COM(2020) 66 final, 19 Φεβρουαρίου 2020, <https://eur-lex.europa.eu/legal-content/EL/TXT/?uri=CELEX:52020DC0066>

⁵⁰Κανονισμός (ΕΕ) 2022/868 του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου της 30ής Μαΐου 2022 σχετικά με την ευρωπαϊκή διακυβέρνηση δεδομένων και την τροποποίηση του κανονισμού (ΕΕ) 2018/1724 (πράξη για τη διακυβέρνηση δεδομένων), <https://eur-lex.europa.eu/eli/reg/2022/868/oj>

⁵¹Οδηγία (ΕΕ) 2019/1024 του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου, της 20ής Ιουνίου 2019, για τα ανοικτά δεδομένα και την περαιτέρω χρήση πληροφοριών του δημόσιου τομέα, <https://data.europa.eu/eli/dir/2019/1024/oj>

3.4.4.1 Ευρωπαϊκή διακυβέρνηση δεδομένων (European data governance)

Ειδικότερα, με τον Κανονισμό (ΕΕ) 2022/868 “Ευρωπαϊκή διακυβέρνηση δεδομένων” τίθεται το νομοθετικό πλαίσιο σε Ενωσιακό επίπεδο αφενός για την εμβάθυνση της περαιτέρω χρήσης δεδομένων του Δημοσίου Τομέα και, αφετέρου, για τη δημιουργία μιας ενιαίας αγοράς παροχής υπηρεσιών διαμεσολάβησης δεδομένων και επεξεργασίας δεδομένων για “αλτρουιστικούς σκοπούς”. Ο Κανονισμός για την ευρωπαϊκή διακυβέρνηση δεδομένων αναμένεται να ενισχύσει την ελεύθερη ροή των δεδομένων προσωπικού και μη χαρακτήρα στην Ευρωπαϊκή Ένωση και να ωθήσει την ανάπτυξη αγορών που στηρίζονται στην επεξεργασία δεδομένων⁵². Με σκοπό την ανάπτυξη μιας ενιαίας ψηφιακής αγοράς στην ΕΕ και μιας ανθρωποκεντρικής, αξιόπιστης και ασφαλούς κοινωνίας και οικονομίας βασισμένης στα δεδομένα μέσω της κοινής χρήσης ή της πρόσβασης σε δεδομένα μεταξύ των κρατών μελών ή και σε πανευρωπαϊκό επίπεδο, η Πράξη καθορίζει:

- (α) τις προϋποθέσεις για την περαιτέρω χρήση, εντός της Ένωσης, ορισμένων κατηγοριών δεδομένων που διατηρούνται από φορείς του δημόσιου τομέα,
- (β) ένα πλαίσιο γνωστοποίησης και εποπτείας για την παροχή υπηρεσιών διαμεσολάβησης δεδομένων,
- (γ) ένα πλαίσιο για την εθελοντική εγγραφή οντοτήτων που συλλέγουν και επεξεργάζονται δεδομένα που διατίθενται για αλτρουιστικούς σκοπούς, και
- (δ) ένα πλαίσιο για τη σύσταση του Ευρωπαϊκού Συμβουλίου Καινοτομίας Δεδομένων.

Ήδη από το πρώτο της άρθρο, η Πράξη ορίζει ότι οι διατάξεις της δεν θίγουν την εφαρμογή των κανόνων δικαίου του Κανονισμού (ΕΕ) 2016/679 (“ΓΚΠΔ”) και της Οδηγίας 2002/58/ΕΚ (“Οδηγία για την προστασία της ιδιωτικής ζωής στις ηλεκτρονικές επικοινωνίες”), συμπεριλαμβανομένης της εφαρμογής αυτών σε σύνολα δεδομένων με άρρηκτη σύνδεση προσωπικών και μη δεδομένων. Πρέπει, ωστόσο, να σημειωθεί ότι η Πράξη δεν δημιουργεί υποχρέωση στα κράτη μέλη για την περαιτέρω χρήση δεδομένων του δημόσιου τομέα. Επομένως, εναπόκειται σε κάθε κράτος μέλος να αποφασίζει αν τα δεδομένα καθίστανται προσβάσιμα για περαιτέρω χρήση, μεταξύ άλλων όσον αφορά τους σκοπούς και το πεδίο εφαρμογής της πρόσβασης. Οι κανόνες της Πράξης για την περαιτέρω χρήση δεδομένων του Δημοσίου Τομέα συμπληρώνουν τις διατάξεις της Οδηγίας (ΕΕ) 2019/1024⁵³ για τα Ανοικτά Δεδομένα και την περαιτέρω χρήση πληροφοριών του Δημοσίου Τομέα, θέτοντας το πλαίσιο για την ανάπτυξη μιας αγοράς δεδομένων του δημόσιου τομέα, η οποία αναμένεται να αγγίξει τα 194 δις σε οικονομική αξία μέχρι το 2030⁵⁴.

Περαιτέρω, η Πράξη θεσπίζει κανόνες για την προώθηση του αλτρουισμού των δεδομένων⁵⁵, που συνίστανται στην εθελοντική κοινοποίηση είτε προσωπικών δεδομένων με βάση τη συναίνεση των υποκειμένων των δεδομένων, είτε δεδομένων μη προσωπικού χαρακτήρα από κατόχους δεδομένων χωρίς αποζημίωση για σκοπούς γενικού συμφέροντος. Τέτοιοι σκοποί δημοσίου συμφέροντος δύνανται λόγω χάρη να αφορούν υγειονομική περίθαλψη, καταπολέμηση της κλιματικής αλλαγής, βελτίωση της κινητικότητας, διευκόλυνση της ανάπτυξης, παραγωγής και διάδοσης επίσημων στατιστικών, βελτίωση παροχής των δημόσιων υπηρεσιών, χάραξη δημόσιας πολιτικής ή σκοπούς επιστημονικής έρευνας υπέρ του γενικού συμφέροντος⁵⁶.

⁵² Ευρωπαϊκή διακυβέρνηση δεδομένων: Κανονισμός (ΕΕ) 2022/868, Ernst & Young https://www.ey.com/el_gr/tax/tax-alerts/praksi-gia-tin-diakivernisi-ton-dedomenon

⁵³ <https://eur-lex.europa.eu/legal-content/EL/TXT/?uri=CELEX:32019L1024>

⁵⁴ <https://digital-strategy.ec.europa.eu/en/policies/open-data>

⁵⁵ ΚΕΦΑΛΑΙΟ IV, Αλτρουισμός δεδομένων, Άρθρα 16 επ., Πράξη για τη διακυβέρνηση δεδομένων, <https://data.europa.eu/eli/reg/2022/868/oj>

⁵⁶ Άρθρο 2, ορισμός 16 “Αλτρουισμός δεδομένων”, Πράξη για τη διακυβέρνηση δεδομένων, <http://data.europa.eu/eli/reg/2022/868/oj>

3.4.4.2 Κανονισμός για τα Δεδομένα (Data Act)

Επιπλέον στις 22 Δεκεμβρίου 2023 δημοσιεύθηκε στην Επίσημη Εφημερίδα της ΕΕ ο Κανονισμός για τη Δίκαιη Πρόσβαση σε Δεδομένα και τη δίκαιη χρήση τους (Κανονισμός για τα Δεδομένα)⁵⁷, που αποσκοπεί στην τόνωση της καινοτομίας με την εξάλειψη των φραγμών στα βιομηχανικά δεδομένα. Αφειρητά υπήρξε η διαπίστωση ότι ο όγκος των δεδομένων που παράγονται από ανθρώπους και μηχανές αυξάνεται εκθετικά και αποτελεί κρίσιμο παράγοντα καινοτομίας για τις επιχειρήσεις και τις δημόσιες αρχές (π.χ. για τη διαμόρφωση έξυπνων πόλεων). Η εν λόγω νομοθεσία θα επιτρέψει στους χρήστες την πρόσβαση στα δεδομένα που παράγουν, καθώς το 80% των βιομηχανικών δεδομένων που συλλέγονται δεν χρησιμοποιούνται ποτέ, σύμφωνα με την Ευρωπαϊκή Επιτροπή.

Οι νέοι κανόνες αναμένεται να δημιουργήσουν επιπλέον ΑΕγχΠ ύψους 270 δισ. ευρώ για τα κράτη μέλη της ΕΕ έως το 2028⁵⁸, αντιμετωπίζοντας τα νομικά, οικονομικά και τεχνικά ζητήματα που οδηγούν σε υποχρησιμοποίηση των δεδομένων. Οι καταναλωτές και οι επιχειρήσεις παράγουν δεδομένα χρησιμοποιώντας προϊόντα και υπηρεσίες. Με την πράξη για τα δεδομένα, θα επωφεληθούν από:

- φθηνότερες τιμές για υπηρεσίες δευτερογενούς αγοράς και αποκατάσταση των συνδεδεμένων αντικειμένων τους,
- νέες ευκαιρίες χρήσης υπηρεσιών που βασίζονται στην πρόσβαση στα εν λόγω τα δεδομένα,
- καλύτερη πρόσβαση σε δεδομένα που συλλέγονται ή παράγονται από συσκευή.

Η νέα ευρωπαϊκή νομοθεσία θεσπίζει κανόνες για την ανταλλαγή δεδομένων που παράγονται μέσω της χρήσης συνδεδεμένων προϊόντων ή συναφών υπηρεσιών (π.χ. το διαδίκτυο των πραγμάτων, βιομηχανικά μηχανήματα) και επιτρέπει στους χρήστες να έχουν πρόσβαση στα δεδομένα που παράγουν. Θα συμβάλει στην ανάπτυξη νέων υπηρεσιών, ιδίως στον τομέα της τεχνητής νοημοσύνης, όπου απαιτούνται τεράστιες ποσότητες δεδομένων για την “εκπαίδευση” αλγορίθμων. Έχει επίσης ως στόχο να καταστήσει φθηνότερες τις υπηρεσίες που παρέχονται μετά την πώληση, αλλά και τις επισκευές συνδεδεμένων συσκευών. Σε εξαιρετικές περιστάσεις ή καταστάσεις έκτακτης ανάγκης, όπως πλημμύρες και πυρκαγιές, οι φορείς του δημόσιου τομέα μπορούν να έχουν πρόσβαση και να χρησιμοποιούν δεδομένα που κατέχει ο ιδιωτικός τομέας.

Ως προς την προστασία του εμπορικού απορρήτου και την πρόληψη παράνομων μεταφορών δεδομένων, το Ευρωπαϊκό Κοινοβούλιο σε δελτίο τύπου⁵⁹ αναφέρει ότι η νομοθεσία περιέχει έναν σαφή ορισμό του εμπορικού απορρήτου και των κατόχων σχετικών δικαιωμάτων, ώστε να προληφθεί η παράνομη μεταφορά και διαρροή δεδομένων σε χώρες με ασθενέστερους κανόνες προστασίας τους. Επίσης θα αποτρέψει εμπορικούς ανταγωνιστές σε συγκεκριμένους τομείς από το να εκμεταλλεύονται την πρόσβαση σε δεδομένα για να αντιγράψουν τις υπηρεσίες ή συσκευές των αντιπάλων τους.

Επίσης η νέα νομοθεσία διευκολύνει τη δυνατότητα αλλαγής παρόχων υπολογιστικού νέφους (εταιρείες που προσφέρουν υπηρεσίες δικτύου, υποδομών, ή επιχειρηματικών εφαρμογών) και διασφαλίζει ότι οι εταιρείες αυτές δε θα μπορούν να μεταφέρουν παράνομα δεδομένα εκτός συνόρων. Διασφαλίστηκε ακόμη ότι οι πελάτες υπηρεσιών υπολογιστικού νέφους θα έχουν την δυνατότητα να διαπραγματεύονται συμβάσεις και να αποφεύγουν να “εγκλωβίζονται” από έναν συγκεκριμένο πάροχο.

⁵⁷Κανονισμός (ΕΕ) 2023/2854 του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου, της 13ης Δεκεμβρίου 2023, για εναρμονισμένους κανόνες σχετικά με τη δίκαιη πρόσβαση σε δεδομένα και τη δίκαιη χρήση τους, <http://data.europa.eu/eli/reg/2023/2854/oj>

⁵⁸https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_el

⁵⁹<https://www.europarl.europa.eu/news/el/press-room/20231106IPR09025/data-act-kanones-gia-kaluteri-prosvasi-kai-chrisi-dedomenon>

Συγκεκριμένα, η Πράξη για τα Δεδομένα περιλαμβάνει τα εξής⁶⁰:

- Μέτρα που επιτρέπουν στους χρήστες συνδεδεμένων συσκευών να έχουν πρόσβαση στα δεδομένα που παράγονται από τις εν λόγω συσκευές και από υπηρεσίες που σχετίζονται με τις εν λόγω συσκευές. Οι χρήστες θα μπορούν να μοιράζονται τα εν λόγω δεδομένα με τρίτους, ενισχύοντας τις υπηρεσίες δευτερογενούς αγοράς και την καινοτομία. Ταυτόχρονα, οι κατασκευαστές εξακολουθούν να ενθαρρύνονται ώστε να επενδύουν στην παραγωγή δεδομένων υψηλής ποιότητας, ενώ το εμπορικό απόρρητο τους εξακολουθεί να προστατεύεται.
- Μέτρα για την παροχή προστασίας από καταχρηστικές συμβατικές ρήτρες που επιβάλλονται μονομερώς.
- Μηχανισμούς με τους οποίους οι φορείς του δημόσιου τομέα θα μπορούν να έχουν πρόσβαση και να χρησιμοποιούν δεδομένα που βρίσκονται στην κατοχή του ιδιωτικού τομέα σε περιπτώσεις έκτακτης ανάγκης σε τομείς δημόσιου συμφέροντος, όπως πλημμύρες και δασικές πυρκαγιές, ή κατά την εφαρμογή νομικής εντολής, όταν τα απαιτούμενα δεδομένα δεν είναι άμεσα διαθέσιμα με άλλα μέσα.
- Νέους κανόνες που παρέχουν στους πελάτες την ελευθερία να αλλάζουν παρόχους υπηρεσιών επεξεργασίας δεδομένων υπολογιστικού νέφους.
- Μέτρα για την προώθηση της ανάπτυξης προτύπων διαλειτουργικότητας για την κοινοχρησία και την επεξεργασία δεδομένων, σύμφωνα με τη στρατηγική της ΕΕ για την τυποποίηση.

3.5 Διαβούλευση και πιλοτικά εργαλεία που προηγήθηκαν της πρότασης της Πράξης για την TN

Πριν από την παρουσίαση του κειμένου του κανονισμού της Πράξης για την TN, είναι σημαντικό να ανιχνεύσουμε τον σημαντικό όγκο εργασίας που προηγήθηκε από την Ευρωπαϊκή Ένωση προκειμένου να καταλήξει σε αυτό, μέσα από μια συμμετοχική διαδικασία διαβούλευσης και εμπειρογνωμοσύνης.

Το 2019, ακολουθώντας το περίγραμμα της Ευρωπαϊκής Επιτροπής για την TN στην Ευρώπη, συνεργάστηκαν δύο συμβουλευτικά όργανα για να δημοσιεύσουν τις κατευθυντήριες γραμμές δεοντολογίας για την αξιόπιστη TN. Η Ομάδα Εμπειρογνομόνων Υψηλού Επιπέδου (AI HLEG)⁶¹ για την Τεχνητή Νοημοσύνη συνέταξε τις Κατευθυντήριες Γραμμές⁶², αφού συμβουλευτήκε τα μέλη της Ευρωπαϊκής Συμμαχίας⁶³ για την Τεχνητή Νοημοσύνη, η οποία είναι ένα φόρουμ πολλών ενδιαφερομένων μερών που δημιουργήθηκε για την παροχή ανατροφοδότησης σχετικά με τις ρυθμιστικές πρωτοβουλίες που αφορούν την Τεχνητή Νοημοσύνη. Οι Κατευθυντήριες Γραμμές καθόρισαν επτά βασικές απαιτήσεις που πρέπει να πληρούν τα συστήματα τεχνητής νοημοσύνης για να θεωρούνται αξιόπιστα: ανθρώπινη δράση και εποπτεία, τεχνική ευρωστία και ασφάλεια, προστασία της ιδιωτικής ζωής και διακυβέρνηση δεδομένων, διαφάνεια, ποικιλομορφία, μη διάκριση και δικαιοσύνη, κοινωνική και περιβαλλοντική ευημερία και λογοδοσία.

Με βάση την ανατροφοδότηση που έλαβε, η ομάδα AI HLEG παρουσίασε τον τελικό Κατάλογο Αξιολόγησης για Αξιόπιστη Τεχνητή Νοημοσύνη (ALTAI) τον Ιούλιο του 2020. Ο ALTAI είναι ένα πρακτικό εργαλείο που μεταφράζει τις κατευθυντήριες γραμμές δεοντολογίας σε έναν

⁶⁰https://ec.europa.eu/commission/presscorner/detail/el/ip_23_3491

⁶¹High-level expert group on artificial intelligence, <https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai>

⁶²Ethics guidelines for trustworthy AI, <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

⁶³The European AI Alliance, <https://digital-strategy.ec.europa.eu/en/policies/european-ai-alliance>

προσιτό και δυναμικό κατάλογο ελέγχου με αυτοαξιολόγηση⁶⁴. Ο κατάλογος ελέγχου⁶⁵ μπορεί να χρησιμοποιηθεί από τους προγραμματιστές και τους φορείς ανάπτυξης ΤΝ που θέλουν να εφαρμόσουν τις βασικές απαιτήσεις στην πράξη. Αυτός ο νέος κατάλογος⁶⁶ είναι επίσης διαθέσιμος ως διαδικτυακό εργαλείο⁶⁷.

Οι δεοντολογικές κατευθυντήριες γραμμές (και το εργαλείο αξιολόγησης που δημιουργήθηκε για την εφαρμογή τους των κατευθυντήριων γραμμών) συνέβαλαν στην πλαισίωση των συζητήσεων και στη διάρθρωση της συζήτησης για τις επόμενες φάσεις της νομοθετικής δράσης. Τον Φεβρουάριο του 2020, η Ευρωπαϊκή Επιτροπή βασίστηκε σε αυτές τις Κατευθυντήριες Γραμμές για τη δημοσίευση της Λευκής Βίβλου της με τίτλο “Για την τεχνητή νοημοσύνη: Μια ευρωπαϊκή προσέγγιση για την αριστεία και την εμπιστοσύνη”. Η Λευκή Βίβλος ανακοίνωσε την επικείμενη κανονιστική δράση και παρουσίασε τα βασικά στοιχεία του μελλοντικού πλαισίου. Μεταξύ αυτών των βασικών στοιχείων ήταν η προσέγγιση βάσει κινδύνου, η οποία πρότεινε ότι θα πρέπει να επιβληθούν υποχρεωτικές νομικές απαιτήσεις -που προέρχονται από τις ηθικές αρχές- σε συστήματα τεχνητής νοημοσύνης υψηλού κινδύνου.

Τη Λευκή Βίβλο ακολούθησε διαδικασία δημόσιας διαβούλευσης στην οποία συμμετείχαν 1.200 ενδιαφερόμενοι από διάφορα υπόβαθρα - πολίτες, ακαδημαϊκοί, κράτη μέλη της ΕΕ, η κοινωνία των πολιτών, καθώς και επιχειρήσεις και βιομηχανίες. Αυτός ο γύρος διαβούλευσης αποσκοπούσε στο να επηρεάσει τη σύνταξη του νόμου για την τεχνητή νοημοσύνη. Τα υποβληθέντα σχόλια ώθησαν την Ευρωπαϊκή Επιτροπή να εγκαταλείψει τη σχεδιαζόμενη τομεακή προσέγγιση υπέρ μιας ευρύτερης και απλούστερης προσέγγισης: Σύμφωνα με την προτεινόμενη πράξη, όλες οι εφαρμογές της ΤΝ υψηλού κινδύνου, όπου και αν εμφανίζονται πρέπει να συμμορφώνονται με αυξημένες υποχρεώσεις- εξαιρουμένων των χρήσεων αρμοδιότητας του στρατού και ορισμένων χρήσεων από τον eu-LISA⁶⁸, τον Ευρωπαϊκό Οργανισμό για τη λειτουργική διαχείριση συστημάτων ΤΠ μεγάλης κλίμακας στον χώρο ελευθερίας, ασφάλειας και δικαιοσύνης τα οποία αποτελούν βασικά μέσα για την εφαρμογή των πολιτικών ασύλου, διαχείρισης των συνόρων και μετανάστευσης της ΕΕ.

3.6 Πράξη για την ΤΝ: Συμφραζόμενα, ιστορικό και ρύθμιση

3.6.1 Συμφραζόμενα

Το σχέδιο κανονισμού που παρουσίασε η Ευρωπαϊκή Επιτροπή τον Απρίλιο του 2021 αποτελεί βασικό στοιχείο της πολιτικής της ΕΕ για την προώθηση της ανάπτυξης και της υιοθέτησης ασφαλών και νόμιμων εφαρμογών ΤΝ στην ενιαία αγορά, που σέβονται τα θεμελιώδη δικαιώματα. Τον Δεκέμβριο 2022 το Συμβούλιο της Ευρωπαϊκής Ένωσης καθόρισε την κοινή του θέση (“γενική προσέγγιση”) σχετικά με την Πράξη για την Τεχνητή Νοημοσύνη με στόχο να διασφαλιστεί ότι τα συστήματα ΤΝ που διατίθενται στην αγορά της ΕΕ και χρησιμοποιούνται στην Ένωση είναι ασφαλή και τηρούν την ισχύουσα νομοθεσία για τα θεμελιώδη δικαιώματα και τις αξίες της Ένωσης⁶⁹. Με την έγκριση της γενικής προσέγγισης τον Δεκέμβριο 2022, το Συμβούλιο άρχισε διαπραγματεύσεις με το Ευρωπαϊκό Κοινοβούλιο (“τριμερείς διάλογοι”) και το Κοινοβούλιο κα-

⁶⁴Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment, <https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment>

⁶⁵<https://futurium.ec.europa.eu/en/european-ai-alliance/pages/welcome-altai-portal>

⁶⁶Download the Assessment List for Trustworthy Artificial Intelligence (ALTAI), https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=68342

⁶⁷<https://altai.insight-centre.org/>

⁶⁸<https://www.eulisa.europa.eu/>

⁶⁹Πράξη για την τεχνητή νοημοσύνη: Το Συμβούλιο ζητεί την προώθηση ασφαλών συστημάτων ΤΝ που σέβονται τα θεμελιώδη δικαιώματα, Βρυξέλλες, 6 Δεκεμβρίου 2022, <https://www.consilium.europa.eu/el/press/press-releases/2022/12/06/artificial-intelligence-act-council-calls-for-promoting-safe-ai-that-respects-fundamental-rights/>

θόρισε τη δική του θέση για την επίτευξη συμφωνίας επί του προτεινόμενου κανονισμού.

Παράλληλα, το ίδιο διάστημα έκανε την εμφάνισή του το GPT-4 το οποίο προκάλεσε αναθεώρηση των ευκαιριών και των απειλών που παρουσιάζει η χρήση της τεχνητής νοημοσύνης και κινητοποίησε τους οργανισμούς ρύθμισης και τους φορείς εξουσίας να προβληματιστούν για την οριοθέτηση και την εποπτεία της ΤΝ. Αυτό είχε ως αποτέλεσμα να συσπειρώσει όλους τους εμπλεκόμενους από τη δημόσια και την ιδιωτική σφαίρα, ώστε να συνεργαστούν για τη μεθοδική παρακολούθηση των εξελίξεων⁷⁰.

Αξιοπρόσεκτο είναι ότι το Ευρωπαϊκό Κοινοβούλιο έκανε ρητή αναφορά στο GPT⁷¹ ως παράδειγμα μοντέλου παραγωγικής ΤΝ (Generative AI) που χρήζει εξατομικευμένων ρυθμίσεων με πρόσθετες υποχρεώσεις διαφάνειας. Συγκεκριμένα αυτό αναφέρεται στο δελτίο τύπου⁷² που εξέδωσε το Ευρωπαϊκό Κοινοβούλιο στις 11 Μαΐου 2023, με αφορμή την έγκριση διαπραγματευτικής εντολής από ορισμένες επιτροπές του σχετικά με την Πράξη για την Τεχνητή Νοημοσύνη. Ενδεικτική επίσης είναι στενή παρακολούθηση των εξελίξεων από τον διεθνή ηλεκτρονικό τύπο καθώς και η άμεση ανταπόκριση από τις κυβερνήσεις διεθνώς και από την ακαδημαϊκή κοινότητα, αφού στο Digital Bridge του Politico⁷³ πληροφορούμαστε ότι η Κίνα ήδη πρότεινε νέους κανόνες για την παραγωγική τεχνητή νοημοσύνη (Generative AI) στις 11 Απριλίου 2023. Εκτός αυτού υπάρχει παραπομπή στη μετάφραση των ανωτέρω κινεζικών μέτρων⁷⁴, την οποία δημοσίευσε το ερευνητικό πρόγραμμα DigiChina του Ινστιτούτου Διεθνών Σχέσεων Freeman Spogli του Πανεπιστημίου Stanford, στις 12 Απριλίου 2023.

Μέσα από την ακόλουθη συνοπτική παρουσίαση μπορούμε να δούμε ποια είναι η νομοθετική προσέγγιση της ΕΕ για την τεχνητή νοημοσύνη και για τη ρύθμισή της. Στις 21 Απριλίου 2021 η Ευρωπαϊκή Επιτροπή παρουσίασε το πρώτο νομικό πλαίσιο για την τεχνητή νοημοσύνη, την Πράξη για την Τεχνητή Νοημοσύνη. Η Πράξη για την ΤΝ αντιμετωπίζει τους κινδύνους που απορρέουν από τις διάφορες χρήσεις των συστημάτων τεχνητής νοημοσύνης και αποσκοπεί στην προώθηση της καινοτομίας στον τομέα της τεχνητής νοημοσύνης. Αυτή η απαρίθμηση των τεχνικών τεχνητής νοημοσύνης που εμπίπτουν στο ουσιαστικό πεδίο εφαρμογής του προτεινόμενου κανονισμού προκάλεσε κάποια αναστάτωση. Για παράδειγμα, ορισμένοι παρατηρητές ισχυρίστηκαν ότι η ΕΕ “προτείνει να ρυθμίσει τη χρήση της μπεϋζιανής εκτίμησης”, μιας πρωτίστως στατιστικής μεθόδου, για να αποδοκιμάσουν την υπερβολική έκταση του προτεινόμενου κανονισμού⁷⁵.

Ενώ οι επικρίσεις για την ευρύτητα του ορισμού της πράξης δεν είναι εντελώς αβάσιμες, διότι μπορεί να δημιουργήσει αβεβαιότητα στους προγραμματιστές συστημάτων τεχνητής νοημοσύνης, αξίζει να σημειωθεί ότι μια συγκεκριμένη τεχνολογία που εμπίπτει στο πεδίο εφαρμογής της πρότασης δεν σημαίνει απαραίτητα ότι θα υπόκειται σε νέες νομικές υποχρεώσεις. Μόνο ορισμένα συστήματα τεχνητής νοημοσύνης που ενέχουν αυξημένο επίπεδο κινδύνου σύμφωνα με τα κριτήρια της Επιτροπής θα υπόκεινται σε νομικές απαιτήσεις. Για παράδειγμα, τα εργαλεία φιλτραρίσματος ανεπιθύμητης αλληλογραφίας που βασίζονται σε μπεϋζιανά μοντέλα δεν θα υπόκεινται σε νέες νομικές απαιτήσεις, ενώ όταν χρησιμοποιούνται για την ταξινόμηση ασθενών στα επείγοντα

⁷⁰ChatGPT broke the EU plan to regulate AI, 3 Μαρτίου 2023, <https://www.politico.eu/article/eu-plan-regulate-chatgpt-openai-artificial-intelligence-act/>

⁷¹Definitions, σελ. 54612, G. Yenduri et al., “GPT (Generative Pre-Trained Transformer)— A Comprehensive Review on Enabling Technologies, Potential Applications, Emerging Challenges, and Future Directions,” in IEEE Access, vol. 12, pp. 54608-54649, 2024, doi: 10.1109/ACCESS.2024.3389497.

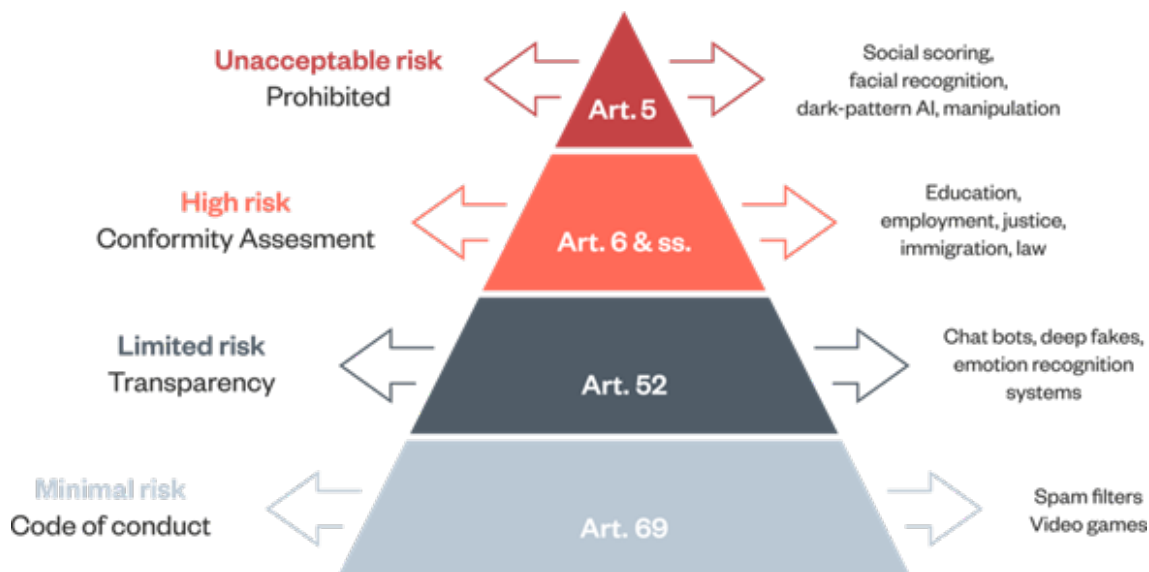
⁷²AI Act: a step closer to the first rules on Artificial Intelligence, 11 Μαΐου 2023, <https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence>

⁷³TTC primer — EU-India summit — AI Act gets real, 11 Μαΐου 2023, <https://www.politico.eu/newsletter/digital-bridge/ttc-primer-eu-india-summit-ai-act-gets-real>

⁷⁴Translation: Measures for the Management of Generative Artificial Intelligence Services (Draft for Comment), Απρίλιος 2023, <https://digichina.stanford.edu/work/translation-measures-for-the-management-of-generative-artificial-intelligence-services-draft-for-comment-ap>

⁷⁵Eve Gaumond, Artificial Intelligence Act: What Is the European Approach for AI?, 4 June 2021, <https://www.lawfareblog.com/artificial-intelligence-act-what-european-approach-ai>

περιστατικά πιθανότατα θα υπόκεινται.



Εικόνα 3.1. A ‘risk-based’ approach

<https://www.adalovelaceinstitute.org/resource/eu-ai-act-explainer/>

Αυτές οι διακρίσεις καταδεικνύουν την προσέγγιση της Ευρωπαϊκής Επιτροπής με βάση τον κίνδυνο. Μια τέτοια διαβάθμιση των κινδύνων είναι η κεντρική ιδέα του προτεινόμενου κανονισμού. Αυτή η κλιμακωτή προσέγγιση, η οποία προτάθηκε για πρώτη φορά στην Ευρώπη από τη Γερμανία⁷⁶, σκιαγραφείται επίσης στο νόμο των Ηνωμένων Πολιτειών για την λογοδοσία των αλγορίθμων του 2019⁷⁷ και στην καναδική οδηγία για την αυτοματοποιημένη λήψη αποφάσεων⁷⁸. Η εμφάνιση παρόμοιων προσεγγίσεων για τη ρύθμιση της τεχνητής νοημοσύνης σε άλλες χώρες μπορεί να υποδηλώνει ότι η ευρωπαϊκή ρύθμιση θα αποτελέσει τον κανόνα στη διεθνή σκηνή.

3.6.2 Ιστορικό

Συνοπτικά η εξέλιξη της επεξεργασίας της πρότασης για την Πράξη της Τεχνητής Νοημοσύνης αποτυπώθηκε ως εξής από το διαδικτυακό εργαλείο του Ευρωπαϊκού Κοινοβουλίου “Legislative train”⁷⁹ (που αποτυπώνει μεταφορικά υπό τη μορφή τρένου και βαγονιών την πρόοδο του νομοθετικού έργου και ενημερώνεται κάθε μήνα): Η υιοθέτηση συστημάτων Τεχνητής Νοημοσύνης έχει μεγάλες δυνατότητες να αποφέρει κοινωνικά οφέλη, οικονομική ανάπτυξη και να ενισχύσει την καινοτομία της ΕΕ και την παγκόσμια ανταγωνιστικότητα. Ταυτόχρονα, είναι κοινώς αποδεκτό ότι τα ειδικά χαρακτηριστικά ορισμένων συστημάτων τεχνητής νοημοσύνης εγείρουν ορισμένες ανησυχίες, ιδίως όσον αφορά την ασφάλεια και την προστασία των θεμελιωδών δικαιωμάτων. Ξεκίνησε ένας προβληματισμός στην ΕΕ σχετικά με τον τρόπο αντιμετώπισης αυτών των ανησυχιών. Η Ευρωπαϊκή Επιτροπή δημοσίευσε τον Φεβρουάριο του 2020 τη Λευκή Βίβλο για την Τεχνητή Νοημοσύνη⁸⁰ και πρότεινε τη δημιουργία ενός ευρωπαϊκού ρυθμιστικού πλαισίου για αξιόπιστη

⁷⁶https://www.bmjbv.de/SharedDocs/Downloads/DE/Themen/Fokusthemen/Gutachten_DEK_EN_lang.pdf

⁷⁷<https://www.congress.gov/bill/116th-congress/house-bill/2231/text>

⁷⁸<https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592>

⁷⁹20 Μαρτίου 2023, <https://www.europarl.europa.eu/legislative-train/spotlight-JD22/file-regulation-on-artificial-intelligence?sid=6701>

⁸⁰ΛΕΥΚΗ ΒΙΒΛΟΣ Τεχνητή Νοημοσύνη - Ευρωπαϊκή προσέγγιση της αριστείας και της εμπιστοσύνης, COM(2020)65 final, Βρυξέλλες, 19 Φεβρουαρίου 2020, <https://op.europa.eu/el/publication-detail/-/publication/ac957f13-53c6-11ea-aece-01aa75ed71a1>

τεχνητή νοημοσύνη. Το Ευρωπαϊκό Κοινοβούλιο ενέκρινε τον Οκτώβριο του 2020 τρία νομοθετικά ψηφίσματα σχετικά με την τεχνητή νοημοσύνη που αφορούν α) τη δεοντολογία⁸¹, β) την αστική ευθύνη⁸² και γ) την πνευματική ιδιοκτησία⁸³ και ζήτησε από την Επιτροπή να δημιουργήσει ένα ολοκληρωμένο ευρωπαϊκό νομικό πλαίσιο ηθικών αρχών για την ανάπτυξη και τη χρήση της τεχνητής νοημοσύνης, της ρομποτικής και συναφών τεχνολογιών.

3.6.2.1 Η θέση της Ευρωπαϊκής Επιτροπής

Σε αυτό το πλαίσιο, η Ευρωπαϊκή Επιτροπή παρουσίασε μια πρόταση για έναν νέο κανονισμό περί τεχνητής νοημοσύνης, την Πράξη για την Τεχνητή Νοημοσύνη τον Απρίλιο 2021, που περιλαμβάνει έναν τεχνολογικά ουδέτερο ορισμό των συστημάτων τεχνητής νοημοσύνης και βασίζεται σε τέσσερα επίπεδα κινδύνων που μπορεί να προκαλέσουν:

1. Απαγορευμένες πρακτικές ΤΝ (μη αποδεκτή χρήση). Οι επιβλαβείς χρήσεις της τεχνητής νοημοσύνης που αντιβαίνουν στις αξίες της ΕΕ (όπως η κοινωνική βαθμολογία από τις κυβερνήσεις) θα απαγορευθούν λόγω του μη αποδεκτού βαθμού κινδύνου που δημιουργούν.
2. ΑΙ υψηλού κινδύνου. Ορισμένα συστήματα τεχνητής νοημοσύνης που δημιουργούν δυσμενείς επιπτώσεις στην ασφάλεια των ανθρώπων ή στα θεμελιώδη δικαιώματά τους θεωρούνται υψηλού κινδύνου. Προκειμένου να διασφαλιστεί η εμπιστοσύνη και το σταθερό υψηλό επίπεδο προστασίας της ασφάλειας και των θεμελιωδών δικαιωμάτων, μια σειρά υποχρεωτικών απαιτήσεων (συμπεριλαμβανομένης της αξιολόγησης της συμμόρφωσης) θα ισχύουν για όλα τα συστήματα υψηλού κινδύνου.
3. Τεχνητή νοημοσύνη περιορισμένου κινδύνου. Ορισμένα συστήματα τεχνητής νοημοσύνης θα υπόκεινται σε περιορισμένο σύνολο υποχρεώσεων (π.χ. διαφάνεια).
4. Τεχνητή νοημοσύνη ελάχιστου κινδύνου. Όλα τα άλλα συστήματα τεχνητής νοημοσύνης μπορούν να αναπτυχθούν και να χρησιμοποιηθούν στην ΕΕ χωρίς πρόσθετες νομικές υποχρεώσεις από την ισχύουσα νομοθεσία.

3.6.2.2 Παραδείγματα βαθμών κινδύνου

Παραδείγματα των ανωτέρω αντλούμε από άρθρο της ιστοσελίδας του Ευρωπαϊκού Κοινοβουλίου⁸⁴ όπως εκτίθενται παρακάτω.

1. Παραδείγματα μη αποδεκτού κινδύνου:
 - γνωστική συμπεριφορική χειραγώγηση ατόμων ή συγκεκριμένων ευάλωτων ομάδων όπως παιχνίδια που ενεργοποιούνται με φωνή που ενθαρρύνουν επικίνδυνη συμπεριφορά στα παιδιά
 - κοινωνική βαθμολογία όπως η ταξινόμηση ατόμων με βάση τη συμπεριφορά, την κοινωνικοοικονομική κατάσταση, τα προσωπικά χαρακτηριστικά.
2. Παραδείγματα υψηλού κινδύνου:
 - Συστήματα τεχνητής νοημοσύνης που χρησιμοποιούνται σε προϊόντα που εμπίπτουν στη νομοθεσία της ΕΕ για την ασφάλεια των προϊόντων που περιλαμβάνει παιχνίδια, αεροπορία, αυτοκίνητα, ιατρικές συσκευές και ανελκυστήρες.

⁸¹<https://eur-lex.europa.eu/legal-content/EL/TXT/?uri=CELEX:52020IP0275>

⁸²<https://eur-lex.europa.eu/legal-content/EL/TXT/?uri=CELEX:52020IP0276>

⁸³<https://eur-lex.europa.eu/legal-content/EL/TXT/?uri=CELEX:52020IP0277>

⁸⁴<https://www.europarl.europa.eu/news/el/headlines/society/20230601ST093804/praxi-technitis-noimosunis-tis-ee-protos-kanonismos-gia-tin-techniti-noimosuni>

- Συστήματα τεχνητής νοημοσύνης που εμπίπτουν σε συγκεκριμένους τομείς που θα πρέπει να καταχωρηθούν σε μια βάση δεδομένων της ΕΕ (αναλυτικά στο ΠΑΡΑΡΤΗΜΑ ΙΙΙ της Πράξης για την ΤΝ) όπως για παράδειγμα:
 - (α) διαχείριση και λειτουργία υποδομών ζωτικής σημασίας (οδική κυκλοφορία και παροχή νερού, φυσικού αερίου, θέρμανσης και ηλεκτρικής ενέργειας)
 - (β) εκπαίδευση και επαγγελματική κατάρτιση (πρόσβαση ή τοποθέτηση ή αξιολόγηση φυσικών προσώπων σε ιδρύματα εκπαίδευσης και επαγγελματικής κατάρτισης)
 - (γ) απασχόληση, διαχείριση εργαζομένων και πρόσβαση στην αυτοαπασχόληση (δικασίες πρόσληψης ή αξιολόγησης εργαζομένων).
3. Παραδείγματα περιορισμένου κινδύνου:
Οι χρήστες θα πρέπει να ενημερώνονται όταν αλληλεπιδρούν με την τεχνητή νοημοσύνη. Αυτό περιλαμβάνει συστήματα ΑΙ που δημιουργούν ή χειρίζονται περιεχόμενο εικόνας, ήχου ή βίντεο (π.χ. deepfakes).
 4. Παράδειγμα ελάχιστου κινδύνου:
Το φίλτρο στα εισερχόμενα της ηλεκτρονικής αλληλογραφίας.

3.6.2.3 Η θέση του Συμβουλίου της ΕΕ

Το Συμβούλιο είχε ήδη εγκρίνει την κοινή του θέση (“γενική προσέγγιση”) σχετικά με τον νόμο ΑΙ στις 6 Δεκεμβρίου 2022⁸⁵. Το κείμενο του Συμβουλίου, μεταξύ άλλων:

1. Περιορίζει τον ορισμό σε συστήματα που αναπτύχθηκαν μέσω προσεγγίσεων μηχανικής μάθησης και προσεγγίσεων που βασίζονται στη λογική και στη γνώση,
2. Επεκτείνει σε ιδιωτικούς φορείς την απαγόρευση χρήσης τεχνητής νοημοσύνης για κοινωνική βαθμολόγηση,
3. Προσθέτει ένα οριζόντιο επίπεδο πάνω από την ταξινόμηση υψηλού κινδύνου για να διασφαλίσει ότι τα συστήματα τεχνητής νοημοσύνης που δεν είναι πιθανό να προκαλέσουν σοβαρές παραβιάσεις θεμελιωδών δικαιωμάτων ή άλλους σημαντικούς κινδύνους δεν καταγράφονται,
4. Διευκρινίζει τις απαιτήσεις για συστήματα τεχνητής νοημοσύνης υψηλού κινδύνου,
5. Προσθέτει νέες διατάξεις για να ληφθούν υπόψη καταστάσεις όπου τα συστήματα τεχνητής νοημοσύνης μπορούν να χρησιμοποιηθούν για πολλούς διαφορετικούς σκοπούς (τεχνητή νοημοσύνη γενικού σκοπού),
6. Αποσαφηνίζει το πεδίο εφαρμογής της πράξης για την τεχνητή νοημοσύνη (π.χ. ρητό αποκλεισμό της εθνικής ασφάλειας, της άμυνας και των στρατιωτικών σκοπών από το πεδίο εφαρμογής του νόμου για την τεχνητή νοημοσύνη) και τις διατάξεις που σχετίζονται με την επιβολή του νόμου,
7. Απλοποιεί το πλαίσιο συμμόρφωσης για τον νόμο ΑΙ,
8. Προσθέτει νέες διατάξεις για να αυξήσει τη διαφάνεια και να επιτρέπει τα παράπονα των χρηστών,

⁸⁵Πράξη για την τεχνητή νοημοσύνη: Το Συμβούλιο ζητεί την προώθηση ασφαλών συστημάτων ΤΝ που σέβονται τα θεμελιώδη δικαιώματα, 6 Δεκεμβρίου 2022, <https://www.consilium.europa.eu/el/press/press-releases/2022/12/06/artificial-intelligence-act-council-calls-for-promoting-safe-ai-that-respects-fundamental-rights/>

9. Τροποποιεί ουσιαστικά τις διατάξεις που αφορούν μέτρα στήριξης της καινοτομίας (π.χ. ρυθμιστικά δοκιμαστήρια τεχνητής νοημοσύνης-sandboxes).

3.6.2.4 Η Θέση του Κοινοβουλίου της ΕΕ

Στο Κοινοβούλιο, οι συζητήσεις διεξήχθησαν υπό την καθοδήγηση της Επιτροπής Εσωτερικής Αγοράς και Προστασίας των Καταναλωτών και της Επιτροπής Πολιτικών Ελευθεριών, Δικαιοσύνης και Εσωτερικών Υποθέσεων στο πλαίσιο μιας διαδικασίας κοινής επιτροπής. Το Κοινοβούλιο ενέκρινε τη διαπραγματευτική του θέση (499 ψήφοι υπέρ, 28 κατά και 93 αποχές) τον Ιούνιο του 2023 με ουσιαστικές τροποποιήσεις στο κείμενο της Επιτροπής, μεταξύ άλλων:

1. Οι ευρωβουλευτές τροποποίησαν τον ορισμό των συστημάτων τεχνητής νοημοσύνης για να τον ευθυγραμμίσουν με τον ορισμό που συμφωνήθηκε από τον Οργανισμό Οικονομικής Συνεργασίας και Ανάπτυξης (ΟΟΣΑ).
2. Οι ευρωβουλευτές τροποποίησαν ουσιαστικά τον κατάλογο των συστημάτων τεχνητής νοημοσύνης που απαγορεύονται στην ΕΕ. Το Κοινοβούλιο επιθυμεί να απαγορεύσει τη χρήση συστημάτων βιομετρικής ταυτοποίησης στην ΕΕ τόσο για χρήση σε πραγματικό χρόνο όσο και για εκ των υστέρων χρήση (εκτός από περιπτώσεις σοβαρών εγκλημάτων και προδικαστικής άδειας για εκ των υστέρων χρήση) και όχι μόνο για χρήση σε πραγματικό χρόνο, όπως πρότεινε η Επιτροπή.
3. Επιπλέον, το Κοινοβούλιο επιθυμεί να απαγορεύσει όλα τα συστήματα βιομετρικής κατηγοριοποίησης που χρησιμοποιούν ευαίσθητα χαρακτηριστικά, τα συστήματα προληπτικής αστυνόμευσης, τα συστήματα αναγνώρισης συναισθημάτων και τα συστήματα τεχνητής νοημοσύνης που χρησιμοποιούν αδιάκριτη απόσπαση βιομετρικών δεδομένων από τα μέσα κοινωνικής δικτύωσης ή τις κάμερες κλειστού κυκλώματος παρακολούθησης για τη δημιουργία βάσεων δεδομένων αναγνώρισης προσώπου.
4. Ενώ η Επιτροπή πρότεινε να κατηγοριοποιούνται αυτόματα ως υψηλού κινδύνου όλα τα συστήματα που εμπίπτουν σε ορισμένους τομείς ή περιπτώσεις χρήσης, το Κοινοβούλιο προσθέτει την πρόσθετη απαίτηση ότι τα συστήματα πρέπει να ενέχουν “σημαντικό κίνδυνο” για να χαρακτηριστούν ως υψηλού κινδύνου. Επιπλέον, το Κοινοβούλιο επιβάλλει σε όσους αναπτύσσουν ένα σύστημα υψηλού κινδύνου στην ΕΕ να διενεργούν εκτίμηση επιπτώσεων στα θεμελιώδη δικαιώματα, συμπεριλαμβανομένης της διαβούλευσης με την αρμόδια αρχή και τα σχετικά ενδιαφερόμενα μέρη.
5. Το Κοινοβούλιο επιθυμεί να ενσωματώσει στην Πράξη ΤΝ μια πολυεπίπεδη προσέγγιση για τη ρύθμιση των συστημάτων ΤΝ γενικού σκοπού. Το Κοινοβούλιο επιθυμεί να επιβάλλει στους παρόχους μοντέλων θεμελίωσης την υποχρέωση να διασφαλίζουν ισχυρή προστασία των θεμελιωδών δικαιωμάτων, της υγείας, της ασφάλειας, του περιβάλλοντος, της δημοκρατίας και του κράτους δικαίου. Επιπλέον, τα παραγωγικά μοντέλα θεμελίωσης ΤΝ (όπως το GPT-4 και άλλα παρόμοια συστήματα) που χρησιμοποιούν μεγάλα γλωσσικά μοντέλα για τη δημιουργία τέχνης, μουσικής και άλλου περιεχομένου θα υπόκεινται σε αυστηρές υποχρεώσεις διαφάνειας. Τέλος, όλα τα μοντέλα θεμελίωσης θα πρέπει να παρέχουν όλες τις απαραίτητες πληροφορίες ώστε οι μεταγενέστεροι πάροχοι να είναι σε θέση να συμμορφωθούν με τις υποχρεώσεις τους βάσει της Πράξης για την ΤΝ.
6. Οι αρμοδιότητες των εθνικών αρχών έχουν ενισχυθεί και το Κοινοβούλιο προτείνει επίσης τη σύσταση Γραφείου ΤΝ (AI Office), ενός νέου φορέα της ΕΕ για την υποστήριξη της εναρμονισμένης εφαρμογής της πράξης ΤΠ, την παροχή καθοδήγησης και το συντονισμό κοινών διασυνοριακών ερευνών.

7. Προκειμένου να υποστηριχθεί η καινοτομία, το Κοινοβούλιο συμφωνεί ότι οι ερευνητικές δραστηριότητες και η ανάπτυξη στοιχείων ΤΝ ελεύθερου και ανοικτού κώδικα θα εξαιρούνται σε μεγάλο βαθμό από τη συμμόρφωση με τους κανόνες της Πράξης ΤΝ και προτρέπει στη δημιουργία ρυθμιστικών δοκιμαστηρίων.

Αφού ακολούθησαν οι διαδικασίες τριμερούς διαλόγου τον Ιούνιο, τον Ιούλιο, τον Σεπτέμβριο και τον Οκτώβριο του 2023, μετά από μακροχρόνιες διαπραγματεύσεις, η Προεδρία του Συμβουλίου και οι διαπραγματευτές του Ευρωπαϊκού Κοινοβουλίου κατέληξαν σε προσωρινή συμφωνία σχετικά με την πράξη ΤΝ στις 9 Δεκεμβρίου 2023. Το Κοινοβούλιο ενέκρινε την πράξη ΤΝ τον Μάρτιο 2024 με 523 ψήφους υπέρ, 46 κατά και 49 αποχές. Η πράξη ΑΙ εν τέλει :

1. Ενσωματώνει στο δίκαιο της ΕΕ έναν ορισμό των συστημάτων ΤΝ που ευθυγραμμίζεται με τον αναθεωρημένο ορισμό και περιέχει επίσης έναν ορισμό των μοντέλων γενικού σκοπού (GPAI).
2. Εφαρμόζεται κυρίως στους παρόχους και τους φορείς ανάπτυξης που θέτουν σε λειτουργία ή διαθέτουν στην αγορά της ΕΕ συστήματα ΤΝ και μοντέλα γενικού σκοπού (GPAI) και οι οποίοι έχουν την έδρα τους ή βρίσκονται στην ΕΕ, καθώς και στους φορείς ανάπτυξης ή παρόχους συστημάτων ΤΝ που είναι εγκατεστημένοι σε τρίτη χώρα, όταν η έξοδος που παράγεται από τα συστήματά τους χρησιμοποιείται στην ΕΕ.
3. Διατηρεί την προσέγγιση με βάση τον κίνδυνο που πρότεινε η Επιτροπή και κατατάσσει τα συστήματα ΤΝ σε διάφορες κατηγορίες κινδύνου, με διαφορετικούς βαθμούς ρύθμισης.
4. Απαγορεύει ένα ευρύτερο φάσμα πρακτικών ΤΝ, όπως είχε αρχικά προτείνει η Επιτροπή, λόγω των επιβλαβών επιπτώσεών τους.
5. Προσδιορίζει μια σειρά περιπτώσεων χρήσης στις οποίες τα συστήματα ΤΝ πρέπει να θεωρούνται υψηλού κινδύνου, επειδή μπορούν δυνητικά να δημιουργήσουν δυσμενείς επιπτώσεις στην υγεία, την ασφάλεια ή τα θεμελιώδη δικαιώματα των ανθρώπων.
6. Προσδιορίζει έναν αριθμό συστημάτων ΤΝ που ενέχουν περιορισμένους κινδύνους λόγω της έλλειψης διαφάνειας (π.χ. βαθιές απάτες, τεχνητό περιεχόμενο), τα οποία θα υπόκεινται σε απαιτήσεις πληροφόρησης και διαφάνειας.
7. Επιτρέπει τη χρήση συστημάτων που παρουσιάζουν ελάχιστους κινδύνους για τους ανθρώπους (π.χ. φίλτρα ανεπιθύμητης αλληλογραφίας) και συμμορφώνονται με την ισχύουσα νομοθεσία (π.χ. ΓΚΠΔ).
8. Παρέχει ειδικούς κανόνες για τα μοντέλα ΤΝ γενικού σκοπού και για τα μοντέλα γενικού σκοπού με «ικανότητες υψηλού αντίκτυπου» που θα μπορούσαν να αποτελέσουν συστημικό κίνδυνο και να έχουν σημαντικό αντίκτυπο στην εσωτερική αγορά. Εξαιρέσεις ισχύουν για τα μοντέλα γενικού σκοπού ελεύθερου και ανοικτού κώδικα.
9. Ζητά από τα κράτη μέλη να δημιουργήσουν ρυθμιστικά δοκιμαστήρια και επιτρέπει τη δοκιμή συστημάτων ΤΝ υψηλού κινδύνου στον πραγματικό κόσμο, ώστε να διευκολυνθεί η ανάπτυξη, η κατάρτιση, η δοκιμή και η επικύρωση καινοτόμων συστημάτων ΤΝ.
10. Η εφαρμογή της πράξης για την ΤΝ θα είναι ευθύνη πολλών φορέων τόσο σε εθνικό επίπεδο όσο και σε επίπεδο ΕΕ.
11. Η εφαρμογή της πράξης ΤΝ θα γίνει σταδιακά σε διάστημα δύο ετών (ξεκινώντας με τη σταδιακή κατάργηση των απαγορευμένων συστημάτων εντός έξι μηνών από την έναρξη ισχύος της πράξης) και θα απαιτήσει από την Ευρωπαϊκή Επιτροπή να εκδώσει διάφορες εκτελεστικές, κατ' εξουσιοδότηση και κατευθυντήριες γραμμές.

Η πράξη ΤΝ εγκρίθηκε επίσημα από το Κοινοβούλιο κατά τη σύνοδο ολομέλειας του Μαρτίου 2024 (ανακοινώθηκε ορθή επανάληψη στη σύνοδο του Απριλίου 2024) πριν από την τελική έγκριση από το Συμβούλιο.

3.6.3 Ρύθμιση

Ενώ οι βασικοί πυλώνες του νόμου για την τεχνητή νοημοσύνη έχουν πλέον συμφωνηθεί, η νομοθετική του μορφή δεν έχει ακόμη αποκρυσταλλωθεί. Τα επόμενα βήματα θα αφιερωθούν σε λεπτομερείς τεχνικές συζητήσεις για τη βελτίωση του τελικού κειμένου. Μετά από αυτή τη φάση, το Ευρωπαϊκό Κοινοβούλιο και το Συμβούλιο της ΕΕ πρέπει να εγκρίνουν επισήμως το κείμενο, το οποίο μετά τη δημοσίευσή του στην Επίσημη Εφημερίδα της Ευρωπαϊκής Ένωσης αναμένεται να τεθεί σε ισχύ το δεύτερο τρίμηνο του 2024. Θα ακολουθήσει μια επισκόπηση⁸⁶ των βασικών συμφωνιών που επιτεύχθηκαν κατά τη διάρκεια των διαπραγματεύσεων, των επόμενων βημάτων και κάποιων συμπερασμάτων για ανάληψη δράσης από οργανισμούς που χρησιμοποιούν τεχνολογίες τεχνητής νοημοσύνης εφόσον επιθυμούν να συμμορφωθούν κατά τη διάρκεια αυτής της μεταβατικής περιόδου.

3.6.3.1 Πεδίο εφαρμογής

Ο νόμος περί τεχνητής νοημοσύνης υιοθετεί τα κύρια στοιχεία του ορισμού της τεχνητής νοημοσύνης από τον ΟΟΣΑ. Το πεδίο εφαρμογής δεν ισχύει για κανένα λογισμικό ή αλγόριθμο εάν δεν υπάρχει εκπαίδευση και εάν δεν παρέχεται έξοδος (output) με βάση τις πληροφορίες της εκπαίδευσης. Ωστόσο, θα είναι απαραίτητη μια αξιολόγηση καθώς δεν προσδιορίζεται η ανθρώπινη παρέμβαση, η οποία αν αναφερόταν, θα είχε αποκλείσει τα παραγωγικά συστήματα τεχνητής νοημοσύνης.

Σύμφωνα με την προσωρινή συμφωνία, τα δωρεάν και ανοιχτού κώδικα συστήματα τεχνητής νοημοσύνης θα έχουν περιορισμένη εξαίρεση από το πεδίο εφαρμογής του κανονισμού. Αυτή η εξαίρεση δεν θα καλύπτει παρόχους μοντέλων ανοιχτού κώδικα συστημάτων υψηλού κινδύνου, απαγορευμένων εφαρμογών, λύσεων τεχνητής νοημοσύνης που ενέχουν κίνδυνο χειραγώγησης ή μοντέλων τεχνητής νοημοσύνης γενικού σκοπού με συστημικούς κινδύνους. Υποχρεώσεις όσον αφορά τη διαφάνεια και τα πνευματικά δικαιώματα θα ισχύουν επίσης για τους παρόχους τέτοιων συστημάτων τεχνητής νοημοσύνης ανοιχτού κώδικα.

3.6.3.2 Μοντέλα GPAI/θεμελίωσης και απαιτήσεις διαφάνειας

Σχετικά με ένα από τα πιο αμφισβητούμενα ζητήματα, τη ρύθμιση της τεχνητής νοημοσύνης γενικού σκοπού (GPAI) ή των μοντέλων θεμελίωσης, οι διαπραγματευτές συμφώνησαν να υιοθετήσουν την αρχική πρόταση του Ευρωπαϊκού Κοινοβουλίου για μια κλιμακωτή προσέγγιση που διακρίνει μεταξύ όλων των μοντέλων GPAI και μοντέλων με δυνητικούς συστημικούς κινδύνους. Η συμπερίληψη της ρύθμισης για τα μοντέλα θεμελίωσης ήρθε σχετικά αργά στη διαδικασία, όταν το Ευρωπαϊκό Κοινοβούλιο το ενσωμάτωσε στο κείμενο που δημοσίευσε τον Ιούνιο του 2023. Οι ευρέως δημοσιευμένες αντιρρήσεις των γαλλικών, γερμανικών και ιταλικών κυβερνήσεων στα τέλη Νοεμβρίου του 2023 έθεσαν σημαντικές αμφιβολίες για την τελική του τύχη του⁸⁷. Ερμήνευσαν τη συμπερίληψη οποιωνδήποτε κανόνων για το GPAI ως πιθανό εμπόδιο στην καινοτομία, δυνητικά καταπνίγοντας το ανταγωνιστικό πλεονέκτημα που θα μπορούσαν διαφορετικά να έχουν ορισμένοι επιτυχημένοι προγραμματιστές τεχνητής νοημοσύνης της ΕΕ. Η εναλλακτική

⁸⁶<https://www.lexology.com/library/detail.aspx?f600392e-90c4-4b36-9457-c76703758756>

⁸⁷Exclusive: Germany, France and Italy reach agreement on future AI regulation, Andreas Rinke, 20 Νοεμβρίου 2023, <https://www.reuters.com/technology/germany-france-italy-reach-agreement-future-ai-regulation-2023-11-18/>

που πρότειναν, ήταν να βασιστούμε σε προαιρετικούς κώδικες δεοντολογίας. Στην τελική πολιτική συμφωνία, εξασφαλίστηκε συμβιβασμός σύμφωνα με τον οποίο αυτοί οι κώδικες θα έχουν συμπληρωματικό ρόλο αντί να αντικαθιστούν τις ουσιαστικές υποχρεώσεις.

Για τα μοντέλα GPAI που έχουν υψηλό αντίκτυπο και ενέχουν συστημικό κίνδυνο, οι διαπραγματευτές του Ευρωπαϊκού Κοινοβουλίου κατάφεραν να επιβάλουν πιο αυστηρές απαιτήσεις κατά τη διάρκεια των διαπραγματεύσεων. Εάν τα μοντέλα GPAI έχουν φτάσει σε ένα ορισμένο όριο υπολογιστικών πόρων που χρησιμοποιούνται στην εκπαίδευση, θα απαιτείται η διεξαγωγή αξιολογήσεων μοντέλων, η αξιολόγηση και ο μετριασμός των συστημικών κινδύνων, η διεξαγωγή δοκιμών αντιπαράθεσης (adversarial testing), η αναφορά στην Ευρωπαϊκή Επιτροπή για σοβαρά συμβάντα, η διασφάλιση της κυβερνοασφάλειας και η αναφορά στην ενεργειακή απόδοση τέτοιων μοντέλων⁸⁸. Τα όρια υπολογισμού ορίζονται σε $(10)^{25}$ floating point operations (FLOPs), που θα καταγράφουν μόνο τα πολύ μεγαλύτερα μοντέλα. Έρευνες⁸⁹ έχουν δείξει ότι διάφορες “αναδυόμενες ιδιότητες” εμφανίζονται κάπου πάνω από $(10)^{24}$ FLOPs. Ωστόσο, δεδομένων των τάσεων για αποτελεσματικότερη χρήση των υπολογιστικών πόρων και των αναδυόμενων ιδιοτήτων που εκπαιδεύονται σε μικρότερα μοντέλα, αυτά τα όρια μπορεί να χρειαστεί να μειωθούν με την πάροδο του χρόνου. Η Επιτροπή θα έχει τη δυνατότητα να ενημερώσει αυτά τα κατώτατα όρια και να προσθέσει άλλα κριτήρια μέσω κατ’ εξουσιοδότηση πράξεων.

Συνοπτικά, λοιπόν, σε ό,τι αφορά τα μοντέλα GPAI: -Όλοι οι πάροχοι πρέπει να παρέχουν τεχνική τεκμηρίωση, οδηγίες χρήσης, να συμμορφώνονται με την οδηγία για τα πνευματικά δικαιώματα και να δημοσιεύουν περίληψη σχετικά με το περιεχόμενο που χρησιμοποιείται για την κατάρτιση. - Οι πάροχοι μοντέλων με ελεύθερη και ανοικτή άδεια χρήσης πρέπει μόνο να συμμορφώνονται με τα δικαιώματα πνευματικής ιδιοκτησίας και να δημοσιεύουν την περίληψη των δεδομένων κατάρτισης, εκτός εάν παρουσιάζουν συστημικό κίνδυνο. - Όλοι οι πάροχοι μοντέλων που παρουσιάζουν συστημικό κίνδυνο -ανοικτής ή κλειστής άδειας- πρέπει επίσης να διενεργούν αξιολογήσεις μοντέλων, αντιπαραθετικές δοκιμές, να παρακολουθούν και να αναφέρουν σοβαρά περιστατικά και να διασφαλίζουν την κυβερνοασφάλεια.

3.6.3.3 Απαιτήσεις για συστήματα τεχνητής νοημοσύνης υψηλού κινδύνου

Καθορίστηκαν σαφείς ευθύνες για συστήματα τεχνητής νοημοσύνης που προσδιορίζονται ως υψηλού κινδύνου, λόγω των σημαντικών δυνατοτήτων τους να βλάψουν την ανθρώπινη ασφάλεια και τα θεμελιώδη δικαιώματα. Οι διαπραγματευτές συμφώνησαν σε ένα σύνολο κριτηρίων που έχουν σχεδιαστεί για να ταυτοποιούν τον πράγματι υψηλό κίνδυνο, ενώ οι τομείς που συμφωνήθηκαν ως ευαίσθητοι περιλαμβάνουν περιπτώσεις χρήσης στην εκπαίδευση, την απασχόληση, τις ζωτικές υποδομές, τις βασικές ιδιωτικές και δημόσιες υπηρεσίες, την επιβολή του νόμου, τον έλεγχο των συνόρων, τις δημοκρατικές διαδικασίες και την απονομή δικαιοσύνης⁹⁰. Οι προγραμματιστές συστημάτων τεχνητής νοημοσύνης αυτής της κατηγορίας κινδύνου θα πρέπει να διενεργούν αξιολογήσεις συμμόρφωσης για να διασφαλίσουν ότι το σύστημα πληροί τις βασικές απαιτήσεις για αξιόπιστη τεχνητή νοημοσύνη, πιο συγκεκριμένα: ποιότητα δεδομένων, τεκμηρίωση, ιχνηλασιμότητα, διαφάνεια, ανθρώπινη εποπτεία, ακρίβεια, κυβερνοασφάλεια και ευρωστία. Οι προγραμματιστές θα είναι επίσης υπεύθυνοι για την εφαρμογή προτύπων διαχείρισης ποιότητας και κινδύνου.

Οι προγραμματιστές τέτοιων συστημάτων θα υπόκεινται επίσης σε υποχρεώσεις, συμπεριλαμβανομένης της συνεχούς παρακολούθησης και της διεξαγωγής, σε ορισμένες περιπτώσεις, αξιολο-

⁸⁸<https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai>

⁸⁹Για παράδειγμα βλ. την έρευνα Wei et al: [2206.07682] Emergent Abilities of Large Language Models, <https://arxiv.org/abs/2206.07682>

⁹⁰European Union squares the circle on the world's first AI rulebook, Luca Bertuzzi, 14 Δεκεμβρίου 2023, <https://www.euractiv.com/section/artificial-intelligence/news/european-union-squares-the-circle-on-the-worlds-first-ai-rulebook/>

γήσεων επιπτώσεων στα θεμελιώδη δικαιώματα και αναφοράς δυσλειτουργιών. Οι λεπτομέρειες της αξιολόγησης του αντίκτυπου των θεμελιωδών δικαιωμάτων αναμένουν πλήρη διευκρίνιση, αλλά –σύμφωνα με την Επιτροπή της ΕΕ– θα αποτελούνται από περιγραφή των διαδικασιών ανάπτυξης που σχετίζονται με το σύστημα τεχνητής νοημοσύνης υψηλού κινδύνου, τη χρονική περίοδο και τη συχνότητα κατά την οποία η τεχνητή νοημοσύνη υψηλού κινδύνου σύστημα που προορίζεται να χρησιμοποιηθεί, κατηγορίες φυσικών προσώπων και ομάδων που ενδέχεται να επηρεαστούν από τη χρήση του στο συγκεκριμένο πλαίσιο, ειδικοί κίνδυνοι βλάβης που ενδέχεται να επηρεάσουν τις πληττόμενες κατηγορίες προσώπων ή ομάδα προσώπων, περιγραφή της εφαρμογής του μέτρα ανθρώπινης εποπτείας και τα μέτρα που πρέπει να λαμβάνονται σε περίπτωση υλοποίησης των κινδύνων.

Η Επιτροπή της ΕΕ έχει αναγνωρίσει ότι ενδέχεται να υπάρχει κάποια αλληλοεπικάλυψη με τη διαδικασία που εφαρμόζεται στις αξιολογήσεις επιπτώσεων της επεξεργασίας δεδομένων σύμφωνα με τον Γενικό Κανονισμό Προστασίας Δεδομένων της ΕΕ, αλλά δήλωσε ότι – δεδομένου του ευρύτερου πεδίου εφαρμογής του νόμου περί τεχνητής νοημοσύνης που υπερβαίνει τα προσωπικά δεδομένα – θα χρειαστεί ακόμη μια συγκεκριμένη αξιολόγηση.

3.6.3.4 Εξαιρέσεις για την επιβολή του νόμου

Οι διαπραγματευτές κατέληξαν σε συναίνεση σχετικά με την εφαρμογή ενός συνόλου εγγυήσεων και περιορισμένων εξαιρέσεων για την ανάπτυξη συστημάτων βιομετρικής ταυτοποίησης (RBIs) σε δημόσιους χώρους για την επιβολή του νόμου⁹¹. Αυτό εξαρτάται από την προηγούμενη έγκριση από το δικαστικό σώμα και περιορίζεται σε ένα σχολαστικά καθορισμένο φάσμα σοβαρών εγκληματικών δραστηριοτήτων. Η χρήση “σε μεταγενέστερο χρόνο εξ αποστάσεως” RBI θα περιορίζεται αυστηρά στην εστιασμένη καταδίωξη ατόμων που είτε έχουν καταδικαστεί είτε είναι ύποπτα για σοβαρά ποινικά αδικήματα. Η χρήση των RBI σε “πραγματικό χρόνο” θα πρέπει να συμμορφώνεται με αυστηρούς όρους και θα είναι περιορισμένη ως προς το χρόνο και την τοποθεσία, και θα πρέπει να χρησιμοποιείται για συγκεκριμένους σκοπούς, όπως η πρόληψη μιας συγκεκριμένης και παρούσας τρομοκρατικής απειλής.

3.6.3.5 Απαγορευμένες πρακτικές

Οι διαπραγματευτές συμφώνησαν επίσης ότι η λίστα των συστημάτων τεχνητής νοημοσύνης θα απαγορεύει πλήρως:

- βιομετρικά συστήματα κατηγοριοποίησης που χρησιμοποιούν ευαίσθητα χαρακτηριστικά (π.χ. πολιτικές, θρησκευτικές, φιλοσοφικές πεποιθήσεις, σεξουαλικός προσανατολισμός, φυλή)
- μη στοχευμένη εξαγωγή εικόνων προσώπων από το διαδίκτυο ή πλάνα CCTV για τη δημιουργία βάσεων δεδομένων αναγνώρισης προσώπου.
- αναγνώριση συναισθημάτων στο χώρο εργασίας και στα εκπαιδευτικά ιδρύματα.
- κοινωνική βαθμολογία με βάση την κοινωνική συμπεριφορά ή τα προσωπικά χαρακτηριστικά.
- χειραγώγηση της ανθρώπινης συμπεριφοράς για την παράκαμψη της ελεύθερης βούλησης.
- εκμετάλλευση των ευάλωτων σημείων των ανθρώπων λόγω ηλικίας, αναπηρίας, κοινωνικής ή οικονομικής κατάστασης.

⁹¹<https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai>

Η συμφωνία για τον κατάλογο των απαγορευμένων χρήσεων έγινε με πολλές παλινωδίες μεταξύ των θεσμικών οργάνων, με το Κοινοβούλιο να εισαγάγει αρχικά εκτενέστερους τομείς απαγόρευσης από εκείνους που τελικά συμφωνήθηκαν. Οι συμβιβασμοί που επετεύχθησαν σε σχέση με την τελική λίστα, καθώς και η εξαίρεση από την επιβολή του νόμου για τις RBI, υποδηλώνουν την ένταση των διαπραγματεύσεων και τον μαραθώνιο τριμερή διάλογο που έλαβε χώρα.

3.6.3.6 Κυρώσεις

Η μη τήρηση του κανονισμού μπορεί να οδηγήσει σε πρόστιμο έως και 35 εκατομμύρια ευρώ ή 7% του παγκόσμιου κύκλου εργασιών, ανάλογα με τη φύση της παράβασης και το μέγεθος της εταιρείας. Για συστήματα TN υψηλού κινδύνου, η ποινή μπορεί να ανέλθει σε 15 εκατ. ευρώ ή 3%. Προβλέπονται εξαιρέσεις για μικρότερες επιχειρήσεις, με περιορισμένες κυρώσεις για ΜΜΕ και νεοφυείς επιχειρήσεις. Όσον αφορά τις ποινές, έχει επιτευχθεί μια ισορροπία μεταξύ της ανάγκης ρύθμισης της τεχνητής νοημοσύνης και του στόχου να μην περιοριστεί η ανάπτυξη της τεχνολογίας στην ΕΕ. Για τον ίδιο λόγο, προβλέπονται πλέον οι λεγόμενες λύσεις “sandboxing” (ρυθμιστικά δοκιμαστήρια), όπου οι λύσεις μπορούν να δοκιμαστούν ενώ επωφελούνται από ένα ειδικό καθεστώς.

3.6.3.7 Τι έπεται;

Οι υπεύθυνοι χάραξης πολιτικής της ΕΕ ξεκίνησαν τεχνικές συζητήσεις κατά την εβδομάδα της 11ης Δεκεμβρίου 2023 για να επιδιώξουν να οριστικοποιήσουν τις λεπτομέρειες της νομικής διατύπωσης. Μετά τη συμφωνία επί του τελικού κειμένου, ο προγραμματισμός ήταν να λάβει επίσημη έγκριση τόσο από το Κοινοβούλιο όσο και από το Συμβούλιο πριν από τη λήξη της νομοθετικής περιόδου τον Απρίλιο του 2024, ώστε να τεθεί σε ισχύ 20 ημέρες μετά τη δημοσίευσή του στην Επίσημη Εφημερίδα της ΕΕ. Οι περισσότερες από τις διατάξεις του θα εφαρμοστούν μετά από 24 μήνες. Οι κανόνες για τα απαγορευμένα συστήματα TN θα εφαρμοστούν μετά από 6 μήνες, οι κανόνες για τα GPAI μετά από 12 μήνες και οι κανόνες για τα συστήματα TN υψηλού κινδύνου μετά από 36 μήνες.

Ο νόμος για την TN θα συνοδεύεται από μια σύνθετη και πολυεπίπεδη δομή διακυβέρνησης που θα περιλαμβάνει πολλαπλές οντότητες, όπως κοινοποιούντες και κοινοποιημένους οργανισμούς, οργανισμούς αξιολόγησης της συμμόρφωσης, συμβούλιο TN, γραφείο TN, εθνικές αρμόδιες αρχές και αρχές εποπτείας της αγοράς. Φορείς όπως το Γραφείο TN θα υποστηρίζουν επίσης τους φορείς που εμπíπτουν στο πεδίο εφαρμογής μέσω της ανάπτυξης κωδίκων πρακτικής με βάση το διάλογο με τους ενδιαφερόμενους φορείς. Επιπλέον, οι εν λόγω φορείς θα διαδραματίσουν επίσης ρόλο στα διάφορα μέτρα για τη στήριξη της καινοτομίας, όπως τα ρυθμιστικά πλαίσια TN και τα μέτρα για τις ΜΜΕ και τις νεοσύστατες επιχειρήσεις. Το νέο Ευρωπαϊκό Γραφείο Τεχνητής Νοημοσύνης, το οποίο θα συμμετέχει στην Επιτροπή, θα επιφορτιστεί με τον συντονισμό των εποπτικών αρχών των κρατών μελών της ΕΕ, την εποπτεία της εφαρμογής του νόμου για την τεχνητή νοημοσύνη και την επιβολή των κανόνων για το GPAI. Η Επιτροπή έχει επίσης εγκαινιάσει ένα “Σύμφωνο AI” (AI Pact) για φορείς που επιθυμούν να υιοθετήσουν πρώτοι τον κανονισμό. Αυτό το σύστημα βρίσκεται ήδη σε επικοινωνία με τα ενδιαφερόμενα μέρη που θα δηλώσουν ότι θα εργαστούν για την έγκαιρη συμμόρφωση, δημιουργώντας μια κοινότητα ανταλλαγής βέλτιστων πρακτικών πριν από την εφαρμογή της Πράξης για την Τεχνητή Νοημοσύνη.

3.7 Ζητήματα κυβερνοασφάλειας στην Πράξη για την TN (άρθρο 15)

Τα συστήματα τεχνητής νοημοσύνης είναι υπολογιστικά συστήματα και, ως τέτοια, κληρονομούν όλους τους κινδύνους κυβερνοασφάλειας που συνδέονται ήδη με τα παραδοσιακά ψηφιακά συστήματα που λειτουργούν σε παρόμοια πλαίσια. Επιπλέον, η υιοθέτηση τεχνολογιών TN εισάγει

νέους κινδύνους κυβερνοασφάλειας που συνδέονται με την εμφάνιση νέων κατηγοριών ευπαθειών και επιθέσεων, που μπορούν να πλήξουν ειδικά την ΤΝ και να θέσουν σε κίνδυνο κρίσιμες υποδομές των κρατών. Η κυβερνοασφάλεια της ΤΝ είναι ένας αναδυόμενος τομέας που αποσκοπεί στην έρευνα και την αντιμετώπιση αυτών των ειδικών ευπαθειών της ΤΝ, συμπεριλαμβανομένων των επιθέσεων μηχανικής μάθησης, της δηλητηρίασης δεδομένων ή των backdoors που ενσωματώνονται σε μοντέλα ΤΝ. Επιπλέον η κυβερνοασφάλεια μπορεί να αποτρέψει τεχνικές εξαπάτησης όπως το deepfake αλλά και τα robot φωνής είτε κειμένου (chat) που παραπλανούν, ειδικά μετά την άνθιση των μοντέλων παραγωγικής ΤΝ (generative AI).

Η εσωτερική δομή των συστημάτων ΤΝ περιλαμβάνει μια σειρά από στοιχεία. Παρόλο που τα μοντέλα ΤΝ συγκαταλέγονται μεταξύ αυτών και αποτελούν βασικά συστατικά των συστημάτων ΤΝ, ωστόσο δεν αποτελούν από μόνα τους συστήματα ΤΝ. Η απαίτηση κυβερνοασφάλειας του νόμου ΤΝ ισχύει για το σύστημα ΤΝ στο σύνολό του και όχι απευθείας για τα εσωτερικά συστατικά του. Ειδικότερα το άρθρο 15 αναφέρεται στην “Ακρίβεια, στιβαρότητα και κυβερνοασφάλεια” και ορίζει ότι τα συστήματα ΤΝ υψηλού κινδύνου αναπτύσσονται σύμφωνα με την αρχή της ασφάλειας ήδη από τον σχεδιασμό και εξ ορισμού (by design and by default). Λαμβανομένου υπόψη του επιδιωκόμενου σκοπού τους, θα πρέπει να επιτυγχάνουν κατάλληλο επίπεδο ακρίβειας, στιβαρότητας, ασφάλειας και κυβερνοασφάλειας καθ’ όλη τη διάρκεια του κύκλου ζωής τους. Η διαδικασία αυτή απαιτεί μια ολοκληρωμένη και συνεχή προσέγγιση με τη χρήση αποδεδειγμένων πρακτικών και διαδικασιών κυβερνοασφάλειας σε συνδυασμό με ειδικούς για την ΤΝ ελέγχους.

Αν και η κατάσταση της τεχνολογίας για την ασφάλεια των μοντέλων ΤΝ έχει περιορισμούς, τα συστήματα ΤΝ μπορούν να επιτύχουν συμμόρφωση με την απαίτηση κυβερνοασφάλειας της Πράξης για την ΤΝ, εφόσον οι κίνδυνοι κυβερνοασφάλειας μετριάζονται αποτελεσματικά μέσω άλλων μέτρων που δεν αναπτύσσονται αποκλειστικά σε επίπεδο μοντέλου ΤΝ.

Σύμφωνα με το καθιερωμένο σύστημα ρύθμισης της ασφάλειας των προϊόντων στην ΕΕ⁹², τα εναρμονισμένα πρότυπα αποτελούν ένα από τα κύρια μέσα για την επίτευξη συμμόρφωσης και συμμόρφωσης με τις νομοθετικές απαιτήσεις. Τα εναρμονισμένα πρότυπα θα αναπτυχθούν από τους ευρωπαϊκούς οργανισμούς τυποποίησης μετά από αίτημα τυποποίησης της Ευρωπαϊκής Επιτροπής που δημοσιεύθηκε τον Μάιο του 2023. Είναι σημαντικό να σημειωθεί ότι το αίτημα τυποποίησης καθιστά σαφές ότι η εμβέλεια της τυποποίησης της ΤΝ καθορίζεται από την τεχνολογική ωριμότητα των τεχνολογιών ΤΝ, αναφερόμενη ρητά στην κατάσταση της τεχνολογίας ως βάση της τυποποίησης⁹³.

Οι απαιτήσεις κυβερνοασφάλειας, ακρίβειας και ευρωστίας για συστήματα ΤΝ υψηλού κινδύνου, όπως ορίζονται στο άρθρο 15 της Πράξης για την ΤΝ, συνδέονται με την τεχνική διάσταση των συστημάτων ΤΝ και απαιτούν βαθιά κατανόηση της εσωτερικής λειτουργίας των συστημάτων ΤΝ, των καθιερωμένων τεχνικών πρακτικών και προτύπων. Παρόλο που τα καθιερωμένα πρότυπα και πρακτικές στον τομέα της ασφάλειας στον κυβερνοχώρο μπορεί να εφαρμόζονται στα συστήματα ΤΝ όπως και σε άλλα συστήματα λογισμικού, υπάρχουν τεχνολογικές προκλήσεις ειδικά για την ΤΝ, οι οποίες δεν έχουν ακόμη αποτελέσει αντικείμενο καθιερωμένων πρακτικών ασφάλειας ή ειδικών προτύπων. Επί του παρόντος, για σκοπούς μηχανικής ασφάλειας, όπως η πρακτική εφαρμογή διαδικασιών και τεχνικών για την ασφάλεια συστημάτων, πολλές προσεγγίσεις και εργαλεία ασφάλειας ειδικά για την ΤΝ δεν μπορούν να θεωρηθούν αρκετά ώριμες ώστε να χρησιμοποιηθούν άμεσα για τη σωστή διασφάλιση ορισμένων μοντέλων ΤΝ μεμονωμένα.

Η κυβερνοασφάλεια της ΤΝ είναι ένας αναδυόμενος τομέας που στοχεύει να καλύψει αυτό το κενό και που στηρίζεται σε μεγάλο βαθμό σε τρέχουσες ερευνητικές δραστηριότητες σε το-

⁹²Κανονισμός 1025/2012 του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου, της 25ης Οκτωβρίου 2012, σχετικά με την ευρωπαϊκή τυποποίηση, <https://eur-lex.europa.eu/eli/reg/2012/1025/2023-07-09>

⁹³Junklewitz, H., Hamon, R., André, A., Evas, T., Soler Garrido, J. and Sanchez Martin, J.I., Cybersecurity of Artificial Intelligence in the AI Act, Publications Office of the European Union, Luxembourg, 2023, doi:10.2760/271009, JRC134461, <https://publications.jrc.ec.europa.eu/repository/handle/JRC134461>

μείς όπως η μηχανική ασφάλειας (security engineering) ή η αντιπαραθετική μηχανική μάθηση (adversarial machine learning). Στην πραγματικότητα, ο κύριος σκοπός της Πράξης για την TN είναι να αντιμετωπίσει τους ειδικούς κινδύνους που σχετίζονται με την τεχνολογία TN, και συνεπώς αναμένεται ότι απαιτούνται εκτιμήσεις που υπερβαίνουν την καθιερωμένη πρακτική στην ασφάλεια λογισμικού για την αντιμετώπιση των απαιτήσεών της.

Το Παράρτημα I της Πράξης για την TN καλύπτει τόσο τη μηχανική μάθηση όσο και άλλες προσεγγίσεις της TN, όπως τα συστήματα βασισμένα στη γνώση. Επί του παρόντος η μηχανική μάθηση αντιπροσωπεύει τις πιο εξέχουσες τεχνικές προσεγγίσεις στην TN και θέτει επίσης τις περισσότερες από τις προκλήσεις για την κυβερνοασφάλεια TN που σχετίζονται με την Πράξη για την TN.

3.8 Συμπέρασμα

Τα τεράστια οφέλη και συγχρόνως οι πολυεπίπεδοι κίνδυνοι που απορρέουν από τη χρήση της τεχνολογίας της TN ώθησαν πολλούς φορείς της δημόσιας σφαίρας, σε εθνικό και υπερεθνικό επίπεδο να διατυπώσουν αρχές και εφαρμοστικά πλαίσια για τη βέλτιστη ρύθμισή της. Ειδικά η ΕΕ θέλησε να αναλάβει ηγετικό ρόλο και επεξεργαστεί το πρώτο πλήρες νομοθετικό πλαίσιο, με εν δυνάμει παγκόσμιο αντίκτυπο, που ρυθμίζει την TN με γνώμονα την ατομική, κοινωνική και οικονομική ευημερία, την ασφάλεια των πολιτών της, τον σεβασμό των θεμελιωδών δικαιωμάτων, του κράτους δικαίου και του περιβάλλοντος αλλά και την προώθηση της καινοτομίας.

Ο αναμενόμενος παγκόσμιος αντίκτυπος της Πράξης για την TN βρίσκεται σε συνέχεια της παγκόσμιας επιρροής του ΓΚΠΔ (GDPR) που αποτέλεσε τη βάση της θεωρίας του Brussels Effect, ένα φαινόμενο που αναλύεται στην επικράτηση των νομοθετικών προτύπων της ΕΕ σε παγκόσμια κλίμακα. Με την Πράξη για την TN η Ευρωπαϊκή Ένωση εφαρμόζει τη διαβάθμιση των συστημάτων TN βάσει κινδύνου και εισάγει ένα ρυθμιστικό πλαίσιο που είναι αλληλένδετο με Πράξεις της ΕΕ για τη ρύθμιση συναφών πολιτικών, όπως η ρύθμιση των δεδομένων του δημοσίου τομέα αλλά και των βιομηχανικών δεδομένων, η οριοθέτηση της λειτουργίας των μεγάλων πλατφορμών, των ψηφιακών υπηρεσιών και της ψηφιακής αγοράς και τέλος, αλλά με ύψιστη σπουδαιότητα, την κυβερνοασφάλεια.

Κεφάλαιο 4

Παρατηρητήρια, πρότυπα και δείκτες προς μια αξιόπιστη και ανθρωποκεντρική ΤΝ

- Οι τεχνολογικές και κοινωνικοοικονομικές εξελίξεις που σχετίζονται με την ΤΝ είναι ιδιαίτερα δυναμικές και απαιτούν εξειδικευμένη εμπειρογνώσια προκειμένου να καταγράφονται και να αξιολογούνται, ειδικά όταν στόχος είναι η εξυπηρέτηση του δημοσίου συμφέροντος, η αξιοπιστία και η ανθρωποκεντρική διάσταση. Αρχικά θα παρουσιάσουμε μια τεκμηρίωση των εννοιών “αξιόπιστη” και “ανθρωποκεντρική” ΤΝ και στη συνέχεια μια ανάλυση σχετικά με την προστασία των θεμελιωδών δικαιωμάτων και τη σύνδεσή τους με τα συστήματα ΤΝ.
- Στο πλαίσιο της παρακολούθησης της υιοθέτησης των τεχνολογιών ΤΝ, με σκοπό τον σχεδιασμό μιας τεκμηριωμένης δημόσιας πολιτικής, θα παρουσιάσουμε αντιπροσωπευτικά παραδείγματα σχετικά με τη δημιουργία παρατηρητηρίων, δεδομένου ότι συγκαταλέγονται στους φορείς παρακολούθησης που συγκεντρώνουν στοιχεία για την αξιολόγηση των εφαρμοσμένων πολιτικών. Επίσης θα εξετάσουμε τη θέσπιση προτύπων στην υπηρεσία του στόχου για αξιόπιστη και ανθρωποκεντρική τεχνητή νοημοσύνη.
- Τέλος θα τεκμηριώσουμε την καθιέρωση δεικτών που μετρούν την επίδοση της ΤΝ, βάσει πηγών του ευρωπαϊκού ρυθμιστικού πλαισίου, οι οποίοι σταδιακά προσανατολίζονται όλο και πιο στοχευμένα στη ρητή δέσμευση της ΕΕ για μια αξιόπιστη και ανθρωποκεντρική Τεχνητή Νοημοσύνη.

4.1 Αξιόπιστη και ανθρωποκεντρική τεχνητή νοημοσύνη

Η Ευρωπαϊκή Επιτροπή δημοσίευσε τον Φεβρουάριο του 2020 τη Λευκή Βίβλο για την Τεχνητή Νοημοσύνη¹ και πρότεινε τη δημιουργία ενός ευρωπαϊκού ρυθμιστικού πλαισίου για αξιόπιστη τεχνητή νοημοσύνη. Όρισε ως βασικό σκοπό το να καταστεί η Ευρώπη παγκόσμιος κόμβος αξιόπιστης τεχνητής νοημοσύνης που θα εγγυάται την υγεία, την ασφάλεια και τα θεμελιώδη δικαιώματα των πολιτών της. Ως παράλληλος στόχος τέθηκε να καθιερωθεί η Ευρώπη πρωτοπόρος και ηγέτιδα δύναμη στη ρύθμιση και την ανάπτυξη υπηρεσιών και λογισμικού τεχνητής νοημοσύνης, μέσα στο παγκόσμιο ανταγωνιστικό περιβάλλον. Συγχρόνως βεβαίωσε ρητά την προσήλωση σε μια ΤΝ που θα έχει θετικό αντίκτυπο για τους Ευρωπαίους πολίτες και θα εναρμονίζεται με τις ευρωπαϊκές αξίες.

¹<https://eur-lex.europa.eu/legal-content/EL/ALL/?uri=CELEX:52020DC0065>

Ανατρέχοντας στα κείμενα για την υιοθέτηση πολιτικών για την ΤΝ από την ΕΕ, συναντούμε τη διαπίστωση ότι η ΤΝ μπορεί να ωφελήσει την κοινωνία και την οικονομία στο σύνολό τους. Πρόκειται για μια τεχνολογία στρατηγικής σημασίας που αναπτύσσεται και χρησιμοποιείται με διαρκώς ταχύτερους ρυθμούς σε ολόκληρο τον κόσμο. Ωστόσο συνεπάγεται νέες προκλήσεις και εγείρει νομικά και δεοντολογικά ζητήματα. Η ΤΝ, κατά μια διασταλτική έννοια, δίνει στις μηχανές τη δυνατότητα να “εκπαιδεύονται” χρησιμοποιώντας δεδομένα και εκτελώντας αλγόριθμους. Αυτό με τη σειρά του οδηγεί στη λήψη και εφαρμογή αποφάσεων χωρίς ανθρώπινη παρέμβαση.

Εντούτοις, οι αποφάσεις αυτές μπορεί να προέρχονται από ελλιπή και αναξιόπιστα δεδομένα, μπορούν να αποτελέσουν αντικείμενο παραποίησης μέσω κυβερνοεπιθέσεων ή μπορεί να είναι μεροληπτικές ή εσφαλμένες. Επομένως, η απερισκεπτη εφαρμογή της αναπτυσσόμενης τεχνολογίας θα μπορούσε να έχει προβληματικές συνέπειες και να κάνει τους πολίτες επιφυλακτικούς ως προς την αποδοχή ή την εφαρμογή της².

Τα συστήματα ΤΝ θα πρέπει να βοηθούν τα άτομα να κάνουν καλύτερες και πιο τεκμηριωμένες επιλογές σύμφωνα με τους στόχους τους. Θα πρέπει να συμβάλλουν στη δημιουργία μιας ευημερούσας και ισότιμης κοινωνίας, υποστηρίζοντας την ανθρώπινη παρέμβαση και τα θεμελιώδη δικαιώματα και να μην περιορίζουν την ανθρώπινη αυτονομία. Στο επίκεντρο της λειτουργίας του συστήματος θα πρέπει να βρίσκεται η ελευθερία, η αυτονομία, η συνολική ευημερία του χρήστη και η ανθρώπινη εποπτεία. Χάρη στην ανθρώπινη εποπτεία διασφαλίζεται ότι τα συστήματα ΤΝ δεν υπονομεύουν την ανθρώπινη αυτονομία και δεν προκαλούν άλλες δυσμενείς επιπτώσεις. Ανάλογα με το συγκεκριμένο σύστημα που βασίζεται στην ΤΝ και τον τομέα εφαρμογής του, θα πρέπει να εξασφαλίζονται κατάλληλα επίπεδα μέτρων ελέγχου, συμπεριλαμβανομένης της προσαρμοστικότητας, της ακρίβειας και της εξηγησιμότητας των συστημάτων που βασίζονται στην ΤΝ. Πρέπει να διασφαλιστεί ότι οι δημόσιες αρχές έχουν την ικανότητα να ασκούν τις εποπτικές εξουσίες τους σύμφωνα με την εντολή που έχουν λάβει.

Η εποπτεία μπορεί να επιτευχθεί μέσω μηχανισμών διακυβέρνησης, όπως η διασφάλιση μιας προσέγγισης στην οποία ο άνθρωπος είναι εκείνος που παρεμβαίνει (human-in-the-loop), εκείνος που εποπτεύει (human-on-the-loop) ή εκείνος που ελέγχει (human-in-command), όπως αναλύεται στις “Κατευθυντήριες γραμμές δεοντολογίας για αξιόπιστη Τεχνητή Νοημοσύνη”. Η προσέγγιση της ανθρώπινης παρέμβασης (HITL) αναφέρεται στην ανθρώπινη παρέμβαση σε κάθε κύκλο λήψης αποφάσεων του συστήματος, κάτι που σε πολλές περιπτώσεις δεν είναι ούτε εφικτό ούτε και επιθυμητό. Η προσέγγιση της ανθρώπινης εποπτείας (HOTL) αναφέρεται στην ικανότητα ανθρώπινης παρέμβασης κατά τη διάρκεια του κύκλου σχεδιασμού του συστήματος και παρακολούθησης της λειτουργίας του. Η προσέγγιση του ανθρώπινου ελέγχου (HIC) αναφέρεται στην ικανότητα εποπτείας της συνολικής δραστηριότητας του συστήματος ΤΝ (συμπεριλαμβανομένου του οικονομικού, κοινωνικού, νομικού και δεοντολογικού αντικτύπου του), καθώς και στην ικανότητα λήψης αποφάσεων σχετικά με το πότε και πώς χρησιμοποιείται το σύστημα σε κάθε περίπτωση.

4.1.1 Κατευθυντήριες γραμμές δεοντολογίας για αξιόπιστη ΤΝ

Πριν την υποβολή της πρότασης για την “Πράξη για την ΤΝ”, η Ευρωπαϊκή Επιτροπή το 2018 διόρισε³ την “Ομάδα εμπειρογνομόνων υψηλού επιπέδου για την τεχνητή νοημοσύνη” (HLEG, High Level Expert Group) για την παροχή συμβουλών σχετικά με τη στρατηγική της για τεχνητή νοημοσύνη. Η ομάδα εμπειρογνομόνων εκπόνησε τις “Κατευθυντήριες γραμμές δεοντολογίας για αξιόπιστη τεχνητή νοημοσύνη” όπου προτείνει μια ανθρωποκεντρική προσέγγιση για την τεχνητή νοημοσύνη και απαριθμεί επτά βασικές απαιτήσεις που πρέπει να πληρούν τα συστήματα τεχνητής νοημοσύνης για να είναι αξιόπιστα.

² ΑΝΑΚΟΙΝΩΣΗ ΤΗΣ ΕΠΙΤΡΟΠΗΣ, COM/2019/168 final, Οικοδόμηση εμπιστοσύνης στην ανθρωποκεντρική τεχνητή νοημοσύνη, <https://eur-lex.europa.eu/legal-content/EL/ALL/?uri=CELEX:52019DC0168>

³ <https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai>

Διατύπωσε κατ' αρχάς την άποψη ότι η τεχνολογία ΤΝ θα πρέπει να αναπτύσσεται με επίκεντρο τον άνθρωπο και συνεπώς να είναι άξια της εμπιστοσύνης του κοινού. Αυτό σημαίνει ότι οι εφαρμογές της ΤΝ δεν θα πρέπει απλώς να συνάδουν με τη νομοθεσία, αλλά και να τηρούν δεοντολογικές αρχές και να διασφαλίζουν ότι η χρήση τους δεν προκαλεί ακούσια βλάβη. Οι εφαρμογές της ΤΝ θα πρέπει να ενδυναμώνουν τους πολίτες και να σέβονται τα θεμελιώδη δικαιώματά τους. Στόχος τους θα πρέπει να είναι η ενίσχυση των ικανοτήτων των πολιτών, και όχι η υποκατάστασή τους, καθώς και η διευκόλυνση της πρόσβασης των ατόμων με αναπηρία.

Σύμφωνα με τις “Κατευθυντήριες γραμμές”⁴ για την επίτευξη αξιοπιστίας η ΤΝ πρέπει:

1. να συμμορφώνεται με τη νομοθεσία,
2. να τηρεί τις αρχές δεοντολογίας και τις αξίες,
3. να είναι στιβαρή, τόσο από τεχνικής όσο και από κοινωνικής άποψης διότι τα συστήματα ΤΝ μπορούν να προκαλέσουν ακούσια βλάβη.

Με βάση τις προαναφερθείσες τρεις συνιστώσες και τις ευρωπαϊκές αξίες, που μνημονεύονται στις “Κατευθυντήριες γραμμές”, προσδιορίζονται επτά βασικές απαιτήσεις τις οποίες θα πρέπει να τηρούν οι εφαρμογές ΤΝ προκειμένου να θεωρούνται αξιόπιστες. Οι επτά βασικές απαιτήσεις είναι οι εξής:

1. Ανθρώπινη παρέμβαση και εποπτεία
2. Τεχνική αρτιότητα και ασφάλεια
3. Διακυβέρνηση της ιδιωτικής ζωής και των δεδομένων
4. Διαφάνεια
5. Σεβασμός στη διαφορετικότητα, απαγόρευση των διακρίσεων και δικαιοσύνη
6. Κοινωνική και περιβαλλοντική ευημερία
7. Λογοδοσία

4.1.2 Αξιόπιστη τεχνητή νοημοσύνη

Για να επιτευχθεί αυτό, θα πρέπει να διασφαλιστεί η αξιοπιστία της ΤΝ ώστε η ανάπτυξή της να ενσωματώνει τις αξίες στις οποίες βασίζονται οι κοινωνίες μας και τις οποίες διακηρύσσει η Ευρωπαϊκή Ένωση. Είναι οι αξίες του σεβασμού της ανθρώπινης αξιοπρέπειας, της ελευθερίας, της δημοκρατίας, της ισότητας, του κράτους δικαίου και του σεβασμού των δικαιωμάτων του ανθρώπου, συμπεριλαμβανομένων των δικαιωμάτων των ατόμων που ανήκουν σε μειονότητες. Οι αξίες αυτές διέπουν τις κοινωνίες όλων των κρατών μελών, οι οποίες χαρακτηρίζονται από πλουραλισμό, απαγόρευση των διακρίσεων, ανεκτικότητα, δικαιοσύνη, αλληλεγγύη και ισότητα. Ο Χάρτης των Θεμελιωδών Δικαιωμάτων⁵ της ΕΕ συγκεντρώνει σε ένα ενιαίο κείμενο όλα τα προσωπικά, ατομικά, πολιτικά, οικονομικά και κοινωνικά δικαιώματα των πολιτών, τα οποία τα κράτη μέλη της ΕΕ και τα θεσμικά όργανα της ΕΕ υποχρεούνται νομικά να σέβονται κατά την εφαρμογή της νομοθεσίας της ΕΕ.

Η κοινή βάση που συνενώνει τα εν λόγω δικαιώματα μπορεί να θεωρηθεί ότι εδράζεται στον σεβασμό για την ανθρώπινη αξιοπρέπεια, αποτυπώνοντας αυτό που περιγράφεται ως “ανθρωποκεντρική προσέγγιση”, σύμφωνα με την οποία ο άνθρωπος απολαμβάνει τη μοναδική και αναφαίρετη

⁴Κατευθυντήριες γραμμές δεοντολογίας για αξιόπιστη Τεχνητή Νοημοσύνη, High Level Expert Group, 2019, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60424

⁵Χάρτης των Θεμελιωδών Δικαιωμάτων της Ευρωπαϊκής Ένωσης, <https://eur-lex.europa.eu/legal-content/EL/TXT/?uri=CELEX:12012P/TXT>

ηθική υπόσταση υπεροχής στους τομείς των ατομικών, πολιτικών, οικονομικών και κοινωνικών δικαιωμάτων. Αυτό συνεπάγεται επίσης ότι λαμβάνεται υπόψη το φυσικό περιβάλλον και τα υπόλοιπα έμβια όντα που αποτελούν μέρος του ανθρώπινου οικοσυστήματος, καθώς και μια βιώσιμη προσέγγιση που καθιστά δυνατή την ευημερία των μελλοντικών γενεών.

4.1.3 Ανθρωποκεντρική ΤΝ και η αρχή του σεβασμού της ανθρώπινης αυτονομίας

Τα θεμελιώδη δικαιώματα στα οποία βασίζεται η ΕΕ έχουν γνώμονα τον σεβασμό της ελευθερίας και της αυτονομίας των ανθρώπων. Στην ανθρωποκεντρική ΤΝ οι άνθρωποι που αλληλεπιδρούν με τα συστήματα ΤΝ πρέπει να είναι σε θέση να διατηρούν πλήρη και αποτελεσματική αυτοδιάθεση και να είναι σε θέση να συμμετέχουν στη δημοκρατική διαδικασία⁶. Τα συστήματα ΤΝ δεν θα πρέπει αδικαιολόγητα να προβαίνουν σε καθυπόταξη, εξαναγκασμό, εξαπάτηση, χειραγώγηση, κοινωνικό προγραμματισμό ή αγελοποίηση των ανθρώπων⁷. Αντ' αυτού, τα συστήματα ΤΝ θα πρέπει να σχεδιάζονται έτσι ώστε να αυξάνουν, να συμπληρώνουν και να ενισχύουν τις ανθρώπινες γνωστικές, κοινωνικές και πολιτισμικές δεξιότητες.

Η κατανομή των λειτουργιών μεταξύ των ανθρώπων και των συστημάτων ΤΝ θα πρέπει να ακολουθεί τις αρχές του ανθρωποκεντρικού σχεδιασμού και να προσφέρει ουσιαστικές ευκαιρίες για ανθρώπινη επιλογή. Αυτό σημαίνει εξασφάλιση της ανθρώπινης εποπτείας και ελέγχου των διαδικασιών εργασίας στα συστήματα ΤΝ. Εκτιμάται επίσης ότι τα συστήματα ΤΝ θα μετασηματίσουν ουσιαστικά τον τομέα της εργασίας, επομένως θα πρέπει με τη λειτουργία τους να στηρίζουν τους ανθρώπους στο εργασιακό περιβάλλον και να αποσκοπούν στη δημιουργία ουσιαστικού έργου.

4.1.4 Ο κατάλογος αυτοαξιολόγησης για την αξιόπιστη ΤΝ (ALTAI)

Σε συνέχεια των ανωτέρω και στο πλαίσιο της αποστολής που ανατέθηκε στην Ομάδα εμπειρογνομόνων υψηλού επιπέδου για την τεχνητή νοημοσύνη (HLEG), αρχικά εκπονήθηκαν οι “Κατευθυντήριες γραμμές δεοντολογίας για αξιόπιστη τεχνητή νοημοσύνη” που παρουσιάστηκαν παραπάνω. Το έγγραφο προτείνει μια ανθρωποκεντρική προσέγγιση όσον αφορά την ΤΝ και απαριθμεί 7 βασικές απαιτήσεις που θα πρέπει να πληρούν τα συστήματα ΤΝ προκειμένου να είναι αξιόπιστα, τις οποίες ήδη αναλύσαμε. Στη βάση αυτών, η Ομάδα HLEG προχώρησε στη δημιουργία ενός πρακτικού εργαλείου που μεταφράζει τις Κατευθυντήριες Γραμμές Δεοντολογίας σε έναν προσβάσιμο και δυναμικό κατάλογο αυτοαξιολόγησης. Ο κατάλογος αυτός μπορεί να χρησιμοποιηθεί από προγραμματιστές και φορείς ανάπτυξης τεχνητής νοημοσύνης που επιθυμούν να εφαρμόσουν τις βασικές απαιτήσεις. Αυτή η νέα λίστα είναι διαθέσιμη επίσης ως διαδικτυακή εφαρμογή⁸ και σε μορφή PDF⁹.

4.2 Ειδικότερα για τα Θεμελιώδη Δικαιώματα

Δεδομένου ότι γίνεται συχνή επίκληση στην προστασία των θεμελιωδών δικαιωμάτων τα οποία διασφαλίζει η προτεινόμενη “Πράξη για την ΤΝ” και σε συνέχεια της πρόβλεψης σχετικής

⁶“The principle of respect for human autonomy: “...Humans interacting with AI systems must be able to keep full and effective self-determination over themselves, and be able to partake in the democratic process”, Κατευθυντήριες γραμμές δεοντολογίας για αξιόπιστη Τεχνητή Νοημοσύνη, High Level Expert Group, 2019, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60424

⁷“AI systems should not unjustifiably subordinate, coerce, deceive, manipulate, condition or herd humans”, Κατευθυντήριες γραμμές δεοντολογίας για αξιόπιστη Τεχνητή Νοημοσύνη, High Level Expert Group, 2019, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60424

⁸<https://futurium.ec.europa.eu/en/european-ai-alliance/pages/welcome-altai-portal>

⁹https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=68342 (αναλυτικό) και https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60440 (συνοπτικό)

αρμοδιότητας των παρατηρητηρίων TN, ακολουθεί απόσπασμα από τον προτεινόμενο κανονισμό με σύντομη σχετική ανάλυση.

4.2.1 Τα θεμελιώδη δικαιώματα στην αιτιολογική έκθεση της Πράξης για την TN

Στο κεφάλαιο 3.5 της Αιτιολογικής Έκθεσης της Πρότασης της Πράξης για την TN και υπό τον τίτλο “Θεμελιώδη Δικαιώματα” (μέρος του τίτλου 3. “Αποτελέσματα των εκ των υστέρων αξιολογήσεων, των διαβουλεύσεων με τα ενδιαφερόμενα μέρη και των εκτιμήσεων επιπτώσεων”) αναφέρεται ότι:

“Η χρήση της TN με τα ιδιαίτερα χαρακτηριστικά της (π.χ. αδιαφάνεια, πολυπλοκότητα, εξάρτηση από δεδομένα, αυτόνομη συμπεριφορά) μπορεί να επηρεάσει αρνητικά ορισμένα θεμελιώδη δικαιώματα που κατοχυρώνονται στον Χάρτη των Θεμελιωδών Δικαιωμάτων της ΕΕ¹⁰. Με ένα σύνολο απαιτήσεων για αξιόπιστη TN και συνακόλουθες υποχρεώσεις για όλους τους συμμετέχοντες στην αξιακή αλυσίδα, η πρόταση θα ενισχύσει και θα προωθήσει την προάσπιση των δικαιωμάτων που προστατεύονται από τον Χάρτη, όπως:

- το δικαίωμα στην ανθρώπινη αξιοπρέπεια (άρθρο 1),
- τον σεβασμό της ιδιωτικής ζωής και την προστασία των δεδομένων προσωπικού χαρακτήρα (άρθρα 7 και 8),
- την απαγόρευση των διακρίσεων (άρθρο 21),
- την ισότητα γυναικών και ανδρών (άρθρο 23),
- τα δικαιώματα της ελευθερίας της έκφρασης (άρθρο 11) και της ελευθερίας του συνέρχεσθαι (άρθρο 12),
- το δικαίωμα πραγματικής προσφυγής και αμερόληπτου δικαστηρίου,
- τα δικαιώματα της υπεράσπισης και του τεκμηρίου αθωότητας (άρθρα 47 και 48),
- τη γενική αρχή της χρηστής διοίκησης,
- τα δικαιώματα των εργαζομένων σε δίκαιες και πρόσφορες συνθήκες εργασίας (άρθρο 31),
- υψηλό επίπεδο προστασίας των καταναλωτών (άρθρο 28),
- τα δικαιώματα του παιδιού (άρθρο 24) και
- την ένταξη των ατόμων με αναπηρίες (άρθρο 26),
- το δικαίωμα προστασίας του περιβάλλοντος και βελτίωσης της ποιότητας του περιβάλλοντος (άρθρο 37), καθώς έχει σημασία, μεταξύ άλλων και για την υγεία και την ασφάλεια των ανθρώπων.

Οι υποχρεώσεις για εκ των προτέρων δοκιμές, διαχείριση κινδύνου και ανθρώπινη εποπτεία θα διευκολύνουν επίσης τον σεβασμό άλλων θεμελιωδών δικαιωμάτων, ελαχιστοποιώντας τον κίνδυνο εσφαλμένων ή μεροληπτικών αποφάσεων με τη βοήθεια TN σε κρίσιμους τομείς όπως η εκπαίδευση και η κατάρτιση, η απασχόληση, οι σημαντικές υπηρεσίες, η επιβολή του νόμου και η δικαιοσύνη. Σε περίπτωση που εξακολουθούν να υφίστανται παραβιάσεις των θεμελιωδών δικαιωμάτων, θα καταστεί δυνατή η αποτελεσματική προσφυγή για τα θιγόμενα πρόσωπα με τη διασφάλιση της διαφάνειας και της ιχνηλασιμότητας των συστημάτων TN σε συνδυασμό με αυστηρούς εκ των υστέρων ελέγχους.

¹⁰https://eur-lex.europa.eu/eli/treaty/char_2016/oj

Τίτλος Χάρτη	Άρθρο	Περιγραφή δικαιώματος	
I - ΑΞΙΟΠΡΕΠΕΙΑ	1	ανθρώπινη αξιοπρέπεια	ΠΡΟΣΤΑΙΝΕΤΑΙ
II - ΕΛΕΥΘΕΡΙΕΣ	7, 8	σεβασμός της ιδιωτικής ζωής και προστασία των δεδομένων προσωπικού χαρακτήρα	
	11, 12	ελευθερία της έκφρασης και ελευθερία του συνέρχεσθαι	
III - ΙΣΟΤΗΤΑ	21	απαγόρευση των διακρίσεων	
	23	ισότητα γυναικών και ανδρών	
	24	τα δικαιώματα του παιδιού	
	26	ένταξη των ατόμων με αναπηρίες	
IV - ΑΛΛΗΛΕΓΓΥΗ	28	υψηλό επίπεδο προστασίας των καταναλωτών	
	31	δικαιώματα των εργαζομένων σε δίκαιες και πρόσφορες συνθήκες εργασίας	
	37	προστασία και βελτίωση ποιότητας του περιβάλλοντος (→ υγεία και ασφάλεια)	
V - ΔΙΚΑΙΩΜΑΤΑ ΤΩΝ ΠΟΛΙΤΩΝ	41	χρηστή διοίκηση	ΠΕΡΙΟΡΙΖΕΤΑΙ
VI - ΔΙΚΑΙΟΣΥΝΗ	47, 48	διασφάλιση πραγματικής προσφυγής και αμερόληπτου δικαστηρίου, υπεράσπιση και τεκμήριο αθωότητας	
II - ΕΛΕΥΘΕΡΙΕΣ	13	ελευθερία της τέχνης και της επιστήμης	
	16	επιχειρηματική ελευθερία	
	17 §2	προστασία της διανοητικής ιδιοκτησίας	

Εικόνα 4.1. Τα δικαιώματα του Χάρτη Θεμελιωδών Δικαιωμάτων που άπτονται της Πράξης ΤΝ

Η παρούσα πρόταση¹¹ επιβάλλει ορισμένους περιορισμούς στην επιχειρηματική ελευθερία (άρθρο 16) και στην ελευθερία της τέχνης και της επιστήμης (άρθρο 13) για τη διασφάλιση της συμμόρφωσης με επιτακτικούς λόγους δημόσιου συμφέροντος, όπως η υγεία, η ασφάλεια, η προστασία των καταναλωτών και η προστασία άλλων θεμελιωδών δικαιωμάτων (“υπεύθυνη καινοτομία”), κατά την ανάπτυξη και τη χρήση τεχνολογίας ΤΝ υψηλού κινδύνου. Οι περιορισμοί αυτοί είναι αναλογικοί και περιορίζονται στο ελάχιστο αναγκαίο για την πρόληψη και τον μετριασμό των σοβαρών κινδύνων ασφάλειας και των πιθανών παραβιάσεων των θεμελιωδών δικαιωμάτων.

Επίσης, οι αυξημένες υποχρεώσεις διαφάνειας δεν θα επηρεάσουν δυσανάλογα το δικαίωμα προστασίας της διανοητικής ιδιοκτησίας (άρθρο 17 παράγραφος 2), δεδομένου ότι θα περιορίζονται μόνο στις ελάχιστες απαραίτητες πληροφορίες ώστε τα άτομα να ασκούν το δικαίωμά τους σε πραγματική προσφυγή και στην αναγκαία διαφάνεια έναντι των αρχών εποπτείας και επιβολής, σύμφωνα με τις αντίστοιχες εντολές τους. Κάθε αποκάλυψη πληροφοριών θα πραγματοποιείται σύμφωνα με τη σχετική νομοθεσία στον τομέα, συμπεριλαμβανομένης της οδηγίας 2016/943 περί προστασίας της τεχνολογίας και των επιχειρηματικών πληροφοριών που δεν έχουν αποκαλυφθεί (εμπορικό απόρρητο) από την παράνομη απόκτηση, χρήση και αποκάλυψή τους. Όταν οι δημόσιες αρχές και οι κοινοποιημένοι οργανισμοί χρειαστεί να αποκτήσουν πρόσβαση σε εμπιστευτι-

¹¹(Σ.σ.) Αναφέρεται στο αρχικά κατατεθέν νομοθετικό κείμενο ως “Πρόταση” Κανονισμού. Βλ. τίτλο του κειμένου: “Πρόταση Κανονισμός του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου για τη θέσπιση εναρμονισμένων κανόνων σχετικά με την Τεχνητή Νοημοσύνη (Πράξη για την Τεχνητή Νοημοσύνη) και για την τροποποίηση ορισμένων νομοθετικών πράξεων της Ένωσης”

κές πληροφορίες ή πηγαίο κώδικα για να εξετάσουν τη συμμόρφωση με ουσιαστικές υποχρεώσεις, υπόκεινται σε δεσμευτικές υποχρεώσεις εμπιστευτικότητας.”

Αν και διευκρινίστηκε ότι η παρούσα εργασία επικεντρώθηκε στο αρχικό κείμενο της προταθείσας “Πράξης για την ΤΝ”, παρακολούθησε ωστόσο τις προκλήσεις ως προς τα θεμελιώδη δικαιώματα που έθεσαν οι διαφορετικές πλευρές κατά τη διαδικασία των τριμερών συζητήσεων (trilogues). Σε αυτό το σημείο λοιπόν θα ήταν παράλειψη να μην αναφερθεί ότι το τελικό κείμενο του Κανονισμού, όπως δημοσιεύθηκε τον Ιούλιο 2024, υιοθέτησε σε σημαντικό βαθμό θέσεις του Ευρωπαϊκού Κοινοβουλίου το οποίο διαδραμάτισε καθοριστικό ρόλο στην ενίσχυση της προστασίας των θεμελιωδών δικαιωμάτων. Αυτό είχε αποτέλεσμα τη θέσπιση της εκτίμησης επιπτώσεων στα θεμελιώδη δικαιώματα (Fundamental Rights Impact Assessment - FRIA), κάτι που δεν είχε προβλεφθεί στο αρχικό κείμενο της προταθείσας “Πράξης για την ΤΝ”, την οποία πρέπει να διενεργούν οι φορείς εφαρμογής συστημάτων ΤΝ υψηλού κινδύνου που είναι οργανισμοί δημοσίου δικαίου ή ιδιωτικές οντότητες που παρέχουν δημόσιες υπηρεσίες είτε φορείς εφαρμογής που αναπτύσσουν συστήματα ΤΝ υψηλού κινδύνου όπως τραπεζικές ή ασφαλιστικές οντότητες, προτού θέσουν το σύστημα σε χρήση.

Επίσης η “Σύμβαση Πλαίσιο για την Τεχνητή Νοημοσύνη” του Συμβουλίου της Ευρώπης¹² απαιτεί από τα κράτη-μέλη να υιοθετήσουν νομοθετικά μέτρα για να εξασφαλίσουν ότι οι δραστηριότητες εντός του κύκλου ζωής των συστημάτων τεχνητής νοημοσύνης συνάδουν πλήρως με τα ανθρώπινα δικαιώματα, τη δημοκρατία και το κράτος δικαίου. Στο πλαίσιο αυτό παρέχεται ένα νέο εργαλείο από το Συμβούλιο της Ευρώπης με καθοδήγηση και δομημένη προσέγγιση για τη διεξαγωγή αξιολογήσεων κινδύνου και επιπτώσεων για συστήματα Τεχνητής Νοημοσύνης. Πρόκειται για τη μεθοδολογία HUDERIA¹³ η οποία είναι ειδικά προσαρμοσμένη για την προστασία και την προώθηση των ανθρωπίνων δικαιωμάτων, της δημοκρατίας και του κράτους δικαίου. Μπορεί να χρησιμοποιηθεί τόσο από δημόσιους όσο και από ιδιωτικούς φορείς για να βοηθήσει στον εντοπισμό και την αντιμετώπιση κινδύνων και επιπτώσεων στα ανθρώπινα δικαιώματα, τη δημοκρατία και το κράτος δικαίου καθ’ όλη τη διάρκεια του κύκλου ζωής των συστημάτων τεχνητής νοημοσύνης.

4.2.2 FRA - European Agency for Fundamental Rights

Στο σημείο αυτό επισημαίνεται επιπλέον ο σημαντικός ρόλος του FRA-European Agency for Fundamental Rights¹⁴, ενός ανεξάρτητου κέντρου αναφοράς και αριστείας για την προώθηση και την προστασία των ανθρωπίνων δικαιωμάτων στην ΕΕ που μάλιστα συμμετέχει στην αντιπροσωπεία της ΕΕ ενώπιον της ολομέλειας της Επιτροπής Τεχνητής Νοημοσύνης (CAI) του Συμβουλίου της Ευρώπης¹⁵. Μια εύληπτη οπτικοποίηση των ανθρωπίνων δικαιωμάτων σε σχέση με την ΤΝ, αποτυπώθηκε από το FRA στην ακόλουθη εικόνα με τίτλο **Θεμελιώδη δικαιώματα στην τεχνητή νοημοσύνη: Τι πρέπει να ληφθεί υπόψη**¹⁶

Η απεικόνιση αυτή συνοψίζει την έκθεση του Οργανισμού Θεμελιωδών Δικαιωμάτων της ΕΕ (FRA), σύμφωνα με τον οποίο διαπιστώνεται μια σύγχυση σχετικά με τον αντίκτυπο της τεχνητής νοημοσύνης στα ανθρώπινα δικαιώματα, ακόμη και μεταξύ των οργανισμών που ήδη την χρησιμοποιούν. Προτρέπει συνεπώς τους υπεύθυνους χάραξης πολιτικής να παράσχουν περισσότερη

¹²Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law, <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>


¹³HUDERIA: New tool to assess the impact of AI systems on human rights, <https://www.coe.int/en/web/portal/-/huderia-new-tool-to-assess-the-impact-of-ai-systems-on-human-rights>

¹⁴<https://fra.europa.eu/en/about-fra>

¹⁵<https://fra.europa.eu/en/news/2023/fra-council-europe-ai-committee-plenary>

¹⁶Έκθεση του FRA - Getting the future right – Τεχνητή νοημοσύνη και θεμελιώδη δικαιώματα στην ΕΕ, <https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-rights>

καθοδήγηση σχετικά με τον τρόπο εφαρμογής των υφιστάμενων κανόνων για την τεχνητή νοημοσύνη και να διασφαλίσουν ότι τυχόν μελλοντικοί νόμοι για την τεχνητή νοημοσύνη προστατεύουν τα θεμελιώδη δικαιώματα. Η έκθεση, που δημοσιεύθηκε το 2020, αποτελεί μέρος του έργου του FRA για την τεχνητή νοημοσύνη και τα μεγάλα δεδομένα και βασίζεται σε περισσότερες από 100 συνεντεύξεις με δημόσιους και ιδιωτικούς οργανισμούς από την Εσθονία, τη Φινλανδία, τη Γαλλία, την Ολλανδία και την Ισπανία, οι οποίοι χρησιμοποιούν ΤΝ.

Θεμελιώδη δικαιώματα στην ΤΝ: Τι πρέπει να ληφθεί υπόψη	
Είναι συμβατό;	Ο σχεδιασμός και η χρήση πρέπει να συμμορφώνονται με τους σχετικούς νόμους.
	Κάθε επεξεργασία δεδομένων πρέπει να σέβεται τους νόμους περί προστασίας δεδομένων.
	Εξετάζει τον ευρύτερο αντίκτυπο σε άλλα δικαιώματα.
Είναι δίκαιο;	Δεν κάνει διακρίσεις για λόγους όπως εθνικότητα, ηλικία, αναπηρία, φύλο και σεξουαλικός προσανατολισμός.
	Σέβεται τα δικαιώματα των παιδιών, των ηλικιωμένων και των ατόμων με αναπηρία.
Μπορεί να αμφισβητηθεί;	Οι άνθρωποι γνωρίζουν ότι χρησιμοποιείται η Τεχνητή Νοημοσύνη
	Οι άνθρωποι μπορούν να διαμαρτύρονται για τις αποφάσεις της ΤΝ.
	Οι αποφάσεις που βασίζονται στο σύστημα μπορούν να εξηγηθούν.
Μπορεί να ελεγχθεί;	Αξιολόγηση και τακτική επανεξέταση της χρήσης της ΤΝ για θέματα θεμελιωδών δικαιωμάτων.
	Οι άνθρωποι που εφαρμόζουν την ΤΝ μπορούν να περιγράψουν το σύστημα, τον στόχο του και τα δεδομένα που χρησιμοποιούνται.
Συμμετέχουν εξωτερικοί εμπειρογνώμονες;	Διαβούλευση με εμπειρογνώμονες και ενδιαφερόμενους φορείς.
	Επίβλεψη από εμπειρογνώμονες.

Εικόνα 4.2. Θεμελιώδη δικαιώματα στην τεχνητή νοημοσύνη: Τι πρέπει να ληφθεί υπόψη, FRA

Επισημαίνει ότι η ΤΝ δεν είναι αλάνθαστη, σχεδιάζεται και χρησιμοποιείται από ανθρώπους, και οι άνθρωποι μπορούν να κάνουν λάθη. Αυτός είναι ο λόγος για τον οποίο οι τυχόν θιγόμενοι πολίτες πρέπει να γνωρίζουν πότε χρησιμοποιείται η τεχνητή νοημοσύνη, πώς λειτουργεί και πώς να αμφισβητούν αυτοματοποιημένες αποφάσεις. Επίσης εντοπίζει αστοχίες στη χρήση της τεχνητής νοημοσύνης, για παράδειγμα στην αστυνόμευση, τις ιατρικές διαγνώσεις, τις κοινωνικές υπηρεσίες και τη στοχευμένη διαφήμιση και καλεί την ΕΕ και τις χώρες της ΕΕ να διαμορφώσουν την τεχνητή νοημοσύνη που όχι μόνο σέβεται τα ανθρώπινα και θεμελιώδη δικαιώματα αλλά επιπλέον τα προστατεύει και τα προωθεί.

Τονίζει κάτι που δεν είναι προφανές σε όλους, ότι δηλαδή η τεχνητή νοημοσύνη μπορεί να επηρεάσει πολλά δικαιώματα και όχι μόνο την ιδιωτικότητα ή την προστασία δεδομένων. Επειδή μπορεί να κάνει διακρίσεις ή να εμποδίσει τη δικαιοσύνη, θα πρέπει οποιαδήποτε νομοθεσία για την τεχνητή νοημοσύνη να το λαμβάνει υπόψη και να δημιουργεί αποτελεσματικές διασφαλίσεις ότι η χρήση της σέβεται όλα τα θεμελιώδη δικαιώματα. Στο πλαίσιο αυτό πρέπει να παρέχονται εγγυήσεις ότι οι ενδιαφερόμενοι μπορούν να αμφισβητήσουν τις αποφάσεις που λαμβάνονται από συστήματα τεχνητής νοημοσύνης, κάτι που προϋποθέτει ότι γνωρίζουν πότε και πώς χρησιμοποιεί-

ται η ΤΝ καθώς και με τι τρόπο και πού μπορούν να διαμαρτυρηθούν.

Συνεπώς οι οργανισμοί που χρησιμοποιούν ΤΝ πρέπει να είναι σε θέση να εξηγούν πώς τα συστήματά τους λαμβάνουν αποφάσεις, να σέβονται το δικαίωμα στην ανθρώπινη επανεξέταση όταν χρησιμοποιείται η τεχνητή νοημοσύνη και να καθιστούν σαφείς τις επιπτώσεις της αυτοματοποιημένης λήψης αποφάσεων. Επίσης προτρέπει ιδιωτικούς και δημόσιους οργανισμούς να αξιολογούν την ΤΝ πριν και κατά τη χρήση της για τη μείωση των αρνητικών επιπτώσεων στα θεμελιώδη δικαιώματα και τις τυχόν διακρίσεις στις οποίες μπορεί να οδηγήσει και να προβλέπουν ένα αποτελεσματικό σύστημα εποπτείας και ανθρώπινης επίβλεψης.

Μια συνοπτική περιπτωσιολογική αναφορά των κινδύνων της ΤΝ για τα θεμελιώδη δικαιώματα βρίσκεται στο Παράρτημα 1 της μελέτης της Ευρωπαϊκής Επιτροπής, και συγκεκριμένα της DG Communications Networks, Content and Technology, με τίτλο “Study to Support an Impact Assessment of Regulatory Requirements for Artificial Intelligence in Europe”¹⁷. Το παράρτημα εκτείνεται σε δέκα σελίδες και μπορεί να αναγνωστεί σε πλήρη μορφή, στη σελίδα 191 του εγγράφου που βρίσκεται στον σύνδεσμο <https://ec.europa.eu/newsroom/dae/redirection/document/75755>.

Ο πίνακας παρουσιάζει σε σύνοψη μια ενδεικτική επισκόπηση τεκμηριωμένων περιπτώσεων στις οποίες η ΤΝ ήταν εν μέρει υπεύθυνη για παραβίαση θεμελιωδών δικαιωμάτων. Οι περιπτώσεις ομαδοποιούνται σε κατηγορίες I - VI που αντιστοιχούν στις έξι βασικές κατηγορίες δικαιωμάτων του Χάρτη της ΕΕ, ήτοι:

- αξιοπρέπεια
- ελευθερίες
- ισότητα
- αλληλεγγύη
- δικαιώματα των πολιτών
- δικαιοσύνη

Επισημαίνεται ότι οι περισσότερες περιπτώσεις δεν ανήκουν αποκλειστικά σε μία κατηγορία, καθώς μπορεί να εγκυμονούν κινδύνους για περισσότερα του ενός θεμελιώδη δικαιώματα της ΕΕ. Επιπλέον επισημαίνεται ότι ο πίνακας περιλαμβάνει και παραδείγματα κρίσιμων συστημάτων ΤΝ που λειτουργούν εκτός του νομοθετικού πεδίου εφαρμογής της ΕΕ, καθώς αυτά είναι όλο και πιο διαδεδομένα και υπερβαίνουν τα εθνικά σύνορα. Αποτελείται από πέντε στήλες όπου στην πρώτη περιγράφεται το περιστατικό και παρέχεται παραπομπή με σύνδεσμο, στη δεύτερη περιγράφεται ο μακροπρόθεσμος κίνδυνος και το θιγόμενο δικαίωμα, στην τρίτη περιγράφονται τα εμπλεκόμενα φυσικά ή νομικά πρόσωπα είτε η κυβέρνηση, στην τέταρτη η προέλευση και χρήση των δεδομένων και στην πέμπτη η δυνατότητα και ο βαθμός παρέμβασης του χρήστη ή θιγόμενου προσώπου.

¹⁷DG CNECT, Study to Support an Impact Assessment of Regulatory Requirements for Artificial Intelligence in Europe, 21 April 2021, <https://digital-strategy.ec.europa.eu/en/library/study-supporting-impact-assessment-ai-regulation>

ΠΑΡΑΡΤΗΜΑ 1: ΠΕΡΙΛΗΨΗ ΤΩΝ ΚΙΝΔΥΝΩΝ ΤΗΣ ΤΝ ΓΙΑ ΤΑ ΘΕΜΕΛΙΩΔΗ ΔΙΚΑΙΩΜΑΤΑ

Παραδείγματα συστημάτων τεχνητής νοημοσύνης που ενέχουν (μακροπρόθεσμους) κινδύνους για τα θεμελιώδη δικαιώματα

Περιπτωση	Μακροπρόθεσμος κίνδυνος	Επιρροές	Προέλευση και χρήση των δεδομένων	Βαθμός παρέμβασης
Ι ΔΞΙΟΠΡΕΠΕΙΑ (Χάρτης Θεμελιωδών Δικαιωμάτων της ΕΕ, άρθρα 1-5: ανθρώπινη αξιοπρέπεια, δικαίωμα στη ζωή, δικαίωμα στην ακεραιότητα του προσώπου, απαγόρευση των βασανιστηρίων και της απάνθρωπης ή εξευτελιστικής μεταχείρισης ή τιμωρίας, απαγόρευση της δουλείας και της καταναγκαστικής εργασίας)				
<ul style="list-style-type: none"> Το project "Blue Feed, Red Feed" δείχνει πώς οι ροές του Facebook απεικονίζουν διαφορετικές πραγματικότητες με βάση τις πολιτικές τους προτιμήσεις. Το TikTok χρησιμοποιεί την αναγνώριση προσώπου για να αναλύει τις εικόνες προφίλ για να προτείνει νέους λογαριασμούς με βάση τη φυσική εμφάνιση των ατόμων που ήδη ακολουθεί ο χρήστης. 	Υψηλοί μακροπρόθεσμοι κίνδυνοι για την πολυφωνία των απόψεων και το δικαίωμα στην ψυχική ασφάλεια και ακεραιότητα.	B2C	Τα δεδομένα παρέχονται ως επί το πλείστον οικειοθελώς, επειδή οι χρήστες επιλέγουν την υπηρεσία/δίκτυο. Ωστόσο, δεν υπάρχουν εναλλακτικά δίκτυα. Ως επί το πλείστον, δεν υπάρχει δυνατότητα εξαίρεσης, καθώς τα δεδομένα καταγράφονται από τη χρήση.	Χαμηλός/κανένας βαθμός παρέμβασης για τους χρήστες. Δεν υπάρχει ιχνηλασιμότητα των δεδομένων (επαναχρησιμοποίηση).

Εικόνα 4.3. Απόσπασμα από την “Περίληψη των κινδύνων της ΤΝ για τα θεμελιώδη δικαιώματα”, Παράρτημα μελέτης DG CNECT

4.3 Παρατηρήτρια: τα κείμενα της ΕΕ προβλέπουν

Το 2018 η Ευρωπαϊκή Επιτροπή (σε συντονισμό με τα κράτη μέλη) ανέπτυξε και εγκαινίασε το Παρατηρητήριο για την ΤΝ (AI Watch) προκειμένου να παρακολουθεί τις εξελίξεις που σχετίζονται με τις τεχνολογίες ΤΝ¹⁸ συνολικά σε επίπεδο Ευρωπαϊκής Ένωσης. Το AI Watch παρακολουθεί τη βιομηχανική, τεχνολογική και ερευνητική ικανότητα, τις πρωτοβουλίες χάραξης πολιτικής για την ΤΝ από τα κράτη μέλη, τις επενδύσεις, τις δεξιότητες στην ΤΝ, τις εξελίξεις για την υιοθέτηση της ΤΝ, τον αντίκτυπό της στην οικονομία, στην κοινωνία και στις δημόσιες υπηρεσίες.

Κατά τα πρώτα δύο έτη, 2018-2020, δημοσίευσε μελέτες προκειμένου να υποστηρίξει την τεκμηριωμένη χάραξη πολιτικής και να τροφοδοτήσει με πληροφορίες τον δημόσιο διάλογο σχετικά με την ΤΝ. Σε ό,τι αφορά τη γενικότερη παρακολούθηση της υιοθέτησης και της χρήσης, σε ολόκληρη την Ευρώπη, αναδυόμενων τεχνολογιών για την παροχή δημόσιων υπηρεσιών, συμπεριλαμβανομένων των τεχνολογιών ΤΝ, η Επιτροπή δημιούργησε το 2018 μια πλατφόρμα καινοτόμων δημόσιων υπηρεσιών (IPSO) η οποία στο μεταξύ μετασηματίστηκε στη “Διαλειτουργική Ευρώπη”¹⁹, στην οποία υπάγεται το NIFO²⁰ (Εθνικό Παρατηρητήριο Πλαισίου Διαλειτουργικότητας). Το NIFO είναι ένας από τους μηχανισμούς για την τακτική συλλογή πληροφοριών σχετικά με την πορεία της ψηφιακής δημόσιας διοίκησης και διαλειτουργικότητας εντός της ΕΕ και των συνδεδεμένων χωρών. Στη συνέχεια, το 2020 η Eurostat άρχισε να συλλέγει δεδομένα σχετικά με την υιοθέτηση της ΤΝ στην ΕΕ από επιχειρήσεις²¹.

Ωστόσο οι απαιτήσεις των καιρών επιβάλλουν εγρήγορση, διάχυτη εποπτεία και διεπιστημονικότητα με αποτέλεσμα τη δημιουργία νέων φορέων όπως είναι το πρόσφατο European Centre for Algorithmic Transparency (ECAT)²². Το ECAT αποτελεί μέρος της Ευρωπαϊκής Επιτροπής

¹⁸ Ανακοίνωση της Επιτροπής προς το Ευρωπαϊκό Κοινοβούλιο, το Συμβούλιο, την Ευρωπαϊκή Οικονομική και Κοινωνική Επιτροπή και την Επιτροπή των Περιφερειών, Προώθηση μιας ευρωπαϊκής προσέγγισης της τεχνητής νοημοσύνης, Βρυξέλλες, 21 Απριλίου 2021, <https://eur-lex.europa.eu/legal-content/EL/TXT/?uri=CELEX:52021DC0205>

¹⁹ <https://joinup.ec.europa.eu/interoperable-europe>

²⁰ About NIFO - National Interoperability Framework Observatory, <https://joinup.ec.europa.eu/collection/nifo-national-interoperability-framework-observatory/about>

²¹ <https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20210413-1>

²² Ευρωπαϊκό Κέντρο Αλγοριθμικής Διαφάνειας, <https://algorithmic-transparency.ec.europa.eu/>

και φιλοξενείται από το Κοινό Κέντρο Ερευνών (JRC) σε στενή συνεργασία με τη Γενική Διεύθυνση CONNECT (Directorate General for Communications Networks, Content and Technology). Συνεισφέρει την επιστημονική και τεχνική εμπειρογνώσια του στον αποκλειστικό εποπτικό και εκτελεστικό ρόλο της Ευρωπαϊκής Επιτροπής²³ όσον αφορά τις υποχρεώσεις που προβλέπονται από τον κανονισμό για τις ψηφιακές υπηρεσίες (DSA)²⁴, στο πλαίσιο του οποίου²⁵ απαιτούνται έλεγχοι αλγοριθμικής λογοδοσίας και διαφάνειας, ειδικά όσον αφορά τις πολύ μεγάλες επιγραμμικές πλατφόρμες (VLOP) και τις πολύ μεγάλες μηχανές αναζήτησης (VLOSE). Πόσο μάλλον όταν ο τρόπος με τον οποίο τα αλγοριθμικά συστήματα διαμορφώνουν την ορατότητα και την προώθηση του περιεχομένου, καθώς και ο κοινωνικός και ηθικός αντίκτυπος αυτού, αποτελούν τομείς αυξανόμενου προβληματισμού.

Τέλος, παρατηρείται μια συνεχής αύξηση της δημιουργίας παρατηρητηρίων ως εργαλείων παρακολούθησης για τη συλλογή πληροφοριών σχετικά με την κατάσταση της ψηφιακής δημόσιας διοίκησης και την προώθηση της διαλειτουργικότητας. Αυτά τα παρατηρητήρια²⁶, όπως το Public Sector Tech Watch, το LIFO και το OSOR, πέραν του προαναφερθέντος NIFO, δίνουν τη δυνατότητα στους ενδιαφερόμενους να παραμένουν ενημερωμένοι για τις τελευταίες εξελίξεις και τάσεις στις ψηφιακές δημόσιες υπηρεσίες.

Εξ αυτών, ειδικά το Public Sector Tech Watch (PSTW)²⁷ είναι ένα παρατηρητήριο αφιερωμένο στην παρακολούθηση, ανάλυση και διάδοση της χρήσης αναδυόμενων τεχνολογιών (π.χ. Blockchain, Τεχνητής Νοημοσύνης κ.λπ.) στον δημόσιο τομέα στην Ευρώπη. Στόχος του είναι να γίνει ο “μοναδικός σταθμός” για όλους τους ενδιαφερόμενους φορείς, στελέχη του δημόσιου τομέα, υπεύθυνους χάραξης πολιτικής, ιδιωτικές εταιρείες, ακαδημαϊκό κόσμο, που ενδιαφέρονται για τις τελευταίες τεχνολογικές εξελίξεις υπό το πρίσμα της βελτίωσης των λειτουργιών και της παροχής υπηρεσιών του δημόσιου τομέα. Το PSTW διοικείται από τη Directorate General for Informatics (DIGIT) και το Joint Research Centre (JRC) της Ευρωπαϊκής Επιτροπής.

4.3.1 Ο ρόλος και η λειτουργία ενός Παρατηρητηρίου TN στην ΕΕ

Στο έγγραφο εργασίας που συνέταξε η υπηρεσία της Γενικής Διεύθυνσης Επικοινωνιακών Δικτύων, Περιεχομένου και Τεχνολογιών (DG CONNECT), το οποίο συνόδευε την πρόταση κανονισμού για την “Πράξη για την TN”, επισημαίνεται²⁸ η σημασία της παρακολούθησης και αξιολόγησης των πραγματικών επιπτώσεων του προτεινόμενου κανονιστικού πλαισίου.

Τονίζεται ότι ένας καλά δομημένος μηχανισμός παρακολούθησης και αξιολόγησης είναι ιδιαίτερα σημαντικός για την αξιολόγηση του βαθμού στον οποίο το κανονιστικό πλαίσιο της TN επιτυγχάνει τους στόχους του. Σκοπός είναι τα συστήματα τεχνητής νοημοσύνης στην Ευρώπη να μην παραβιάζουν τα θεμελιώδη δικαιώματα των πολιτών και να μην προκαλούν βλάβη στην υγεία των ανθρώπων εξαιτίας της λειτουργίας τους. Επίσης η TN θα πρέπει να λειτουργεί προς όφελος των καταναλωτών και των επιχειρήσεων μέσα σε ένα πλαίσιο αλγοριθμικής διαφάνειας και λογοδοσίας που δεν θα καταπνίγει την καινοτομία και θα είναι αξιόπιστο ώστε να αποτελέσει ένα παγκόσμιο πρότυπο αναφοράς.

index_en

²³ Διημερίδα παρουσίασης του Ecat, 18-19 Απριλίου 2023, https://algorithmic-transparency.ec.europa.eu/presenting-ecat_en

²⁴ <https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package>

²⁵ Κανονισμός 2022/2065 του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου της 19ης Οκτωβρίου 2022 σχετικά με την ενιαία αγορά ψηφιακών υπηρεσιών, <https://eur-lex.europa.eu/legal-content/EL/TXT/?uri=CELEX:32022R2065>

²⁶ <https://joinup.ec.europa.eu/topic/digital-government/observatories>

²⁷ <https://joinup.ec.europa.eu/collection/public-sector-tech-watch>

²⁸ Commission, Staff working document impact assessment, Accompanying the 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, SWD(2021) 84 final, Brussels, 21 April 2021, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021SC0084>

Το ίδιο έγγραφο προτείνει να είναι υπεύθυνη η Ευρωπαϊκή Επιτροπή για την παρακολούθηση των αποτελεσμάτων της ρύθμισης για την ΤΝ που θα θεσπιστεί, δημιουργώντας ένα σύστημα καταγραφής αυτόνομων εφαρμογών ΤΝ επάνω σε μια δημόσια βάση δεδομένων σε επίπεδο ΕΕ, προσανατολισμένου κυρίως σε τυχόν επιπτώσεις στα θεμελιώδη δικαιώματα. Αυτό θα επιτρέψει επίσης στις αρμόδιες αρχές, τους χρήστες και άλλους ενδιαφερόμενους να επαληθεύουν αν ένα σύστημα ΤΝ υψηλού κινδύνου συμμορφώνεται με τις νέες απαιτήσεις, ώστε να ασκούν ενισχυμένη εποπτεία επί αυτών των εφαρμογών ΤΝ που θέτουν περιορισμούς στα θεμελιώδη δικαιώματα. Για την τροφοδότηση αυτής της βάσης δεδομένων, οι προμηθευτές ΤΝ θα είναι υποχρεωμένοι να παρέχουν ουσιαστικές πληροφορίες σχετικά με το σύστημα και την αξιολόγηση της συμμόρφωσης που διενεργήθηκε.

Επιπλέον, οι πάροχοι ΤΝ θα υποχρεούνται να ενημερώνουν τις εθνικές αρμόδιες αρχές για σοβαρά περιστατικά ή επιδόσεις ΤΝ που συνιστούν παραβίαση των θεμελιωδών δικαιωμάτων αμέσως μόλις λάβουν γνώση, καθώς και για τυχόν ανακλήσεις ή αποσύρσεις συστημάτων ΤΝ από την αγορά. Στη συνέχεια, οι εθνικές αρμόδιες αρχές θα διερευνούν τα περιστατικά και τις παραβάσεις, θα συλλέγουν όλες τις απαραίτητες πληροφορίες και θα τις διαβιβάζουν τακτικά με κατάλληλα μεταδεδομένα στο συμβούλιο της ΕΕ για την ΤΝ, κατανεμημένες ανά τομέα εφαρμογών (π.χ. προσλήψεις, βιομετρική αναγνώριση κλπ) και υπολογιζόμενες:

1. σε απόλυτους αριθμούς,
2. ως ποσοστό των εφαρμογών που αναπτύχθηκαν και
3. ως ποσοστό των πολιτών που αφορά.

Η Επιτροπή προτείνεται να συμπληρώνει αυτές τις πληροφορίες σχετικά με τις εφαρμοσμένες περιπτώσεις χρήσης ΤΝ υψηλού κινδύνου, με μια ολοκληρωμένη ανάλυση της συνολικής αγοράς τεχνητής νοημοσύνης. Για τον σκοπό αυτό, θα μετράει την υιοθέτηση της τεχνητής νοημοσύνης με τακτικές έρευνες με αφετηρία την πρώτη ποσοτική έρευνα που έγινε ποτέ σε επίπεδο ΕΕ σχετικά με την υιοθέτηση των τεχνολογιών ΤΝ από τις ευρωπαϊκές επιχειρήσεις και η οποία διεξήχθη από την Επιτροπή το πρώτο τρίμηνο του 2020²⁹. Συγκεκριμένα ερωτήθηκαν συνολικά 9.640 επιχειρήσεις σε 30 ευρωπαϊκές χώρες, την περίοδο Ιανουαρίου-Μαρτίου 2020. Για τις ανάγκες της έρευνας μετρήθηκαν οι εξής πέντε KPIs (βασικοί δείκτες επίδοσης) που αφορούν τη χρήση τεχνολογιών ΤΝ από τις επιχειρήσεις:

- Ενημέρωση για την τεχνητή νοημοσύνη
- Υιοθέτηση της ΤΝ
- Προμήθεια ΤΝ
- Εξωτερικά εμπόδια στην υιοθέτηση της ΤΝ
- Εσωτερικά εμπόδια στην υιοθέτηση της ΤΝ.

²⁹European Commission, Ipsos Survey, European enterprise survey on the use of technologies based on artificial intelligence, 2020, <https://digital-strategy.ec.europa.eu/en/library/european-enterprise-survey-use-technologies-based-artificial-intelligence>. Για την εν λόγω έρευνα διαβάζουμε στις παραπομπές:

α) 27 του εγγράφου Commission, Staff working document impact assessment, <httpst://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021SC0084> και

β) 36 του εγγράφου Ανακοίνωση της Επιτροπής προς το Ευρωπαϊκό Κοινοβούλιο, το Συμβούλιο, την Οικονομική και Κοινωνική Επιτροπή και την Επιτροπή των Περιφερειών, Προώθηση μιας ευρωπαϊκής προσέγγισης της τεχνητής νοημοσύνης, <https://eur-lex.europa.eu/legal-content/EL/TXT/?uri=CELEX:52021DC0205>

Επίσης το έγγραφο ορίζει ότι θα χρησιμοποιηθούν στοιχεία από τις αρμόδιες εθνικές αρχές, τη Eurostat, το Joint Research Center (Κοινό Κέντρο Ερευνών) μέσω του AI Watch και τον ΟΟΣΑ για την αξιολόγηση του βαθμού στον οποίο το κανονιστικό πλαίσιο επιτυγχάνει τους στόχους του, ώστε η Ευρωπαϊκή Ένωση και τα κράτη μέλη να αναπτύξουν από κοινού μια νέα κουλτούρα αλγοριθμικής διαφάνειας και λογοδοσίας χωρίς να καταπνίγουν την καινοτομία.

Επιπλέον επισημαίνει ότι θα δοθεί ιδιαίτερη προσοχή στη διεθνή συμβατότητα των συλλογών δεδομένων, ώστε τα δεδομένα να καταστούν συγκρίσιμα μεταξύ των κρατών μελών και άλλων προηγμένων οικονομιών. Το κοινό παρατηρητήριο του ΟΟΣΑ/ΕΕ για την τεχνητή νοημοσύνη αποτελεί ένα πρώτο βήμα προς την κατεύθυνση αυτή.

Για όλα τα παραπάνω θεσπίζεται η υποχρέωση των παρόχων ΤΝ να ενημερώνουν τις εθνικές αρμόδιες αρχές για σοβαρά περιστατικά ή επιδόσεις ΤΝ που συνιστούν παραβίαση των υποχρεώσεων θεμελιωδών δικαιωμάτων αμέσως μόλις λάβουν γνώση, καθώς και για τυχόν ανακλήσεις ή αποσύρσεις συστημάτων ΤΝ από την αγορά. Ειδικότερα οι πάροχοι συστημάτων ΤΝ υψηλού κινδύνου οφείλουν να καταρτίζουν δήλωση συμμόρφωσης ΕΕ και να τοποθετούν τη σήμανση συμμόρφωσης CE σύμφωνα με τα άρθρα 47 επ. της Πράξης για την ΤΝ. Συγκεκριμένα στα άρθρα αυτά γίνεται περαιτέρω εξειδίκευση και παραπομπή στο περιεχόμενο της δήλωσης συμμόρφωσης όπως ορίζεται στο παράρτημα V της Πράξης για την ΤΝ³⁰ και αφορά τα συστήματα υψηλού κινδύνου.

4.4 Τυποποίηση και πρότυπα

Σε συνέχεια των επτά απαιτήσεων, που προέρχονται από το έγγραφο “Κατευθυντήριες γραμμές δεοντολογίας για αξιόπιστη τεχνητή νοημοσύνη” όπως αναλύθηκαν παραπάνω και τις οποίες θα πρέπει να τηρούν τα συστήματα ΤΝ για να πραγματωθεί αξιόπιστη ΤΝ, επισημαίνεται ότι για την εφαρμογή τους μπορούν να χρησιμοποιούνται τόσο τεχνικές όσο και μη τεχνικές μέθοδοι. Ειδικότερα στις μη τεχνικές μεθόδους συγκαταλέγονται τα πρότυπα, π.χ. για τον σχεδιασμό, την κατασκευή και τις επιχειρηματικές πρακτικές, που μπορεί να λειτουργούν ως σύστημα διαχείρισης ποιότητας για τους χρήστες της ΤΝ, τους καταναλωτές, τους οργανισμούς, τα ερευνητικά ιδρύματα και τις κυβερνήσεις, προσφέροντας την ικανότητα να αναγνωρίζουν και να ενθαρρύνουν τη δεοντολογική συμπεριφορά. Πέρα από τα συμβατικά πρότυπα, υφίστανται προσεγγίσεις από κοινού ρύθμισης: συστήματα διαπίστευσης, επαγγελματικοί κώδικες δεοντολογίας ή πρότυπα για τον σχεδιασμό που συμμορφώνεται με τα θεμελιώδη δικαιώματα. Τρέχοντα παραδείγματα είναι π.χ. τα πρότυπα ISO ή η σειρά προτύπων IEEE 7000, αλλά στο μέλλον ενδέχεται να κριθεί κατάλληλη η χρήση κάποιας σήμανσης όπως, ενδεχομένως, «Αξιόπιστη ΤΝ», με την οποία θα επιβεβαιώνεται διά της αναφοράς σε συγκεκριμένα τεχνικά πρότυπα ότι, για παράδειγμα, το σύστημα συμμορφώνεται με τις αρχές της ασφάλειας, της τεχνικής στιβαρότητας και της επεξηγησιμότητας (παράγραφος 106, σελ. 29)³¹.

Πράγματι, η προτεινόμενη Πράξη για την ΤΝ δεν είναι ο πρώτος νόμος για τις ψηφιακές τεχνολογίες στην Ευρώπη, έχουν ήδη προηγηθεί ο Γενικός Κανονισμός για την Προστασία Δεδομένων (GDPR) 2016/679, ο Κανονισμός 2022/868 για τη Διακυβέρνηση Δεδομένων, ο Κανονισμός για τα Δεδομένα 2023/2854), η Πράξη για τις Ψηφιακές Αγορές (DMA, (2022/1925 και) και Πράξη για τις Ψηφιακές Υπηρεσίες (DSA, 2022/2065) για τη ρύθμιση των διαδικτυακών πλατφορμών. Ωστόσο, η Πράξη για την ΤΝ ακολουθεί διαφορετική διαδρομή από αυτά τα κείμενα, αντλώντας έμπνευση από τους ευρωπαϊκούς κανόνες ασφάλειας προϊόντων. Ειδικότερα, τα συστήματα τεχνητής νοημοσύνης απαιτούν αξιολόγηση συμμόρφωσης που θα βασίζεται σε εναρμονισμένα πρότυπα, δηλαδή τεχνικές προδιαγραφές που καταρτίζονται από Ευρωπαϊκούς οργανισμούς τυ-

³⁰<https://eur-lex.europa.eu/legal-content/EL/TXT/?uri=CELEX:52021PC0206>

³¹Κατευθυντήριες γραμμές δεοντολογίας για αξιόπιστη Τεχνητή Νοημοσύνη, High Level Expert Group, 2019, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60424

ποποίησης και διαθέτουν νομικές ιδιότητες, όπως η δημιουργία τεκμηρίου συμμόρφωσης με τη νομοθεσία³².

Αυτή η διαδικασία αξιολόγησης της συμμόρφωσης θα οδηγήσει στη συνέχεια στη σήμανση ευρωπαϊκής συμμόρφωσης (CE) του συστήματος τεχνητής νοημοσύνης. Όμως, σε αντίθεση με άλλους κανονισμούς για την ασφάλεια των προϊόντων, η Πράξη για την TN δεν αποσκοπεί μόνο στην προστασία από κινδύνους για την ασφάλεια, αλλά και από δυσμενείς επιπτώσεις στα θεμελιώδη δικαιώματα, και συνεπώς η διαφαινόμενη επέκταση των προτύπων στο πεδίο των θεμελιωδών δικαιωμάτων είναι νέα και εγείρει σύνθετα ερωτήματα.

4.4.1 Η σειρά προτύπων IEEE 7000 για δεοντολογικό σχεδιασμό και διακυβέρνηση της TN

Στην κατεύθυνση για σχεδιασμό συστημάτων TN με διαφάνεια, υπευθυνότητα και δικαιοσύνη βρίσκουμε την παγκόσμια πρωτοβουλία για τη δεοντολογία των αυτόνομων και ευφών συστημάτων (Global Initiative on Autonomous and Intelligent Systems) του Ινστιτούτου Ηλεκτρολόγων και Ηλεκτρονικών Μηχανικών (IEEE), ενός υπερεθνικού, πλέον, επιστημονικού οργανισμού με έδρα την Νέα Υόρκη³³. Στόχο έχει να διασφαλίσει ότι κάθε ενδιαφερόμενος που εμπλέκεται στο σχεδιασμό και την ανάπτυξη αυτόνομων και ευφών συστημάτων είναι εκπαιδευμένος και έχει τα απαραίτητα εφόδια για να δίνει προτεραιότητα σε ηθικά ζητήματα, έτσι ώστε αυτές οι τεχνολογίες να προωθούνται προς όφελος της ανθρωπότητας.

4.4.1.1 IEEE: Σύντομο ιστορικό και ο ρόλος του στη δεοντολογία για τον κοινωνικό αντίκτυπο της TN

Το Ινστιτούτο Ηλεκτρολόγων και Ηλεκτρονικών Μηχανικών (IEEE) ιδρύθηκε το 1963 μέσω της συγχώνευσης δύο αμερικανικών οργανισμών: του Institute of Radio Engineers (IRE) και του American Institute of Electrical Engineers (AIEE). Η συγχώνευση αυτή συγκέντρωσε επαγγελματίες στον τομέα της ηλεκτρολογίας και της ηλεκτρονικής, δημιουργώντας έναν ενιαίο φορέα αφιερωμένο στην προώθηση της τεχνολογίας για την ανθρωπότητα. Το IEEE έγινε γρήγορα κορυφαίος οργανισμός για τεχνικούς επαγγελματίες και μηχανικούς παγκοσμίως, καθορίζοντας πρότυπα και προωθώντας την καινοτομία σε διάφορους τεχνολογικούς τομείς.

Το 2016, το IEEE έκανε ένα σημαντικό βήμα στην αντιμετώπιση των ηθικών επιπτώσεων της τεχνολογίας με την εισαγωγή της έννοιας του Ευθυγραμμισμένου με τη Δεοντολογία Σχεδιασμού (Ethically Aligned Design, EAD). Αυτή η πρωτοβουλία στοχεύει στη διασφάλιση ότι ηθικά κριτήρια και παράμετροι ενσωματώνονται στο σχεδιασμό, την ανάπτυξη και την εφαρμογή των προηγμένων τεχνολογιών, ιδιαίτερα της τεχνητής νοημοσύνης (TN). Το πλαίσιο EAD³⁴ παρέχει κατευθυντήριες γραμμές για μηχανικούς και τεχνοεπιστήμονες ώστε να λαμβάνουν υπόψη τον κοινωνικό αντίκτυπο, τα ανθρώπινα δικαιώματα και τις ηθικές αρχές καθ' όλη τη διάρκεια των διαδικασιών σχεδιασμού και υλοποίησης των έργων τους.

Με βάση το πλαίσιο EAD, το IEEE ανέπτυξε τη σειρά προτύπων IEEE 7000. Αυτή η σειρά ασχολείται με συγκεκριμένες πτυχές των ηθικών εκτιμήσεων στο σχεδιασμό της τεχνολογίας. Για παράδειγμα, το IEEE 7010 παρέχει μετρήσεις για την αξιολόγηση του αντίκτυπου της TN στην ανθρώπινη ευημερία, ενώ το IEEE 7001 εστιάζει στη διαφάνεια στα αυτόνομα συστήματα. Αυτά τα πρότυπα προσφέρουν πρακτικά εργαλεία για τους μηχανικούς ώστε να ενσωματώσουν ηθικές

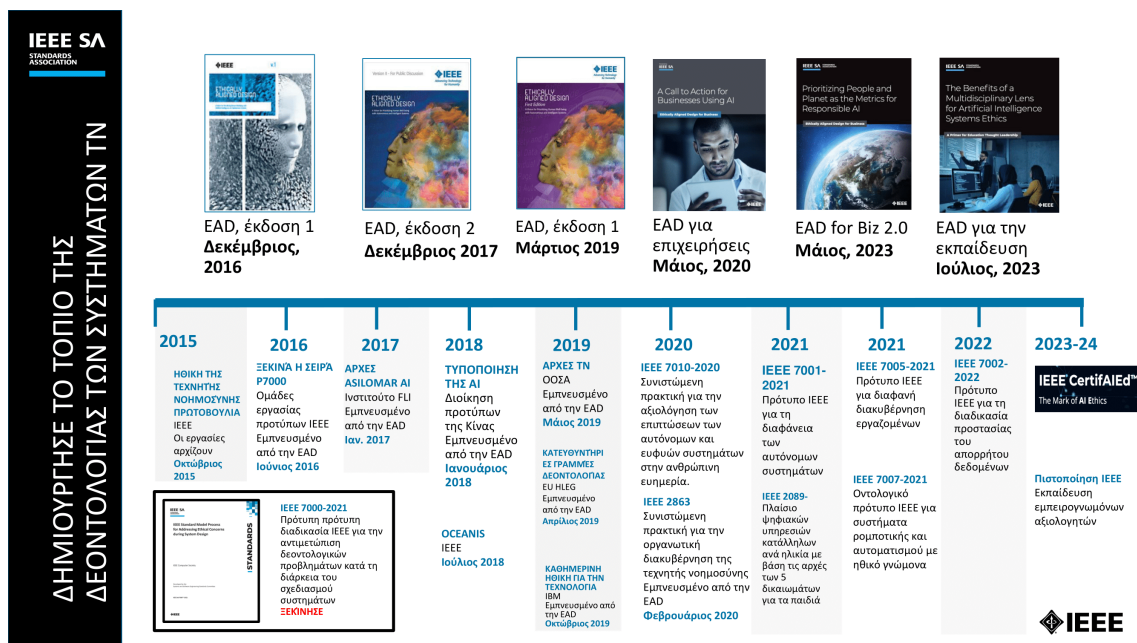
³²M. Gornet and W. Maxwell, 2024, The European approach to regulating AI through technical standards, Internet Policy Review,[online] 13(3), <https://policyreview.info/articles/analysis/regulating-ai-through-technical-standards>

³³<https://standards.ieee.org/initiatives/autonomous-intelligence-systems/>

³⁴The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, Version 2, IEEE, 2017, https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf

αρχές στη δουλειά τους, διασφαλίζοντας ότι η τεχνολογία εξυπηρετεί καλύτερα τα συμφέροντα της ανθρωπότητας.

Η σειρά IEEE 8000 είναι ένα αναδυόμενο έργο που στοχεύει στην περαιτέρω προώθηση των αρχών που καθιερώθηκαν με τη σειρά IEEE 7000. Τα πρότυπα IEEE 8000 θα καλύπτουν ευρύτερες πτυχές της ψηφιακής διακυβέρνησης και των δημόσιων προμηθειών, διασφαλίζοντας ότι η ΤΝ και άλλες προηγμένες τεχνολογίες ευθυγραμμίζονται με το δημόσιο συμφέρον και τις κοινωνικές αξίες. Αυτά τα πρότυπα αναμένεται να καθοδηγήσουν τις κυβερνήσεις και τους οργανισμούς στην υιοθέτηση και ρύθμιση της τεχνολογίας με υπευθυνότητα.



Εικόνα 4.4. Η επιρροή που άσκησε το EAD του IEEE σε άλλες πρωτοβουλίες ρύθμισης της ΤΝ

4.4.1.2 Το πρότυπο IEEE 7010-2020 για την αξιολόγηση των επιπτώσεων της ΤΝ στην ανθρώπινη ευημερία

Το πρότυπο 7010-2020 IEEE με τίτλο “Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being” (Συνιστώμενη πρακτική IEEE για την αξιολόγηση των επιπτώσεων των αυτόνομων και ευφύων συστημάτων στην ανθρώπινη ευημερία), βοηθά τους δημιουργούς συστημάτων ΤΝ και τους υπεύθυνους χάραξης πολιτικής καθώς προβλέπει μια μεθοδολογία αξιολόγησης των επιπτώσεων κατά την αφετηρία του σχεδιασμού. Βάσει αυτής απαιτείται από τους χρήστες να προσδιορίσουν παγκοσμίως επικυρωμένους δείκτες που εστιάζουν σε θέματα ανθρώπινης ευημερίας και περιβαλλοντικής βιωσιμότητας, συμπληρωματικά προς τις υπάρχουσες οικονομικές μετρήσεις και τις μετρήσεις ανάπτυξης. Αυτές συνήθως αποτελούσαν, μέχρι τώρα, τους μόνους βασικούς δείκτες για την ανάπτυξη και τη χρήση των συστημάτων ΤΝ, όπως θα εκτεθεί αναλυτικά παρακάτω, στην ενότητα για τους βασικούς δείκτες επίδοσης. Παγκόσμιοι δείκτες όπως ο Δείκτης Καλύτερης Ζωής του ΟΟΣΑ ή οι Στόχοι Βιώσιμης Ανάπτυξης (ΣΒΑ) του ΟΗΕ παρέχουν δεδομένα που όλοι οι σχεδιαστές συστημάτων ΤΝ μπορούν να χρησιμοποιήσουν για να εξετάσουν πιο ολιστικά ζητήματα βλάβης, κινδύνου και καινοτομίας από ό,τι όταν χρησιμοποιούν μεμονωμένα τους βασικούς δείκτες εκθετικού κέρδους ή ανάπτυξης. Αυτό που μετράμε έχει σημασία και για να ωφελήσει η ΤΝ όλους τους ανθρώπους και την κοινωνία, πρέπει να λάβει υπόψη την ανθρώπινη ευημερία και την περιβαλλοντική βιωσιμότητα.

Το πρότυπο αυτό παρέχει ένα επίσημο πλαίσιο και μια μεθοδολογία για “Υπεύθυνη Τεχνητή Νοημοσύνη”. Πιο συγκεκριμένα παρέχει στους μηχανικούς και στους τεχνολόγους (με έμφαση

στην τεχνητή νοημοσύνη) μια μεθοδολογία για καινοτομία κατά την αντιμετώπιση των θεμάτων “άνθρωποι και πλανήτης”, αποτελώντας μια γέφυρα με το συγκεκριμένο στοιχείο που το συναντούμε στο πλαίσιο ταξινόμησης συστημάτων ΤΝ του ΟΑΣΑ που παρουσιάζεται παρακάτω. Είναι μια μεθοδολογία καινοτομίας, όπου οι χρήστες καλούνται να προσδιορίσουν τις υπάρχουσες μετρήσεις (π.χ. έναν συγκεκριμένο ΣΒΑ) και να φανταστούν/σχεδιάσουν σενάρια γύρω από το προϊόν ΤΝ τους. Ρωτώντας για παράδειγμα: “Πώς σχετίζεται ο ΣΒΑ 4 με το προϊόν σας; Πώς μπορείτε να βεβαιωθείτε ότι βοηθάτε (αντί απλώς να μην βλάπτετε) τους ανθρώπους και τον πλανήτη κατά την ανάπτυξη της τεχνητής νοημοσύνης σας;” επιτρέπεται στους προγραμματιστές, τις εταιρείες και τους υπεύθυνους χάραξης πολιτικής να κινηθούν πέρα από τα υφιστάμενα οικονομικά κριτήρια και τα στερεοτυπικά πλαίσια της αποτελεσματικότητας και της ταχύτητας.

Η ευημερία μπορεί να μετρηθεί αξιόπιστα, σύμφωνα με την IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. Ο συνήθης τρόπος μέτρησης της επιτυχίας των αυτόνομων και ευφών συστημάτων είναι κυρίως η οικονομική αξία, συχνά σε βάρος της ανθρώπινης ευημερίας. Οι δείκτες ευημερίας που προτείνει το πρότυπο 7010-2020 IEEE³⁵ καταγράφουν τόσο θετικές όσο και αρνητικές επιπτώσεις στην ανθρώπινη ευημερία, αλλά και ουδέτερα αποτελέσματα, επιτρέποντας στους δημιουργούς αυτόνομων και ευφών συστημάτων να διαχειρίζονται τις επιδράσεις αυτών των συστημάτων. Οι δείκτες διακρίνονται σε υποκειμενικούς και αντικειμενικούς και η συνδυασμένη χρήση τους επιτρέπει μια ολοκληρωμένη κατανόηση της ευημερίας σε κοινωνικό και ατομικό επίπεδο.

Οι υποκειμενικοί δείκτες καταγράφουν την ικανοποίηση από τη ζωή, το συναίσθημα, την ψυχολογική ευημερία, την ικανοποίηση από τις αποδοχές, την αίσθηση ασφάλειας, την αίσθηση μοναξιάς κλπ. Η έρευνα υποστηρίζει τη χρήση υποκειμενικών δεικτών, όπως αυτοί που βασίζονται σε ερωτηματολόγια, έρευνες, δημοσκοπήσεις ή άλλα μέσα συλλογής δεδομένων από την υποκειμενική εμπειρία. Οι υποκειμενικοί δείκτες που πληρούν αυστηρά επιστημονικά πρότυπα θεωρούνται έγκυρα και αξιόπιστα μέτρα στην επιστημονική κοινότητα. Οι υποκειμενικοί δείκτες χρησιμοποιούνται από εθνικούς και διεθνείς οργανισμούς και κυβερνήσεις για την καλύτερη κατανόηση της ευημερίας. Παραδείγματα ερευνών που περιλαμβάνουν υποκειμενικούς δείκτες ευημερίας είναι η Ευρωπαϊκή Κοινωνική Έρευνα, οι στατιστικές της Eurostat για το εισόδημα, την κοινωνική ένταξη και τις συνθήκες διαβίωσης (Eurostat Income and Living standards) και το πρόγραμμα ευημερίας της Εθνικής Στατιστικής Υπηρεσίας του Ηνωμένου Βασιλείου.

Οι αντικειμενικοί δείκτες καταγράφουν παρατηρήσιμα δεδομένα όπως το εισόδημα, την παραγωγικότητα, το καθεστώς απασχόλησης, την εκπαίδευση, το προσδόκιμο ζωής, τις ώρες εργασίας ανά εβδομάδα, την υγεία, την εγκληματικότητα, τη στέγαση κλπ. Οι δείκτες αυτοί έχουν χρησιμοποιηθεί για την κατανόηση των συνθηκών που επιτρέπουν την ευημερία του πληθυσμού μιας χώρας. Αυτοί οι δείκτες χρησιμοποιούνται τόσο από κυβερνήσεις όσο και από εταιρείες για την αξιολόγηση των συνθηκών που προάγουν την ευημερία.

4.4.2 Τυποποίηση και πρότυπα για την ΤΝ στην ΕΕ

Ερχόμενοι στα της Ευρωπαϊκής Ένωσης, σύμφωνα με το καθιερωμένο σύστημα ρύθμισης της ασφάλειας των προϊόντων στην ΕΕ³⁶, τα εναρμονισμένα πρότυπα αποτελούν ένα από τα κύρια μέσα για την επίτευξη συμμόρφωσης με τις νομοθετικές απαιτήσεις. Τα εναρμονισμένα πρότυπα για την ΤΝ θα αναπτυχθούν από τους ευρωπαϊκούς οργανισμούς τυποποίησης μετά από αίτημα τυποποίησης της Ευρωπαϊκής Επιτροπής που δημοσιεύθηκε τον Μάιο του 2023³⁷. Στο εν λόγω αίτημα επισημαίνεται ότι η προσέγγιση της ΕΕ για την τεχνητή νοημοσύνη δίνει έμφαση στην προώθηση μιας προσέγγισης της τεχνολογίας αυτής με βάση τα ανθρώπινα δικαιώματα, σε συνεργασία

³⁵IEEE Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being,” in IEEE Std 7010-2020, vol., no., pp.1-96, 1 May 2020, doi: 10.1109/IEEESTD.2020.9084219, σελ. 15

³⁶Κανονισμός 1025/2012 του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου, της 25ης Οκτωβρίου 2012, σχετικά με την ευρωπαϊκή τυποποίηση, <https://eur-lex.europa.eu/eli/reg/2012/1025/2023-07-09>

³⁷[https://ec.europa.eu/transparency/documents-register/detail?ref=C\(2023\)3215&lang=en](https://ec.europa.eu/transparency/documents-register/detail?ref=C(2023)3215&lang=en)

με άλλες χώρες σε παγκόσμιο επίπεδο και αναγνωρίζει τη σημασία των προσπαθειών θέσπισης προτύπων και κανονιστικών πλαισίων που καθορίζονται σε διεθνή φόρουμ, όπως για παράδειγμα οι συστάσεις του ΟΟΣΑ για την τεχνητή νοημοσύνη, που εγκρίθηκαν τον Μάιο του 2019, και οι συστάσεις για τη δεοντολογία της τεχνητής νοημοσύνης από την UNESCO, που εγκρίθηκαν τον Νοέμβριο του 2021. Η CEN και η CENELEC δημιούργησαν τη νέα JTC21 (Joint Technical Committee) για την Τεχνητή Νοημοσύνη.

Παρακάτω εξάιρεται η σημασία των προτύπων ως σημαντικών μέσων για την υποστήριξη της εφαρμογής των πολιτικών και της νομοθεσίας της Ένωσης και για τη διασφάλιση υψηλού επιπέδου προστασίας της ασφάλειας και των θεμελιωδών δικαιωμάτων για όλα τα άτομα στην Ένωση. Τονίζεται επίσης ότι τα πρότυπα μπορούν να υποστηρίξουν τη δημιουργία ισότιμων όρων ανταγωνισμού για τον σχεδιασμό και την ανάπτυξη συστημάτων TN, ιδίως για τις μικρομεσαίες επιχειρήσεις που αναπτύσσουν λύσεις TN. Διευκρινίζεται ότι ορισμένα πρότυπα που έχουν αναπτυχθεί σε διεθνές επίπεδο από τον Διεθνή Οργανισμό Τυποποίησης (ISO) και τη Διεθνή Ηλεκτροτεχνική Επιτροπή (IEC) σε σχετικούς τομείς μπορούν να υιοθετηθούν ως ευρωπαϊκά πρότυπα από τη CEN και τη CENELEC, εφόσον είναι πλήρως ευθυγραμμισμένα με τις ευρωπαϊκές αξίες και ιδιαιτερότητες και επιπλέον τονίζεται ότι θα πρέπει να ληφθεί υπόψιν το σχετικό έργο που έχει ήδη επιτελέσει στον τομέα αυτό το Ευρωπαϊκό Ινστιτούτο Τηλεπικοινωνιακών Προτύπων (European Telecommunications Standards Institute, ETSI).

Δεδομένου ότι η διεθνής τυποποίηση μπορεί να συμβάλει στην εδραίωση ενός κοινού οράματος για την αξιόπιστη τεχνητή νοημοσύνη σε ολόκληρο τον κόσμο και να διευκολύνει το εμπόριο και την άρση πιθανών τεχνικών εμποδίων σε σχέση με προϊόντα και υπηρεσίες που λειτουργούν με τεχνητή νοημοσύνη και προκειμένου να προωθηθεί η τεχνική εναρμόνιση στον τομέα της τεχνητής νοημοσύνης και να προετοιμαστεί το απαραίτητο τεχνικό περιβάλλον για την εφαρμογή της ευρωπαϊκής νομοθεσίας, κρίνεται απαραίτητο να ξεκινήσουν οι εργασίες για τα ευρωπαϊκά πρότυπα και την ευρωπαϊκή τυποποίηση. Τα εν λόγω πρότυπα αναμένεται να περιλαμβάνουν τις προδιαγραφές για το σχεδιασμό και την ανάπτυξη συστημάτων τεχνητής νοημοσύνης που χαρακτηρίζονται ως υψηλού κινδύνου, για τα συστήματα διαχείρισης ποιότητας των παρόχων τεχνητής νοημοσύνης και για την αξιολόγηση της συμμόρφωσης των συστημάτων τεχνητής νοημοσύνης.

Παρατίθεται ένας προκαταρκτικός κατάλογος των προτύπων TN από το σχέδιο εργασίας της CEN-CENELEC JTC21³⁸ τον Ιανουάριο του 2023, το οποίο συνδυάζει την υιοθέτηση διεθνών προτύπων από το ISO/IEC με την ανάπτυξη νέων στοιχείων εργασίας από την JTC 21, προκειμένου να καλυφθούν οι ευρωπαϊκές ιδιαιτερότητες.

- ISO/IEC 22989 Artificial Intelligence concepts and terminology
- ISO/IEC 23053 Framework for Artificial Intelligence (AI) system using Machine Learning
- ISO/IEC 42001 AI management system
- ISO/IEC 23894 AI Risk Management
- ISO/IEC 5259-part 1, Data quality for analytics and machine learning (ML) - Overview, terminology and examples
- ISO/IEC 5259-part 2, Data quality for analytics and machine learning (ML) Data quality measures
- ISO/IEC 5259-part 3, Data quality for analytics and machine learning (ML) Data quality management requirements and guidelines
- ISO/IEC 5259-part 4, Data quality for analytics and machine learning (ML) Data quality process framework

³⁸<https://publications.jrc.ec.europa.eu/repository/handle/JRC132833>

- ISO/IEC 27001:2013 Information security management systems
- ISO/IEC 42006 Requirements on bodies performing audit and certification of AI management systems
- CEN-CENELEC Risk AI Risk catalogue and management
- CEN-CENELEC Trustworthiness AI trustworthiness characterisation

Δεν πρέπει πάντως να μας διαφεύγει αν και τα θεμελιώδη δικαιώματα έχουν ήδη αντιμετωπιστεί και προστατεύονται από την ευρωπαϊκή νομοθεσία, η Πράξη για την TN είναι η πρώτη προσπάθεια ενσωμάτωσης των θεμελιωδών δικαιωμάτων σε μια προσέγγιση ασφάλειας προϊόντων, χρησιμοποιώντας εναρμονισμένα πρότυπα και σήμανση CE. Η Ευρωπαϊκή Επιτροπή³⁹ έχει αναγνωρίσει ότι “τα πρότυπα δεν ασχολούνται πλέον μόνο με τεχνικά στοιχεία, αλλά επιπλέον ενσωματώνουν βασικές δημοκρατικές αξίες και συμφέροντα της ΕΕ, καθώς και πράσινες και κοινωνικές αρχές. Αυτό είναι ιδιαίτερα σημαντικό όσον αφορά τα εναρμονισμένα πρότυπα, που έχουν εγκριθεί βάσει της ενωσιακής νομοθεσίας εναρμόνισης, είναι δημοσιευμένα στην Επίσημη Εφημερίδα της Ευρωπαϊκής Ένωσης και τα οποία το Δικαστήριο θεώρησε ως μέρος του δικαίου της ΕΕ για τους σκοπούς της ερμηνείας των εν λόγω προτύπων”.

Ενώ τα πρότυπα προορίζονταν αρχικά για την ασφάλεια των προϊόντων αυτή η έμφαση στην ασφάλεια έχει σταδιακά μετατοπιστεί σε άλλα κριτήρια. Η νομική σημασία των τεχνικών προτύπων στην ΕΕ έχει αυξηθεί, επειδή οι κανονισμοί δεν μπορούν να γίνουν κατανοητοί χωρίς τα σχετικά πρότυπα, γεγονός που τα καθιστά de facto δεσμευτικά, καθώς αν και υπάρχουν άλλοι τρόποι για να αποδειχθεί η συμμόρφωση, στην πράξη είναι δύσκολο αν όχι αδύνατο για τους κατασκευαστές να επιλέξουν διαφορετική οδό εάν θέλουν να παραμείνουν ανταγωνιστικοί. Ένα άλλο πλεονέκτημα είναι ότι το τεκμήριο συμμόρφωσης αντιστρέφει το βάρος της απόδειξης, αφού η εταιρεία δεν χρειάζεται να αποδείξει ότι συμμορφώνεται με τη νομοθεσία, καθώς αυτό τεκμαίρεται αυτόματα. Εάν ένας κατασκευαστής επιλέξει να μην συμμορφωθεί με τα εναρμονισμένα πρότυπα, φέρει το βάρος του να αποδείξει ότι το προϊόν του συμμορφώνεται με τη νομοθεσία, η οποία αντιπροσωπεύει τεράστιο εμπορικό κίνδυνο που κανένας κατασκευαστής δεν θα ανέλαμβανε (Γνωμοδότηση του Γενικού Εισαγγελέα Medina, 2023, §42). Όπως σημείωσε ο Γενικός Εισαγγελέας στην προσφυγή της υπόθεσης Public.Resource.Org (§47), η όλη αρχιτεκτονική του συστήματος τυποποίησης της ΕΕ προϋποθέτει ότι όλοι οι φορείς χρησιμοποιούν εναρμονισμένα πρότυπα (hEN)⁴⁰.

4.5 Βασικοί δείκτες επίδοσης (KPIs) για την TN

4.5.1 Επισκόπηση και τεκμηρίωση των βασικών δεικτών επίδοσης μέσα από ευρωπαϊκά κείμενα

Οι δείκτες επίδοσης ή βασικοί δείκτες επίδοσης (Key Performance Indicators KPIs) είναι ένας τύπος μέτρησης επιδόσεων. Οι βασικοί δείκτες επίδοσης αξιολογούν την επιτυχία ενός οργανισμού ως προς τους στόχους που έχει θέσει. Θα παρακολουθήσουμε συνοπτικά μέσα από ευρωπαϊκά κείμενα την πορεία προς την καθιέρωση των βασικών δεικτών επίδοσης για την TN. Η νέα

³⁹COM/2022/31 final, Η στρατηγική της ΕΕ για την τυποποίηση: Δημιουργία παγκόσμιων προτύπων για μια ανθεκτική, πράσινη και ψηφιακή ενιαία αγορά της ΕΕ, Βρυξέλλες, 2 Φεβρουαρίου 2022, <https://eur-lex.europa.eu/legal-content/EL/TXT/HTML/?uri=CELEX:52022DC0031>

⁴⁰Case C-588/21 P. (2023). Opinion of Advocate General Medina delivered on 22 June 2023(1) Public.Resource.Org, Inc., Right to Know CLG v European Commission. The Court of Justice of the European Union.<https://curia.europa.eu/juris/document/document.jsf?text=&docid=274881&pageIndex=0&doclang=EN&mode=req&dir=&occ=first&part=1&cid=2825624>

Ευρωπαϊκή Επιτροπή το 2019 περιέγραψε πώς ήθελε η ΕΕ να αξιοποιήσει τις ευκαιρίες που παρουσιάζει η ψηφιακή εποχή και έθεσε ως μία από τις προτεραιότητες της για την περίοδο 2019-2024 να δημιουργήσει “Μια Ευρώπη κατάλληλη για την ψηφιακή εποχή”⁴¹. Ένας τέτοιος ψηφιακός μετασχηματισμός βασίζεται στην προϋπόθεση⁴² ότι οι ψηφιακές τεχνολογίες και λύσεις θα πρέπει να ανοίγουν νέες ευκαιρίες για τις επιχειρήσεις, να ενισχύουν την ανάπτυξη αξιόπιστης τεχνολογίας και να προωθούν μια ζωντανή και βιώσιμη οικονομία.

Έχοντας αυτό υπόψη, τον Φεβρουάριο του 2020 η Ευρωπαϊκή Επιτροπή ενέκρινε την “Ευρωπαϊκή Στρατηγική για τα Δεδομένα”⁴³ που επιδιώκει να προωθήσει την ΕΕ ως ηγετικό πρότυπο και τη “Λευκή Βίβλο για την Τεχνητή Νοημοσύνη — Μια ευρωπαϊκή προσέγγιση για την αριστεία και την εμπιστοσύνη”⁴⁴, η οποία υποστηρίζει μια ρυθμιστική και οικονομική προσέγγιση. Το 2021 διατυπώθηκε η “Ψηφιακή πυξίδα για την Ψηφιακή Δεκαετία της ΕΕ”⁴⁵ και έθεσε τους ψηφιακούς στόχους της ΕΕ για το 2030 που εξελίσσονται γύρω από τέσσερα βασικά σημεία: δεξιότητες, ψηφιακός μετασχηματισμός των επιχειρήσεων, ασφαλείς και βιώσιμες ψηφιακές υποδομές και ψηφιοποίηση των δημοσίων υπηρεσιών.

Επισημαίνεται σε αυτό το σημείο εκείνο που προαναφέρθηκε και σχετικά με τα πρότυπα. Δηλαδή ότι σε συνέχεια των επτά απαιτήσεων, που προέρχονται από το έγγραφο “Κατευθυντήριες γραμμές δεοντολογίας για αξιόπιστη τεχνητή νοημοσύνη”⁴⁶ που ήδη αναλύθηκαν, και τις οποίες θα πρέπει να τηρούν τα συστήματα ΤΝ για να πραγματοποιηθεί αξιόπιστη ΤΝ, για την εφαρμογή τους μπορούν να χρησιμοποιούνται τόσο τεχνικές όσο και μη τεχνικές μέθοδοι. Ειδικότερα στις μη τεχνικές μεθόδους συγκαταλέγονται και οι κώδικες δεοντολογίας (παράγραφος 105) και αναφέρεται ότι οι οργανισμοί και τα ενδιαφερόμενα μέρη μπορούν να προσυπογράψουν τις κατευθυντήριες γραμμές και να προσαρμόσουν ανάλογα και τους **βασικούς δείκτες επίδοσης**, ώστε να συμβάλουν στις προσπάθειες για αξιόπιστη ΤΝ μετουσιώνοντάς τες σε μετρήσιμα μεγέθη επιθυμητών αξιών, όπως τα θεμελιώδη δικαιώματα, η διαφάνεια και η αποφυγή βλάβης.

4.5.2 Μεθοδολογία έρευνας για την εξεύρεση βασικών δεικτών επίδοσης-KPIs

Με αφετηρία τα παραπάνω, στην παρούσα μελέτη εξετάστηκαν ως πηγή αναζήτησης βασικών δεικτών επίδοσης (KPIs) για την ΤΝ:

- (α) Ο συνδεδεμένος με την “Ψηφιακή πυξίδα για την Ψηφιακή Δεκαετία της ΕΕ”, **Δείκτης Ψηφιακής Οικονομίας και Κοινωνίας (DESI)**, που παρουσιάζεται διεξοδικά παρακάτω.
- (β) Η **πρώτη ποσοτική έρευνα σχετικά με την υιοθέτηση των τεχνολογιών ΤΝ από τις ευρωπαϊκές επιχειρήσεις** και η οποία διεξήχθη από την Ευρωπαϊκή Επιτροπή την περίοδο Ιανουαρίου-Μαρτίου 2020⁴⁷.
- (γ) Ακολούθως εξετάστηκαν και παρουσιάζονται: α) **οι δείκτες επίδοσης της χρήσης της ΤΝ** που προκύπτουν από το κείμενο της προτεινόμενης “Πράξης για την ΤΝ” αλλά και β) **οι αναφερόμενοι δείκτες για την παρακολούθηση της υλοποίησης της πρότασης “Πράξη για**

⁴¹https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age_el

⁴²https://ec.europa.eu/eurostat/statisticsexplained/index.php?title=Use_of_artificial_intelligence_in_enterprises

⁴³<https://eur-lex.europa.eu/legal-content/EL/ALL/?uri=CELEX:52020DC0066>

⁴⁴<https://eur-lex.europa.eu/legal-content/EL/ALL/?uri=CELEX:52020DC0065>

⁴⁵https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/europes-digital-decade-digital-targets-2030_el

⁴⁶Κατευθυντήριες γραμμές δεοντολογίας για αξιόπιστη Τεχνητή Νοημοσύνη, High Level Expert Group, 2019, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60424

⁴⁷European Commission, Ipsos Survey, European enterprise survey on the use of technologies based on artificial intelligence, 2020, <https://digital-strategy.ec.europa.eu/en/library/european-enterprise-survey-use-technologies-based-artificial-intelligence>

την ΤΝ” που βρίσκονται στο τέλος του κειμένου, στο “Νομοθετικό Δημοσιονομικό Δελτίο”, σημείο 1.4.4. κατ’ επιταγή του εγγράφου εργασίας⁴⁸ που συνέταξε η υπηρεσία DG CNECT (Γενική Διεύθυνση Επικοινωνιακών Δικτύων, Περιεχομένου και Τεχνολογιών) όπου επεσήμανε τη σημασία της παρακολούθησης και αξιολόγησης των πραγματικών επιπτώσεων του προτεινόμενου κανονιστικού πλαισίου.

- (δ) Στη συνέχεια εξετάζονται οι δείκτες του **AI Watch Index** που παρέχει ποσοτικούς δείκτες επίδοσης για τη χάραξη πολιτικής της ΤΝ και αφορά τη συγκριτική αξιολόγηση μεταξύ των χωρών που κάνει η Ευρωπαϊκή Επιτροπή.
- (ε) Το έγγραφο εργασίας (Staff working document) που συνέταξε η υπηρεσία της Γενικής Διεύθυνσης Επικοινωνιακών Δικτύων, Περιεχομένου και Τεχνολογιών (DG CNECT), το οποίο συνόδευε την πρόταση κανονισμού για την “Πράξη για την ΤΝ”, επισημαίνει τη σημασία της παρακολούθησης και αξιολόγησης των πραγματικών επιπτώσεων του προτεινόμενου κανονιστικού πλαισίου και μάλιστα παρέχει έναν **ενδεικτικό κατάλογο δεικτών**⁴⁹
- (στ) Πέραν των Ευρωπαϊκών πηγών, παρατίθεται ο **AI Index του Stanford Institute for Human-Centered Artificial Intelligence (HAI)** που παρακολουθεί και οπτικοποιεί δεδομένα που σχετίζονται με την τεχνητή νοημοσύνη. Σκοπό έχει να παρέχει αμερόληπτα, αυστηρά ελεγμένα και παγκόσμια προέλευσης δεδομένα σε υπεύθυνους χάραξης πολιτικής, ερευνητές, στελέχη, δημοσιογράφους και στο ευρύ κοινό για να αναπτύξουν καλύτερη εποπτεία σχετικά με το σύνθετο πεδίο της τεχνητής νοημοσύνης.
- (ζ) Στην εθνική μας νομοθεσία, ο **Νόμος 4961/2022** θεσπίζει το νομοθετικό πλαίσιο για την ανάπτυξη της τεχνητής νοημοσύνης και εισάγει την έννοια των βασικών δεικτών επίδοσης (Key Performance Indicators KPI) που θα καταρτίζει και θα επικαιροποιεί το υπό σύσταση Παρατηρητήριο του άρθρου 14, τους οποίους θα εξετάσουμε.
- (η) Ο ΟΟΣΑ συμβάλλει σημαντικά στις εργασίες για την ΤΝ, συνεργάζεται με την Ευρωπαϊκή Επιτροπή για την παγκόσμια παρακολούθηση και ανάλυση των εξελίξεων στον τομέα της τεχνητής νοημοσύνης και παίζει σημαντικό ρόλο στη μέτρηση και αξιολόγηση των επιπτώσεων από τη χρήση της, προτείνοντας μετρικές και εργαλεία αξιολόγησης.

4.5.3 Ψηφιακή πυξίδα για την ψηφιακή δεκαετία 2020-30 και δείκτες DESI

Η Επιτροπή παρουσίασε αρχικά την Ψηφιακή Πυξίδα 2030, η οποία περιλαμβάνει στόχους στους τομείς των ψηφιακών δεξιοτήτων, των υποδομών, των επιχειρήσεων και των δημόσιων υπηρεσιών, τον Μάρτιο του 2021. Στη διάρκεια της ομιλίας για την κατάσταση της Ένωσης, τον Σεπτέμβριο του 2021, η πρόεδρος φον ντερ Λάιεν παρουσίασε την πρόταση για την πορεία προς την ψηφιακή δεκαετία με σκοπό τον ψηφιακό μετασχηματισμό της Ευρώπης έως το 2030. Τον Ιούλιο 2023 το Ευρωπαϊκό Κοινοβούλιο και το Συμβούλιο της ΕΕ ήρθαν σε συμφωνία σχετικά με το πρόγραμμα πολιτικής 2030: **Πορεία προς την ψηφιακή δεκαετία**⁵⁰.

Το πρόγραμμα θεσπίζει⁵¹ μηχανισμό παρακολούθησης και συνεργασίας για την επίτευξη των κοινών γενικών και ειδικών στόχων στο πλαίσιο του ψηφιακού μετασχηματισμού της Ευρώπης που ορίζονται στην Ψηφιακή Πυξίδα 2030. Η Ψηφιακή Πυξίδα καλύπτει τον τομέα των δεξιοτήτων και των υποδομών, συμπεριλαμβανομένης της συνδεσιμότητας, την ψηφιοποίηση των επιχειρήσεων

⁴⁸ Commission, Staff working document impact assessment, Accompanying the Proposal for AI Act, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021SC0084>

⁴⁹ Όπως παραπάνω.

⁵⁰ https://ec.europa.eu/commission/presscorner/detail/el/IP_22_4503

⁵¹ ΑΠΟΦΑΣΗ (ΕΕ) 2022/2481 ΤΟΥ ΕΥΡΩΠΑΪΚΟΥ ΚΟΙΝΟΒΟΥΛΙΟΥ ΚΑΙ ΤΟΥ ΣΥΜΒΟΥΛΙΟΥ της 14ης Δεκεμβρίου 2022 για τη θέσπιση του προγράμματος πολιτικής 2030 “Ψηφιακή Δεκαετία”, <https://eur-lex.europa.eu/legal-content/EL/TXT/?uri=CELEX:32022D2481&qid=1686060601697>

και των επιγραμμικών δημόσιων υπηρεσιών, καθώς και τον σεβασμό των ψηφιακών δικαιωμάτων και των αρχών της ΕΕ κατά την επίτευξη των γενικών στόχων.

Η Επιτροπή πρότεινε επίσης διακήρυξη σχετικά με τα ψηφιακά δικαιώματα και τις ψηφιακές αρχές για τον ψηφιακό μετασχηματισμό. Η κοινή διακήρυξη του Ευρωπαϊκού Κοινοβουλίου, του Συμβουλίου και της Επιτροπής σχετικά με τα ψηφιακά δικαιώματα και τις ψηφιακές αρχές για την ψηφιακή δεκαετία, δημοσιεύθηκε τελικά στην Επίσημη Εφημερίδα της Ευρωπαϊκής Ένωσης τον Ιανουάριο 2023⁵² και διασφαλίζει ότι όλοι και όλες μπορούν να συμμετέχουν στο ψηφιακό περιβάλλον, που είναι ασφαλές και συμπεριληπτικό.

Για να παρακολουθεί αποτελεσματικά την πρόοδο που επιτυγχάνεται, η Επιτροπή, από κοινού με τα κράτη μέλη, αναπτύσσει προβλέψεις για κάθε στόχο σε επίπεδο ΕΕ. Με τη σειρά τους, τα κράτη μέλη προτείνουν εθνικούς στρατηγικούς χάρτες πορείας, στους οποίους περιγράφονται οι εθνικές προβλεπόμενες πορείες και δράσεις για την επίτευξη των γενικών και ειδικών στόχων, συμπεριλαμβανομένων των προγραμματισμένων κανονιστικών μέτρων και επενδύσεων.

Η πρόοδος θα μετράται με βασικούς δείκτες επίδοσης (key performance indicators KPIs) βάσει ενός ενισχυμένου Δείκτη Ψηφιακής Οικονομίας και Κοινωνίας (Digital Economy and Society Index - DESI) και θα αξιολογείται με βάση ετήσια έκθεση σχετικά με την “κατάσταση της ψηφιακής δεκαετίας”, η οποία θα υποβάλλεται στο Ευρωπαϊκό Κοινοβούλιο και στο Συμβούλιο.

Ο Δείκτης Ψηφιακής Οικονομίας και Κοινωνίας⁵³ (DESI) συνοψίζει τους δείκτες σχετικά με τις ψηφιακές επιδόσεις της Ευρώπης και παρακολουθεί την πρόοδο των χωρών της ΕΕ⁵⁴. Από το 2014, η Ευρωπαϊκή Επιτροπή παρακολουθεί την πρόοδο των κρατών μελών στον ψηφιακό τομέα και δημοσιεύει ετήσιες εκθέσεις για τον Δείκτη Ψηφιακής Οικονομίας και Κοινωνίας (DESI). Ο δείκτης DESI κατατάσσει τα κράτη μέλη ανάλογα με το επίπεδο ψηφιοποίησής τους και αναλύει τη σχετική πρόοδό τους κατά την τελευταία πενταετία, λαμβάνοντας υπόψη το σημείο εκκίνησής τους.

Η Επιτροπή προσάρμοσε τον DESI για να τον ευθυγραμμίσει με τα τέσσερα βασικά σημεία που ορίζονται στην πρόταση της Επιτροπής για την απόφαση “Πορεία προς το πρόγραμμα πολιτικής της ψηφιακής δεκαετίας”. Η πρόταση θέτει στόχους σε επίπεδο ΕΕ που πρέπει να επιτευχθούν έως το 2030 για την επίτευξη ενός ολοκληρωμένου και βιώσιμου ψηφιακού μετασχηματισμού σε όλους τους τομείς της οικονομίας. Οι βαθμολογίες και οι κατατάξεις DESI των προηγούμενων ετών υπολογίζονται εκ νέου για όλα τα κράτη μέλη, ώστε να αντικατοπτρίζουν τις αλλαγές.

4.5.3.1 Βασικοί δείκτες επίδοσης

Ο δείκτης ψηφιακής οικονομίας και κοινωνίας (DESI) είναι μια ετήσια έκθεση που δημοσιεύεται από την Ευρωπαϊκή Επιτροπή και παρακολουθεί την πρόοδο των κρατών μελών της ΕΕ στον τομέα της ψηφιακής ανάπτυξης. Οι ανά χώρα εκθέσεις του DESI παρέχουν αξιολόγηση των εθνικών ψηφιακών πολιτικών και επισκόπηση των ψηφιακών επενδύσεων και μεταρρυθμίσεων στα σχέδια ανάκαμψης και ανθεκτικότητας⁵⁵. Επιπλέον περιλαμβάνει τα προφίλ των χωρών, τα οποία βοηθούν τα κράτη-μέλη να εντοπίζουν τομείς για δράσεις προτεραιότητας, καθώς και θεματικά κεφάλαια που παρέχουν ανάλυση σε επίπεδο ΕΕ, στους εξής τέσσερις κύριους τομείς πολιτικής:

1. ανθρώπινο κεφάλαιο
2. συνδεσιμότητα
3. ενσωμάτωση ψηφιακών τεχνολογιών
4. ψηφιακές δημόσιες υπηρεσίες.

⁵²<https://digital-strategy.ec.europa.eu/en/policies/digital-principles>

⁵³<https://digital-strategy.ec.europa.eu/el/policies/desi>

⁵⁴<https://digital-strategy.ec.europa.eu/en/policies/countries-digitisation-performance>

⁵⁵https://ec.europa.eu/commission/presscorner/detail/el/qanda_22_4561

4.5.3.2 Ποιες είναι οι πηγές των δεδομένων και πώς διαμορφώνεται ο δείκτης DESI για την TN;

Η πλειονότητα των δεικτών DESI προέρχονται από τη Eurostat, τη στατιστική υπηρεσία της Ευρωπαϊκής Ένωσης. Ορισμένοι δείκτες ευρυζωνικότητας συλλέγονται από τα κράτη μέλη μέσω της επιτροπής επικοινωνιών. Άλλοι δείκτες, όπως ορισμένοι δείκτες ηλεκτρονικής διακυβέρνησης και ευρυζωνικότητας, βασίζονται σε δεδομένα προερχόμενα από μελέτες που εκπονήθηκαν για λογαριασμό της Επιτροπής. Ο πλήρης κατάλογος των δεικτών, οι ακριβείς ορισμοί και οι πηγές διατίθενται στην ιστοσελίδα <https://digital-strategy.ec.europa.eu/el/policies/desi>.

Ο δείκτης σχετικά με την TN υπάγεται στον τομέα πολιτικής “Ενσωμάτωση ψηφιακών τεχνολογιών” και περιγράφεται στους Μεμονωμένους Δείκτες (DESI Individual Indicators)⁵⁶. Στο Μεθοδολογικό Σημείωμα⁵⁷ του DESI 2022 επισημαίνεται ότι έχουν γίνει αρκετές βελτιώσεις στους δείκτες DESI για τις εκθέσεις DESI 2022. Αναφέρει επίσης (σελ. 3) ότι ειδικά στο πλαίσιο της Ενσωμάτωσης της ψηφιακής τεχνολογίας χρησιμοποιήθηκαν για πρώτη φορά επίσημες στατιστικές για την TN από τη Eurostat. Με βάση την πολιτική συμφωνία για την προτεινόμενη πορεία προς την ψηφιακή δεκαετία, ποσοστό 75%, τουλάχιστον, των εταιρειών, θα πρέπει να υιοθετήσει τεχνολογίες TN, υπολογιστικού νέφους και μαζικών δεδομένων έως το 2030⁵⁸.

Σε συνέχεια των παραπάνω και σε συνδυασμό α) με το έγγραφο DESI 2022 σχετικά με τα θεματικά κεφάλαια που παρέχουν ανάλυση σε επίπεδο ΕΕ⁵⁹ και β) με το μεθοδολογικό σημείωμα του DESI 2022⁶⁰ προκύπτει ότι η TN υπάγεται στη διάσταση “Ενσωμάτωση Ψηφιακής Τεχνολογίας”. Επιπλέον το προαναφερθέν έγγραφο DESI 2022 με τα θεματικά κεφάλαια, που παρέχουν ανάλυση σε επίπεδο ΕΕ, αποτυπώνει τις ακόλουθες διαστάσεις του δείκτη ενσωμάτωσης τεχνολογιών TN από τις επιχειρήσεις⁶¹ σύμφωνα με στοιχεία που αντλούνται από τη Eurostat⁶²:

1. Εταιρείες που χρησιμοποιούν τεχνολογίες TN ανά τύπο τεχνολογίας TN
2. Μέγεθος εταιρείας σκοπού για τη χρήση τεχνολογιών TN
3. Οικονομική δραστηριότητα της εταιρείας
4. Τρόπος απόκτησης τεχνολογιών TN.

⁵⁶https://digital-decade-desi.digital-strategy.ec.europa.eu/datasets/desi-2022/charts/desi-components?period=2022&indicator=desi_idt&breakdownGroup=desi_idt&unit=pc_desi_idt

⁵⁷<https://ec.europa.eu/newsroom/dae/redirection/document/88557>

⁵⁸https://ec.europa.eu/commission/presscorner/detail/el/qanda_22_4561

⁵⁹Digital Economy and Society Index (DESI) 2022, Thematic chapters, <https://ec.europa.eu/newsroom/dae/redirection/document/88764>

⁶⁰<https://ec.europa.eu/newsroom/dae/redirection/document/88557>

⁶¹https://ec.europa.eu/eurostat/statistics-explained/images/f/f9/Enterprises_using_AI_Graphs_Tables.xls

⁶²https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Use_of_artificial_intelligence_in_enterprises

ΔΕΙΚΤΗΣ DESI 3b5: TN		Επιχειρήσεις που χρησιμοποιούν οποιαδήποτε τεχνολογία TN: (% των επιχειρήσεων)	
1	Χρήση TN ανά μέγεθος επιχείρησης	1	Μικρές
		2	Μεσαίες
		3	Μεγάλες
2	Οικονομική δραστηριότητα της εταιρείας	1	Πληροφορίες και επικοινωνία
		2	Δραστηριότητες επιστημονικών και τεχνικών υπηρεσιών
		3	Ηλεκτρικό ρεύμα, φυσικό αέριο, ατμός, κλιματισμός και ύδρευση
		4	Διοικητικές και υποστηρικτικές δραστηριότητες
		5	Βιοτεχνία
		6	Κτηματομεσιτικές δραστηριότητες
		7	Διαμονή
		8	Λιανικό εμπόριο
		9	Χονδρικό και λιανικό εμπόριο-επισκευή μηχανοκίνητων οχημάτων και μοτοσικλετών
		10	Κατασκευές
		11	Μεταφορά και αποθήκευση
3	Τεχνολογίες TN	1	Τεχνολογίες TN που αυτοματοποιούν διάφορες ροές εργασίας ή βοηθούν στη λήψη αποφάσεων (π.χ. ρομποτική αυτοματοποίηση διαδικασιών λογισμικού με βάση την TN)
		2	Μηχανική μάθηση (π.χ. βαθιά μάθηση) για την ανάλυση δεδομένων
		3	Τεχνολογίες ανάλυσης του γραπτού λόγου (π.χ. εξόρυξη κειμένου)
		4	Τεχνολογίες TN που αναγνωρίζουν αντικείμενα ή πρόσωπα με βάση εικόνες (αναγνώριση εικόνας, επεξεργασία εικόνας)
		5	Τεχνολογίες TN που μετατρέπουν τον προφορικό λόγο σε μορφή αναγνώσιμη από μηχανήματα (αναγνώριση ομιλίας)
		6	Τεχνολογίες που επιτρέπουν τη φυσική μετακίνηση μηχανών μέσω αυτόνομων αποφάσεων που βασίζονται στην παρατήρηση του περιβάλλοντος (π.χ. αυτοκινούμενα οχήματα)
		7	Τεχνολογίες TN που παράγουν γραπτό ή προφορικό λόγο (παραγωγή φυσικής γλώσσας)
4	Είδος σκοπού για τη χρήση τεχνολογιών TN	1	Για την ασφάλεια ΤΠΕ
		2	Για διαδικασίες παραγωγής
		3	Για την οργάνωση των διαδικασιών διοίκησης επιχειρήσεων
		4	Για μάρκετινγκ ή πωλήσεις
		5	Για τη διαχείριση των επιχειρήσεων
		6	Για την εφοδιαστική αλυσίδα
		7	Για τη διαχείριση ανθρώπινου δυναμικού ή την πρόσληψη προσωπικού
5	Τρόπος απόκτησης τεχνολογιών TN	1	Οι τεχνολογίες TN ήταν εμπορικό λογισμικό ή συστήματα έτοιμα προς χρήση
		2	Οι τεχνολογίες TN αναπτύχθηκαν ή τροποποιήθηκαν από εξωτερικούς παρόχους
		3	Οι τεχνολογίες TN ήταν εμπορικό λογισμικό ή συστήματα που τροποποιήθηκαν από τους ίδιους τους υπαλλήλους
		4	Οι τεχνολογίες TN αναπτύχθηκαν από τους ίδιους τους υπαλλήλους
		5	Οι τεχνολογίες TN ήταν λογισμικό ή συστήματα ανοικτού κώδικα που τροποποιήθηκαν από τους ίδιους τους υπαλλήλους

Πίνακας 4.1. DESI σε συνδυασμό με EUROSTAT: Δείκτες για την TN

Το αναλυτικό άρθρο της ιστοσελίδας της EUROSTAT με τίτλο “Use of artificial intelligence in enterprises” περιέχει διαγράμματα που προκύπτουν από τους δυνατούς συνδυασμούς των προαναφερθεισών διαστάσεων, με βάση τα στοιχεία της Eurostat σχετικά με τη χρήση τεχνολογιών

TN από τις επιχειρήσεις⁶³, κάτι για το οποίο άρχισαν να συγκεντρώνονται στοιχεία από το 2020 και εξής. Οι διαστάσεις αυτές, μαζί με τις υποδιαστάσεις τους, αποδελτιώθηκαν για τους στόχους της παρούσας εργασίας στον παραπάνω πίνακα.

4.5.4 Δείκτες στην έρευνα “European enterprise survey on the use of technologies”

Στο σημείο αυτό θα επανέλθουμε στην πρώτη ποσοτική έρευνα που έγινε ποτέ σε επίπεδο ΕΕ σχετικά με την υιοθέτηση των τεχνολογιών TN από τις ευρωπαϊκές επιχειρήσεις και δημοσιεύθηκε τον Ιούλιο 2020⁶⁴, η οποία αναφέρθηκε παραπάνω. Στην τελική αναφορά⁶⁵ της εν λόγω έρευνας επισημαίνεται ότι στη διαδικασία ανάπτυξης ενός εργαλείου έρευνας για την αξιολόγηση της υιοθέτησης της τεχνητής νοημοσύνης στις επιχειρήσεις, συμπεριελήφθη ο προσδιορισμός βασικών δεικτών επίδοσης (KPIs) προς μέτρηση. Ο στόχος ήταν να επιλεγούν θεματικές που θα προσέφεραν πολύτιμες πληροφορίες για την τρέχουσα κατάσταση της υιοθέτησης της τεχνητής νοημοσύνης και θα συνέβαλαν στη διαμόρφωση των σχετικών πολιτικών.

Προτεραιότητα δόθηκε στους KPIs που παρουσίαζαν συνάφεια με τους στόχους της μελέτης και ήταν μετρήσιμοι. Αφού προσδιορίστηκαν οι πιο αντιπροσωπευτικές μετρικές, σχεδιάστηκαν ερωτήσεις για τη συλλογή πληροφοριών που συνδέονται με αυτούς τους KPIs. Η εστίαση του ερωτηματολογίου περιστράφηκε γύρω από αυτούς τους βασικούς δείκτες επίδοσης για την ολοκληρωμένη αξιολόγηση του τοπίου της υιοθέτησης τεχνητής νοημοσύνης στις επιχειρήσεις. Συγκεκριμένα ερωτήθηκαν συνολικά 9.640 επιχειρήσεις σε 30 ευρωπαϊκές χώρες, την περίοδο Ιανουαρίου-Μαρτίου 2020. Για τις ανάγκες της έρευνας μετρήθηκαν οι εξής πέντε KPIs (βασικοί δείκτες επίδοσης) που αφορούν τη χρήση τεχνολογιών TN από τις επιχειρήσεις:

- Ενημέρωση για την τεχνητή νοημοσύνη
- Υιοθέτηση της TN
- Προμήθεια TN
- Εξωτερικά εμπόδια στην υιοθέτηση της TN
- Εσωτερικά εμπόδια στην υιοθέτηση της TN.

4.5.5 Βασικοί Δείκτες Επίδοσης στην Πρόταση Κανονισμού “Πράξη για την TN”

Ο δείκτης επίδοσης ή βασικός δείκτης επίδοσης (Key Performance Indicators KPIs) είναι ένας τύπος μέτρησης επιδόσεων. Οι βασικοί δείκτες επίδοσης αξιολογούν την επιτυχία ενός οργανισμού ως προς τους στόχους που έχει θέσει. Στο πλαίσιο της πρότασης του Κανονισμού “Πράξη για την TN” δεν γίνεται κάποια εξαντλητική αναφορά σε δείκτες επίδοσης, ωστόσο από το προτεινόμενο κείμενο συνάγονται: α) δείκτες επίδοσης που αφορούν τη χρήση της TN και β) δείκτες επίδοσης της υλοποίησης της “Πράξης για την TN”.

4.5.5.1 Δείκτες επίδοσης της χρήσης της TN στο κείμενο της “Πράξης για την TN”

Η Πράξη για την TN αναφέρεται ακροθιγώς σε δείκτες επίδοσης. Στο αρχικό κείμενο της πρότασης κανονισμού, στο άρθρο 69 με τίτλο “Κώδικες Δεοντολογίας” αναφέρεται έμμεσα σε αυτούς συνδέοντας την εθελοντική συμμόρφωση στα συστήματα TN με τους βασικούς δείκτες

⁶³https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Use_of_artificial_intelligence_in_enterprises

⁶⁴European Commission, Ipsos Survey, European enterprise survey on the use of technologies based on artificial intelligence, 2020, <https://digital-strategy.ec.europa.eu/en/library/european-enterprise-survey-use-technologies-based-artificial-intelligence>

⁶⁵https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=68488

επιδόσεων για στόχους όπως η περιβαλλοντική βιωσιμότητα, η προσβασιμότητα, η συμμετοχικότητα. Στην τροπολογία 633, δηλαδή στην έκδοση του προτεινόμενου Κανονισμού που ψήφισε το Κοινοβούλιο, στο άρθρ. 69, παρ. 1 αναφέρεται “1. Η Επιτροπή, η Υπηρεσία ΤΝ και τα κράτη μέλη ενθαρρύνουν και διευκολύνουν την κατάρτιση κωδικών δεοντολογίας που αποσκοπούν, μεταξύ άλλων όταν καταρτίζονται για να καταδείξουν τον τρόπο με τον οποίο τα συστήματα ΤΝ τηρούν τις αρχές που ορίζονται στο άρθρο 4α και μπορούν επομένως να θεωρηθούν αξιόπιστα”. Στη συνέχεια η Τροπολογία 634, Άρθρο 69, παράγραφος 2, όπως ψηφίστηκε από το Κοινοβούλιο στις 14.6.2023⁶⁶, αναφέρει τα εξής χαρακτηριστικά για αξιόπιστη ΤΝ, τα οποία συμπεραίνουμε ότι θα μπορούσαν να μετουσιωθούν σε βασικούς δείκτες επίδοσης. “Οι κώδικες δεοντολογίας που αποσκοπούν στην προώθηση της εθελοντικής συμμόρφωσης με τις αρχές που διέπουν τα αξιόπιστα συστήματα ΤΝ, έχουν ιδίως τα εξής χαρακτηριστικά”:

- Γραμματισμός (αλφαριθμητισμός) του προσωπικού στον τομέα της ΤΝ
- Εκτίμηση επιπτώσεων σε ευάλωτα άτομα ή ομάδες και ενίσχυση της προσβασιμότητας
- Εκτίμηση επιπτώσεων στη διαφορετικότητα και υποστήριξη ισότητας των φύλων
- Πιθανότητα μεροληψίας
- Συμπεριληπτικός σχεδιασμός υπέρ ποικιλόμορφων ομάδων
- Επιπτώσεις στη δημοκρατία και τους πολιτικούς θεσμούς
- Περιβαλλοντική βιωσιμότητα

4.5.5.2 Αναφορά στους Στόχους Βιώσιμης Ανάπτυξης των Ηνωμένων Εθνών

Επίσης η “Τροπολογία 83, Πρόταση κανονισμού, Αιτιολογική σκέψη 46β (νέα)” εισάγει κάτι που δεν υπήρχε στο αρχικά υποβληθέν κείμενο της πρότασης της Πράξης για την ΤΝ και αφορά την ανάπτυξη μεθοδολογίας καθορισμού βασικών δεικτών επίδοσης και αναφοράς που θα παρέμπει στους Στόχους Βιώσιμης Ανάπτυξης (ΣΒΑ) των Ηνωμένων Εθνών⁶⁷. Επί λέξει αναφέρει τα εξής: “Για την επίτευξη των στόχων του παρόντος κανονισμού και τη συμβολή στην επίτευξη των περιβαλλοντικών στόχων της Ένωσης, με παράλληλη διασφάλιση της ομαλής λειτουργίας της εσωτερικής αγοράς, ενδέχεται να είναι αναγκαίο να θεσπιστούν συστάσεις και κατευθυντήριες γραμμές και, εν τέλει, στόχοι για τη βιωσιμότητα. Για τον σκοπό αυτό, η Επιτροπή δικαιούται να αναπτύξει μεθοδολογία που θα συμβάλει στον καθορισμό βασικών δεικτών επίδοσης και αναφοράς για τους Στόχους Βιώσιμης Ανάπτυξης (ΣΒΑ). Ο στόχος θα πρέπει να είναι, κατά πρώτο λόγο, να καταστεί δυνατή η δίκαιη σύγκριση μεταξύ των επιλογών εφαρμογής της ΤΝ, παρέχοντας κίνητρα για την προώθηση της χρήσης αποδοτικότερων τεχνολογιών ΤΝ για την αντιμετώπιση ζητημάτων που αφορούν την ενέργεια και τους πόρους. Για την επίτευξη αυτού του στόχου, ο παρών κανονισμός θα πρέπει να παρέχει τα μέσα για την καθιέρωση βασικής συλλογής δεδομένων που υποβάλλονται σχετικά με τις εκπομπές από την ανάπτυξη και την κατάρτιση και για την εγκατάσταση.”

4.5.5.3 Δείκτες επίδοσης της υλοποίησης της “Πράξης για την ΤΝ”

Στο πλαίσιο που περιγράφηκε παραπάνω και όπως σκιαγραφείται μέσα από τα προαναφερθέντα κείμενα, η Ευρωπαϊκή Επιτροπή παρουσίασε την πρότασή της για έναν νέο κανονισμό περί τεχνητής νοημοσύνης, την Πράξη για την Τεχνητή Νοημοσύνη τον Απρίλιο του 2021. Η Πράξη για την ΤΝ περιλαμβάνει έναν τεχνολογικά ουδέτερο ορισμό των συστημάτων τεχνητής νοημοσύνης

⁶⁶https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EL.html

⁶⁷<https://sdgs.un.org/goals>

και βασίζεται στην αντιμετώπιση διαφορετικών επιπέδων κινδύνων που μπορεί να προκληθούν και στη ρύθμιση της χρήσης της τεχνητής νοημοσύνης, αποτελώντας το πρώτο ολοκληρωμένο νομικό πλαίσιο στον κόσμο για την τεχνητή νοημοσύνη. Στόχος είναι η μεγιστοποίηση των οφελών των συστημάτων TN, η εξασφάλιση ανταγωνιστικού πλεονεκτήματος για τους κατασκευαστές των συστημάτων TN που ενσωματώνουν αξιόπιστη TN στα προϊόντα και στις υπηρεσίες τους και η πρόληψη και ελαχιστοποίηση των κινδύνων τους.

Η “Πράξη για την TN” συνοδεύεται από το έγγραφο εργασίας⁶⁸ που συντάξε η υπηρεσία DG CONNECT (Γενική Διεύθυνση Επικοινωνιακών Δικτύων, Περιεχομένου και Τεχνολογιών). Στο έγγραφο επισημαίνεται η σημασία της παρακολούθησης και αξιολόγησης των πραγματικών επιπτώσεων του προτεινόμενου κανονιστικού πλαισίου. Πράγματι η Ευρωπαϊκή Επιτροπή διενήργησε εκτίμηση των επιπτώσεων της Πράξης για την Τεχνητή Νοημοσύνη και στο “Νομοθετικό Δημοσιονομικό Δελτίο” (στο σημείο 1.4.4. του κειμένου της προτεινόμενης Πράξης για την TN αναφέρεται ότι “προσδιορίζονται δείκτες για την παρακολούθηση της υλοποίησης της πρότασης”.

Από το εν λόγω νομοθετικό δελτίο προκύπτουν δύο δείκτες, ο ένας εκ των οποίων ευθυγραμμίζεται με τον στόχο της ανθρωποκεντρικής TN και περιγράφεται ως “Αριθμός σοβαρών περιστατικών ή επιδόσεων TN που συνιστούν σοβαρό περιστατικό ή παραβίαση των υποχρεώσεων στον τομέα των θεμελιωδών δικαιωμάτων (ανά εξάμηνο) ανά πεδίο εφαρμογών” και ο οποίος υπολογίζεται με τρεις παραμέτρους: α) σε απόλυτους αριθμούς, β) ως ποσοστό των εφαρμογών που αναπτύχθηκαν γ) ως ποσοστό των θιγομένων πολιτών.

Ο άλλος δείκτης που αναφέρεται είναι οικονομικού περιεχομένου και αναλύεται σε στοιχεία που αφορούν το ύψος των επενδύσεων και τη χρήση της TN από εταιρείες.

Δείκτης	Περιγραφή	Υπολογισμός
1	Αριθμός σοβαρών περιστατικών ή επιδόσεων TN που συνιστούν σοβαρό περιστατικό ή παραβίαση των υποχρεώσεων στον τομέα των θεμελιωδών δικαιωμάτων (ανά εξάμηνο) ανά πεδίο εφαρμογών	α) σε απόλυτους αριθμούς,
		β) ως ποσοστό των εφαρμογών που αναπτύχθηκαν
		γ) ως ποσοστό των ενδιαφερόμενων πολιτών
2	A) Σύνολο επενδύσεων στον τομέα της τεχνητής νοημοσύνης στην ΕΕ (ετησίως).	Με βάση επίσημες πηγές και θα συγκρίνονται με ιδιωτικές εκτιμήσεις
	B) Σύνολο επενδύσεων στον τομέα της τεχνητής νοημοσύνης ανά κράτος μέλος (ετησίως)	
	Γ) Ποσοστό εταιρειών που χρησιμοποιούν TN (ετησίως)	θα συλλέγονται μέσω τακτικών ερευνών των επιχειρήσεων
	Δ) Ποσοστό MME που χρησιμοποιούν TN (ετησίως)	

Πίνακας 4.2. Δείκτες για την παρακολούθηση της υλοποίησης της πρότασης από το «Νομοθετικό Δημοσιονομικό Δελτίο» της προτεινόμενης Πράξης για την TN

4.5.6 Δείκτες του AI WATCH

4.5.6.1 AI Watch, National Strategies on AI: A European Perspective, 2022

Όλες οι χώρες συμφωνούν στη σημασία της αξιολόγησης της προόδου της εφαρμογής των στρατηγικών. Παρά τη συνειδητοποίηση αυτή, λίγες χώρες έχουν θεσπίσει μέσα αξιολόγησης, όπως βασικούς δείκτες επιδόσεων ή ορόσημα για τα επόμενα χρόνια. Συγκεκριμένα⁶⁹, μόνο η

⁶⁸ COMMISSION STAFF WORKING DOCUMENT, IMPACT ASSESSMENT LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) SWD(2021) 84 final, Brussels, 21 April 2021, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021SC0084>

⁶⁹ Jorge Ricart, R., Van Roy, V., Rossetti, F. and Tangi, L., AI Watch. National strategies on Artificial Intelligence: A European perspective. 2022 edition, EUR 31083 EN, Publications Office of the European Union, Luxembourg, 2022, https://ai-watch.ec.europa.eu/publications/ai-watch-national-strategies-artificial-intelligence-european-perspective-2022-edition_en

Φινλανδία, η Γαλλία και η Ουγγαρία έχουν παράσχει έναν κατάλογο βασικών δεικτών επίδοσης με βάση τους οποίους μπορούν να αξιολογηθούν τα αποτελέσματα της πολιτικής. Είναι αξιοσημείωτο το γεγονός ότι η Φινλανδία έχει δημιουργήσει πέντε θεματικές ομάδες εργασίας για τον καθορισμό συγκεκριμένων μέτρων και βασικών δεικτών επιδόσεων, μία εκ των οποίων επικεντρώνεται ειδικά στις ευρωπαϊκές και διεθνείς υποθέσεις.

4.5.6.2 AI Watch Index 2021

Το AI Watch Index 2021⁷⁰ παρέχει ένα δομημένο σύνολο ποσοτικών δεικτών σχετικά με την απόδοση και τη θέση της ΕΕ σε διάφορες διαστάσεις της τεχνητής νοημοσύνης που σχετίζονται με τη χάραξη πολιτική. Το ευρετήριο οργανώνεται γύρω από πέντε διαστάσεις:

1. G: Παγκόσμια άποψη για το τοπίο της τεχνητής νοημοσύνης,
2. I: Βιομηχανία,
3. R: Έρευνα και ανάπτυξη (E&A),
4. T: Τεχνολογία,
5. S: Κοινωνικές πτυχές.

AI Watch Index Διάσταση	AI Watch Index Υποδιάσταση	Δείκτης
G - Παγκόσμια άποψη για το τοπίο της Τεχνητής Νοημοσύνης	Δραστηριότητα TN	G1: Οικονομικοί παίκτες TN G2: Ένταση παίκτη Τεχνητής Νοημοσύνης
	Ισχυροί τομείς της TN	G3: Τομείς εξειδίκευσης της TN: συγκριτικό πλεονέκτημα σε θεματικούς τομείς της TN G4: Θεματικές εστίες TN G5: Συγκριτικό πλεονέκτημα της ΕΕ στο εμπόριο βιομηχανικής ρομποτικής
	Επενδύσεις TN	G6: Επενδύσεις Τεχνητής Νοημοσύνης στην ΕΕ
I - Βιομηχανία	Βιομηχανία	I1: Προφίλ των επιχειρήσεων TN I2: Νεοσύστατες επιχειρήσεις ρομποτικής στην ΕΕ
R - Έρευνα και ανάπτυξη	Δραστηριότητα E&A	R1: Παίκτες TN στην E&A TN R2: Βαθμολογία δραστηριότητας E&A TN
	Δίκτυο συνεργασιών	R3: R4: Συνεργασίες μεταξύ ομότιμων εταιρών R5: Στρατηγική θέση στο δίκτυο συνεργασιών
T - Τεχνολογία	Απόδοση της TN	T1: Απόδοση της έρευνας για την τεχνητή νοημοσύνη
	Προτυποποίηση	T2: Δραστηριότητα προτυποποίησης
S – Κοινωνικές πτυχές	Ποικιλομορφία στην έρευνα	S1: Δείκτης ποικιλομορφίας των φύλων S2: Δείκτης γεωγραφικής ποικιλότητας S3: Δείκτης επιχειρηματικής ποικιλομορφίας S4: Δείκτης ποικιλομορφίας συνεδρίων
	Τριτοβάθμια εκπαίδευση	S5: TN σε πανεπιστημιακά προγράμματα στην ΕΕ S6: Πανεπιστημιακές έδρες με περιεχόμενο TN στην ΕΕ S7: Ένταση TN σε πανεπιστημιακές έδρες στην ΕΕ

Εικόνα 4.5. Σύνοψη των δεικτών του AI Watch Index 2021

⁷⁰https://ai-watch.ec.europa.eu/ai-watch-index-2021_en

4.5.7 Δείκτες του STAFF WORKING PAPER: AI Act Impact Assessment

Στο έγγραφο εργασίας που συντάξε η υπηρεσία της Γενικής Διεύθυνσης Επικοινωνιακών Δι-κτύων, Περιεχομένου και Τεχνολογιών (DG CNECT), το οποίο συνόδευε την πρόταση κανονισμού για την “Πράξη για την ΤΝ”, επισημαίνεται η σημασία της παρακολούθησης και αξιολόγησης των πραγματικών επιπτώσεων του προτεινόμενου κανονιστικού πλαισίου. Στην ενότητα 9 του κειμέ-νου αναφέρει τη δημιουργία μητρώων καταγραφής εφαρμογών ΤΝ αλλά και την ανατροφοδότηση της Ευρωπαϊκής Επιτροπής από τις αρμόδιες εθνικές αρχές σχετικά με επιπτώσεις κυρίως στα θε-μελιώδη δικαιώματα, ώστε να διασφαλιστεί αυξημένη εποπτεία.

Επίσης παρέχει έναν κατάλογο δεικτών που τονίζει ότι είναι ενδεικτικός και όχι αποκλειστικός (The following list of indicators is provisional and non-exhaustive)⁷¹ και ότι λαμβάνοντας υπόψη αυτούς τους δείκτες και συμπληρώνοντας με πρόσθετες ad-hoc πηγές καθώς και με ποιοτικά στοι-χεία προερχόμενα από το Joint Research Center (μέσω του AI Watch) και τον ΟΟΣΑ, η Επιτροπή θα δημοσιεύσει έκθεση αξιολόγησης και επανεξέτασης του πλαισίου ρύθμισης της ΤΝ (Πράξη για την ΤΝ) πέντε έτη μετά την ημερομηνία έναρξης εφαρμογής του. Τέτοιοι ενδεικτικοί δείκτες παρακολούθησης και αξιολόγησης επιπτώσεων από την εφαρμογή της Πράξης για την ΤΝ απο-τυπώνονται στον πίνακα 13 του εγγράφου (ο οποίος παρατίθεται παραπάνω σε μετάφραση), που περιέχει επίσης πεδία τα οποία περιγράφουν τον σκοπό και την πηγή προέλευσης των στοιχείων για τη συμπλήρωση του αντίστοιχου δείκτη.

ΣΚΟΠΟΣ	ΔΕΙΚΤΗΣ	ΠΗΓΗ
Τα συστήματα ΤΝ είναι ασφαλή και σέβονται τα θεμελιώδη δικαιώματα και τις αξίες της ΕΕ (αρνητικοί δείκτες)	Αριθμός σοβαρών περιστατικών ή επιδόσεων ΤΝ που συνιστούν σοβαρό περιστατικό ή παραβίαση υποχρεώσεων θεμελιωδών δικαιωμάτων (6μηνο) ανά πεδία εφαρμογών και υπολογίζεται α) σε απόλυτες τιμές, β) ως ποσοστό των εφαρμογών που έχουν αναπτυχθεί και γ) ως ποσοστό των ενδιαφερομένων πολιτών	Αρμόδιες εθνικές αρχές· Ευρωπαϊκό Συμβούλιο Προστασίας Δεδομένων
Διευκόλυνση των επενδύσεων και της καινοτομίας. (θετικοί δείκτες)	α) Συνολικές επενδύσεις ΤΝ στην ΕΕ (ετησίως) β) Συνολική επένδυση ΤΝ ανά κράτος μέλος (ετησίως) γ) Μερίδιο εταιρειών που χρησιμοποιούν ΤΝ (ετησίως) δ) Μερίδιο ΜΜΕ που χρησιμοποιούν ΤΝ (ετησίως) ε) Έργα μέσω sandboxes που διατίθενται στην αγορά (ετησίως) στ) Αριθμός ΜΜΕ που συμβουλευονται για την ΤΝ σε κόμβους ψηφιακής καινοτομίας και εγκαταστάσεις δοκιμών και πειραματισμών.	Υπηρεσίες της Επιτροπής και AI Watch· Αρμόδιες εθνικές αρχές
Βελτίωση των μηχανισμών διακυβέρνησης και επιβολής (αρνητικοί δείκτες)	Αριθμός ανακλήσεων ή αποσύρσεων συστημάτων ΤΝ από την αγορά (ανά 6μηνο) ανά πεδία εφαρμογών και υπολογίζεται α) σε απόλυτες τιμές, β) ως ποσοστό των εφαρμογών που έχουν αναπτυχθεί και γ) ως ποσοστό των ενδιαφερομένων πολιτών.	Αρμόδιες εθνικές αρχές
Διευκόλυνση της ενιαίας αγοράς	α) Επίπεδο εμπιστοσύνης στην ΤΝ (ετησίως) (θετικός δείκτης) β) Αριθμός εθνικών νομοθεσιών που δύνανται να κατακερματίσουν την ενιαία αγορά (ανά 6μηνο) (αρνητικός δείκτης)	Υπηρεσίες της Επιτροπής και AI Watch

Πίνακας 4.3. Δείκτες παρακολούθησης και αξιολόγησης (μετάφραση του “Table 13: Indicators for monitoring and evaluation” του εγγράφου Staff Working Paper, AI Act Impact Assessment)

4.5.8 AI Index, Stanford Institute for Human-Centered AI (HAI)

Ο AI Index είναι μια ανεξάρτητη πρωτοβουλία του Stanford Institute for Human-Centered Artificial Intelligence (HAI), με επικεφαλής την AI Index Steering Committee, μια διεπιστημο-νική ομάδα εμπειρογνομώνων από όλο τον ακαδημαϊκό χώρο και τη βιομηχανία. Η ετήσια έκθεση

⁷¹ Commission Staff Working Paper, AI Act Impact Assessment Accompanying the Proposal for AI Act, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021SC0084>

παρακολουθεί, συγκεντρώνει και οπτικοποιεί δεδομένα που σχετίζονται με την τεχνητή νοημοσύνη σε παγκόσμιο επίπεδο, διευκολύνοντας τη χάραξη πολιτικής.

Το AI Index συνεργάζεται με πολλούς διαφορετικούς οργανισμούς για την παρακολούθηση της προόδου στην τεχνητή νοημοσύνη. Αυτοί οι οργανισμοί περιλαμβάνουν: το Center for Security and Emerging Technology at Georgetown University, το LinkedIn, το NetBase Quid, το Lightcast και τη McKinsey.

Η έκθεση του 2023⁷² περιελάμβανε νέα ανάλυση για τα μοντέλα θεμελίωσης, συμπεριλαμβανομένου του κόστους γεωπολιτικής και εκπαίδευσης, των περιβαλλοντικών επιπτώσεων των συστημάτων τεχνητής νοημοσύνης, της εκπαίδευσης στην τεχνητή νοημοσύνη στην κατηγορία K-12⁷³ (υποχρεωτική εκπαίδευση ΗΠΑ) και των τάσεων της κοινής γνώμης για την τεχνητή νοημοσύνη. Ο AI Index διεύρυνε επίσης την παρακολούθηση της παγκόσμιας νομοθεσίας για την τεχνητή νοημοσύνη από 25 χώρες το 2022, σε 127 το 2023 και μπορεί κανείς να βρει όλες τις σχετικές πληροφορίες στο <https://aiindex.stanford.edu/report/>.

Επίσης διαπίστωσε, μεταξύ άλλων, τα εξής που προκύπτουν από τη χρήση αντίστοιχων εργαλείων μέτρησης των σχετικών δεικτών-στόχων:

1. Από το 2016 παρατηρείται μια σταθερή αύξηση του συνολικού αριθμού των μετρικών για τη δικαιοσύνη και τη μεροληψία της TN.
2. Η τεχνητή νοημοσύνη συγχρόνως βοηθά και βλάπτει το περιβάλλον, καθώς πολλές εφαρμογές καταναλώνουν πολλή ενέργεια αλλά κάποια μοντέλα TN μπορούν να προτείνουν τρόπους βελτιστοποίησης της χρήσης ενέργειας.
3. Ο αριθμός των περιστατικών που αφορούν την κατάχρηση της τεχνητής νοημοσύνης αυξάνεται ραγδαία. Η αύξηση αυτή αποτελεί απόδειξη τόσο του αυξανόμενου βαθμού στον οποίο η TN διαπλέκεται με τον φυσικό κόσμο όσο και της αυξανόμενης συνειδητοποίησης των τρόπων με τους οποίους η TN μπορεί να χρησιμοποιηθεί καταχρηστικά, όσον αφορά τη δεοντολογία. Η δραματική αύξηση εγείρει επίσης ένα σημαντικό ζήτημα, ότι δηλαδή καθώς η ευαισθητοποίηση έχει αυξηθεί, η παρακολούθηση των συμβάντων ασφαλείας και των βλαβών έχει επίσης βελτιωθεί, γεγονός που υποδηλώνει ότι τα παλαιότερα κρούσματα μπορεί να μην είχαν αναφερθεί επαρκώς.
4. Αύξηση χρήσης των εργαλείων ανίχνευσης τοξικού λόγου, κάτι που μπορεί να σημαίνει αύξηση των περιστατικών αλλά και αύξηση της ευαισθητοποίησης.
5. Μέτρηση εκπροσώπησης φύλου του chatbot (σύμφωνα με έρευνα η επιλογή γυναικείας φωνής ενέχει τον κίνδυνο να αναπαράγονται στερεότυπα και προκαταλήψεις).

4.5.9 Δείκτες στον Ελληνικό Νόμο 4961/2022

Ερχόμενοι στον Έλληνα νομοθέτη, ο Νόμος 4961/2022⁷⁴ μεταξύ άλλων ρυθμίσεων, θεσπίζει το νομοθετικό πλαίσιο για την αξιοποίηση των δυνατοτήτων της τεχνητής νοημοσύνης από φορείς του δημόσιου και ιδιωτικού τομέα και περιλαμβάνει ρυθμίσεις για την ανάπτυξη της τεχνητής νοημοσύνης. Θεσπίζει γενικές ρυθμίσεις για τη δυνατότητα χρήσης συστημάτων τεχνητής νοημοσύνης από φορείς του δημόσιου τομέα και συστήνει Παρατηρητήριο Τεχνητής Νοημοσύνης στο άρθρο 14.

⁷²<https://aiindex.stanford.edu/report/>

⁷³<https://en.wikipedia.org/wiki/K%E2%80%9312>

⁷⁴Νόμος 4961/2022 «Αναδυόμενες τεχνολογίες πληροφορικής και επικοινωνιών, ενίσχυση της ψηφιακής διακυβέρνησης και άλλες διατάξεις», ΦΕΚ Α' 146/27.07.2022, https://www.et.gr/api/DownloadFeksApi/?fek_pdf=20220100146

Στο πλαίσιο της αποστολής του το Παρατηρητήριο, ανάμεσα σε άλλα καθήκοντα, καταρτίζει και επικαιροποιεί βασικούς δείκτες επίδοσης (Key Performance Indicators - KPIs) των δραστηριοτήτων σχετικά με την ΤΝ στην Ελλάδα. Κατ' εφαρμογή των εξουσιοδοτικών διατάξεων του εν λόγω νόμου (άρθ. 113, παρ. 3) καταρτίζεται η Υπουργική Απόφαση για τον καθορισμό της μεθοδολογίας μέτρησης των βασικών δεικτών επίδοσης.

Ορίζεται ότι οι δείκτες που θα παρακολουθεί το Παρατηρητήριο για την Τεχνητή Νοημοσύνη σχετίζονται με τη βιομηχανική, τεχνολογική και ερευνητική ικανότητα της χώρας στον τομέα της ΤΝ, τις πρωτοβουλίες για πολιτικές που σχετίζονται με την ΤΝ, την υιοθέτηση και τις τεχνικές εξελίξεις της ΤΝ και τον αντίκτυπό της. Από αυτές ξεχωρίζουμε δύο πεδία:

1. Τις κοινωνικές πτυχές όπως την ποικιλομορφία στην έρευνα και στα συνέδρια, τη διαφορετικότητα των φύλων, την περιφερειακή διείσδυση της ΤΝ, την επιχειρηματική πολυμορφία, την επίδραση στην απασχόληση και στις αμοιβές, την επίπτωση των δραστηριοτήτων ΤΝ στα θεμελιώδη δικαιώματα.
2. Την εκπαίδευση, όπως εκπαιδευτικές δραστηριότητες σχετικά με την ανάπτυξη και χρήση της ΤΝ που λαμβάνουν χώρα στην Ελλάδα σε όλα τα επίπεδα εκπαίδευσης, τόσο στο ιδιωτικό τομέα όσο και για στελέχη του δημόσιου τομέα.

4.5.10 Δείκτες που αντλούμε από την ταξινόμηση των συστημάτων ΤΝ του ΟΟΣΑ

Ο Οργανισμός Οικονομικής Συνεργασίας και Ανάπτυξης υιοθέτησε το 2019 τις Αρχές για μια ΤΝ καινοτόμο, αξιόπιστη, με σεβασμό στα ανθρώπινα δικαιώματα και τις δημοκρατικές αξίες. Ειδικότερα, σύμφωνα με την αρχή "Ανθρωποκεντρικές αξίες και δικαιοσύνη" (Αρχή 1.2)⁷⁵, "Οι φορείς ΤΝ θα πρέπει να σέβονται το κράτος δικαίου, τα ανθρώπινα δικαιώματα και τις δημοκρατικές αξίες"⁷⁶. Οι Αρχές για την ΤΝ του ΟΟΣΑ αποτέλεσαν την πρώτη δέσμη προτύπων ΤΝ που οι κυβερνήσεις δεσμεύτηκαν να ενσωματώσουν στη χάραξη πολιτικής με σκοπό να προωθήσουν την καινοτόμο και αξιόπιστη χρήση της ΤΝ.

ΑΝΘΡΩΠΟΙ & ΠΛΑΝΗΤΗΣ	Κριτήρια	Περιγραφή
ΧΡΗΣΤΕΣ	Χρήστες του συστήματος AI	Ποιο είναι το επίπεδο ικανοτήτων των χρηστών που αλληλεπιδρούν με το σύστημα;
ΕΝΔΙΑΦΕΡΟΜΕΝΟΙ	Επηρεαζόμενοι ενδιαφερόμενοι	Ποιος επηρεάζεται από το σύστημα (π.χ. καταναλωτές, εργαζόμενοι, κυβερνητικές υπηρεσίες);
ΠΡΟΑΙΡΕΤΙΚΟΤΗΤΑ	Προαιρετικότητα και επανόρθωση	Μπορούν οι χρήστες να εξαιρεθούν, π.χ. να αλλάξουν σύστημα; Μπορούν οι χρήστες να αμφισβητήσουν ή να διορθώσουν τα αποτελέσματα;
ΑΝΘΡΩΠΙΝΑ ΔΙΚΑΙΩΜΑΤΑ	Ανθρώπινα δικαιώματα και δημοκρατικές αξίες	Μπορούν οι εκροές του συστήματος να επηρεάσουν τα θεμελιώδη ανθρώπινα δικαιώματα (π.χ. ανθρώπινη αξιοπρέπεια, ιδιωτική ζωή, ελευθερία έκφρασης, μη διάκριση, δίκαιη δίκη, επανόρθωση, ασφάλεια);
ΕΥΗΜΕΡΙΑ & ΠΕΡΙΒΑΛΛΟΝ	Ευημερία, κοινωνία και περιβάλλον	Μπορούν οι εκροές του συστήματος να επηρεάσουν τομείς της ζωής που σχετίζονται με την ευημερία (π.χ. ποιότητα εργασίας, περιβάλλον, υγεία, κοινωνικές αλληλεπιδράσεις, συμμετοχή των πολιτών, εκπαίδευση);
ΔΙΑΦΟΡΑ	{Δυναμικό μετατόπισης}	Θα μπορούσε το σύστημα να αυτοματοποιήσει εργασίες που εκτελούνται ή εκτελούνταν από ανθρώπους;

Εικόνα 4.6. Βασικές διαστάσεις του πλαισίου του ΟΟΣΑ για την ταξινόμηση των συστημάτων τεχνητής νοημοσύνης

⁷⁵<https://oecd.ai/en/dashboards/ai-principles/P6>

⁷⁶ Αυτές περιλαμβάνουν την ελευθερία, την αξιοπρέπεια και την αυτονομία, την προστασία της ιδιωτικής ζωής και των δεδομένων, την απαγόρευση των διακρίσεων και την ισότητα, την ποικιλομορφία, τη δικαιοσύνη, την κοινωνική δικαιοσύνη και τα διεθνώς αναγνωρισμένα εργασιακά δικαιώματα.

Σε συνέχεια αυτών, ο ΟΟΣΑ ανέπτυξε ένα Πλαίσιο Ταξινόμησης των συστημάτων ΤΝ⁷⁷ που μπορούν να εφαρμοσούν ως εργαλείο για το έργο τους οι φορείς χάραξης πολιτικής, οι ρυθμιστικές αρχές, οι νομοθέτες και άλλοι για τον χαρακτηρισμό των συστημάτων ΤΝ. Το Πλαίσιο Ταξινόμησης⁷⁸ συνδέει τα χαρακτηριστικά των συστημάτων ΤΝ με τις Αρχές ΤΝ του ΟΟΣΑ. Μεταξύ των σκοπών του Πλαισίου Ταξινόμησης που θέτει ο ΟΟΣΑ συγκαταλέγεται η κοινή κατανόηση της τεχνητής νοημοσύνης και των μετρήσεων. Στη λογική αυτή υπάγεται η δημιουργία ενός Καταλόγου εργαλείων και μετρήσεων για αξιόπιστη τεχνητή νοημοσύνη (Catalogue of Tools & Metrics for Trustworthy AI)⁷⁹ με σκοπό να βοηθηθούν οι φορείς να αναπτύξουν και να χρησιμοποιούν αξιόπιστα συστήματα τεχνητής νοημοσύνης που είναι δίκαια, διαφανή, εξηγήσιμα, στιβαρά, ασφαλή, προστατευμένα και σέβονται τα ανθρώπινα δικαιώματα.

Το πλαίσιο ταξινομεί τα συστήματα και τις εφαρμογές τεχνητής νοημοσύνης με βάση τις ακόλουθες διαστάσεις, από τις οποίες κάθε μία έχει τις δικές της ιδιότητες ή υποδιαστάσεις: Δεδομένα & Είσοδος, Οικονομικό Πλαίσιο, Άνθρωπος & Πλανήτης, Μοντέλο ΤΝ και Εργασία & Έξοδος.

Στο πλαίσιο της αναζήτησής μας, η διάσταση που μας ενδιαφέρει φέρει τον τίτλο “Άνθρωποι και Πλανήτης” και εξετάζει τη δυνατότητα των εφαρμοζόμενων συστημάτων ΤΝ να προωθήσουν ανθρωποκεντρική, αξιόπιστη ΤΝ που ωφελεί τους ανθρώπους και τον πλανήτη. Σε κάθε πλαίσιο, προσδιορίζει τα άτομα και τις ομάδες που αλληλεπιδρούν ή επηρεάζονται από ένα εφαρμοσμένο σύστημα ΤΝ. Τα βασικά χαρακτηριστικά περιλαμβάνουν τους χρήστες και τα επηρεαζόμενα ενδιαφερόμενα μέρη, την προαιρετικότητα της εφαρμογής και τον τρόπο με τον οποίο επηρεάζει τα ανθρώπινα δικαιώματα, το περιβάλλον, την ευημερία, την κοινωνία και τον κόσμο της εργασίας. Αναλυτικότερα παρουσιάζονται ως εξής:

- Χρήστες του συστήματος ΑΙ: Ποιο είναι το επίπεδο ικανοτήτων των χρηστών που αλληλεπιδρούν με το σύστημα;
- Ποιος επηρεάζεται από το σύστημα: π.χ. καταναλωτές, εργαζόμενοι, κυβερνητικές υπηρεσίες;
- Προαιρετικότητα και επανόρθωση: Μπορούν οι χρήστες να εξαιρεθούν ή να αλλάξουν σύστημα; Μπορούν οι χρήστες να αμφισβητήσουν ή να διορθώσουν τα αποτελέσματα;
- Ανθρώπινα δικαιώματα και δημοκρατικές αξίες: Μπορούν οι έξοδοι του συστήματος να επηρεάσουν τα θεμελιώδη ανθρώπινα δικαιώματα (π.χ. ανθρώπινη αξιοπρέπεια, ιδιωτική ζωή, ελευθερία έκφρασης);
- Ευημερία, κοινωνία και περιβάλλον: Μπορούν οι έξοδοι του συστήματος να επηρεάσουν τομείς της ζωής που σχετίζονται με την ευημερία (π.χ. ποιότητα εργασίας, περιβάλλον, υγεία, κοινωνικές αλληλεπιδράσεις, συμμετοχή των πολιτών, εκπαίδευση);

Εις επίρρωση των παραπάνω είναι σημαντικό να αναφερθεί η επικαιροποίηση των ανωτέρω αρχών⁸⁰ στην οποία προέβη ο ΟΟΣΑ στις 3 Μαΐου 2024 για να ανταποκριθεί στις πρόσφατες εξελίξεις των τεχνολογιών τεχνητής νοημοσύνης, ιδίως στην εμφάνιση της παραγωγικής τεχνητής νοημοσύνης γενικού σκοπού, ώστε να αντιμετωπίζονται πιο άμεσα οι προκλήσεις που σχετίζονται με το απόρρητο, τα δικαιώματα πνευματικής ιδιοκτησίας, την ασφάλεια και την ακεραιότητα των πληροφοριών. Τα βασικά στοιχεία της επικαιροποίησης εστιάζουν στα ακόλουθα σημεία και σαφέστατα μπορούν να τεκμηριώσουν τη δημιουργία ανάλογων δεικτών σε σχέση με:

⁷⁷<https://oecd.ai/en/classification>

⁷⁸OECD (2022), “OECD Framework for the Classification of AI systems”, OECD Digital Economy Papers, No. 323, OECD Publishing, Paris, <https://doi.org/10.1787/cb6d9eca-en>

⁷⁹<https://oecd.ai/en/catalogue>

⁸⁰OECD updates AI Principles to stay abreast of rapid technological developments, 3 Μαΐου 2024, <https://www.oecd.org/newsroom/oecd-updates-ai-principles-to-stay-abreast-of-rapid-technological-developments.htm>

- διασφαλίσεις απέναντι στην αδικαιολόγητη διακινδύνευση και στην ανεπιθύμητη συμπεριφορά που μπορούν να παραχθούν από ένα σύστημα TN,
- την αντιμετώπιση της παραπληροφόρησης,
- την υπεύθυνη επιχειρηματική συμπεριφορά σε όλο τον κύκλο ζωής του συστήματος τεχνητής νοημοσύνης,
- τη διαφάνεια,
- την περιβαλλοντική βιωσιμότητα,
- τη συνεργασία των δικαιοδοσιών για την προώθηση της διαλειτουργικής διακυβέρνησης καθώς ο αριθμός των πρωτοβουλιών πολιτικής για την τεχνητή νοημοσύνη αυξάνεται παγκοσμίως.

4.6 Συμπέρασμα

Τεκμηριώθηκε η θεσμοθέτηση παρατηρητηρίων για την TN μέσα από θεσμικά και κατευθυντήρια κείμενα της ΕΕ, δεδομένου ότι παρατηρείται ολοένα και περισσότερο καθώς εξυπηρετούν την αναγκαιότητα παρακολούθησης και συλλογής πληροφοριών για τον αντίκτυπο της χρήσης της TN με σκοπό την οικονομική ανάπτυξη, την κοινωνική και ατομική ευημερία αλλά και την αποτύπωση της επίδρασής της στην ψηφιακή δημόσια διοίκηση και την προώθηση της διαλειτουργικότητας. Στόχο έχουν να αποτελέσουν χρήσιμο σημείο αναφοράς για όλους τους ενδιαφερόμενους φορείς, τα στελέχη του δημόσιου τομέα, τους υπεύθυνους χάραξης πολιτικής, τις ιδιωτικές εταιρείες, το ακαδημαϊκό και ερευνητικό περιβάλλον, που ενδιαφέρονται για τις τελευταίες τεχνολογικές εξελίξεις υπό το πρίσμα της βελτίωσης των λειτουργιών και της παροχής υπηρεσιών του δημόσιου τομέα.

Μία προτεραιότητα που προκύπτει σαφώς από τα προγραμματικά ευρωπαϊκά κείμενα είναι η διασφάλιση υψηλού επιπέδου προστασίας προκειμένου να επιτευχθεί αξιόπιστη και ανθρωποκεντρική τεχνητή νοημοσύνη και η απαρίθμηση μάλιστα γενικών αρχών που πρέπει να ισχύουν για όλα τα συστήματα τεχνητής νοημοσύνης. Ορισμένες από αυτές αφορούν άμεσα τα θεμελιώδη δικαιώματα, τα οποία αναδεικνύονται σε βασικό άξονα για τον έλεγχο συμμόρφωσης με τις ευρωπαϊκές αξιακές και νομικές επιταγές. Στην υπηρεσία ενός εκ των προτέρων ελέγχου για την ικανοποίηση αυτών των αρχών, τίθεται η θεσμοθέτηση τεχνικών προτύπων για την αποτελεσματικότερη αξιολόγηση των επιπτώσεων στα θεμελιώδη δικαιώματα ενώ για τον εκ των υστέρων έλεγχο χρησιμοποιούνται βασικοί δείκτες επίδοσης.

Σε συνέχεια των παραπάνω η Πράξη για την TN δείχνει ως μια πρώτη προσπάθεια ενσωμάτωσης των θεμελιωδών δικαιωμάτων σε μια προσέγγιση ασφάλειας προϊόντων, χρησιμοποιώντας τα εναρμονισμένα πρότυπα. Πρόκειται για μια διαδικασία η οποία βρίσκεται σε εξέλιξη μετά το αίτημα της Ευρωπαϊκής Επιτροπής προς τους αρμόδιους φορείς ευρωπαϊκών προτύπων και για ορισμένους αυτό μετατοπίζει τα θέματα ασφάλειας στον ηθικό και νομικό χώρο των δικαιωμάτων, με τον κίνδυνο να απλουστευθεί η ηθική και νομική κρίση σε τεχνικό εργαλείο. Συνεπώς είναι σημαντικό οι φορείς τυποποίησης να αναπτύξουν πρότυπα που να συμβάλλουν στην προστασία των θεμελιωδών δικαιωμάτων μέσω της διάδοσης καλών πρακτικών, αλλά να μην επεκτείνονται σε κοινωνικές, νομικές και αξιακές κρίσεις.

Όσον αφορά την εκ των υστέρων αποτίμηση της χρήσης της TN, μετά από τη δειγματοληπτική περιήγηση σε διαφορετικές πηγές ορισμού σχετικών δεικτών επίδοσης, καταλήγουμε ότι πρόκειται για ένα ζήτημα που σχετίζεται με τον προσανατολισμό, τις προτεραιότητες και τους στόχους του φορέα που τους θέτει και δεν αποτελεί μια μονοσήμαντη ή αναπόδραστη ακολουθία βημάτων. Ειδικά κατά την εφαρμογή δημοσίων πολιτικών, η επιλογή δεικτών απεικονίζει την επιδιωκόμενη

αποτελεσματικότητα σε συσχετισμό με τους διαθέσιμους πόρους και το κόστος. Το βέβαιο είναι ότι αν και γνωρίζουμε ότι η χρήση της τεχνητής νοημοσύνης προσφέρει τεράστιες ευκαιρίες για δημιουργία αξίας, δεν ξέρουμε ακριβώς πώς να χρησιμοποιούμε την τεχνητή νοημοσύνη υπεύθυνα⁸¹. Θεσπίζοντας δείκτες ανθρωποκεντρικούς και κοινωνικούς, συμπεριλαμβανομένων των περιβαλλοντικών και ηθικών, που αποτυπώνουν την υπεύθυνη χρήση, μετατοπίζουμε το ενδιαφέρον στον επόμενο κρίσιμο τομέα που επιβάλλεται να μετρηθεί και αφορά την υπευθυνότητα, την αξιοπιστία και την ανθρωποκεντρική διάσταση της ΤΝ. Μάλιστα για καλύτερη εποπτεία αυτής της ποικιλομορφίας δεικτών, οι οποίοι παραμένουν διασκορπισμένοι σε ετερόκλητες πηγές, ακολουθεί πίνακας όπου έχουν κωδικοποιηθεί όλοι οι προαναφερθέντες δείκτες χωρισμένοι σε δύο τύπους, οικονομικούς-τεχνολογικούς και ανθρωποκεντρικούς-κοινωνικούς, ανά πηγή και με επισήμανση του υποκεφαλαίου της παρούσας εργασίας όπου αναφέρονται.

Δεδομένου ότι η επιβολή ρυθμίσεων στο πεδίο της ψηφιακής στρατηγικής έδωσε στην ΕΕ την ισχύ να θέσει παγκόσμια πρότυπα όπως αναλύθηκε παραπάνω υπό τον τίτλο “Φαινόμενο των Βρυξελλών” (Brussels Effect), αναμένεται ότι η καθιέρωση ρητών ανθρωποκεντρικών δεικτών επίδοσης των συστημάτων ΤΝ από την ΕΕ θα επηρεάσει το παγκόσμιο τοπίο ως προς την οπτική μιας ΤΝ για την ευημερία και τον σεβασμό των θεμελιωδών δικαιωμάτων και του κράτους δικαίου όπου γεννιούνται νέες ανάγκες για στενότερη συνεργασία, αμοιβαία κατανόηση και άρση των αυστηρών ορίων μεταξύ του τεχνικού και του νομικού κόσμου. Ήδη φαίνεται να χαράζει το μονοπάτι προς αυτή την κατεύθυνση η επικαιροποίηση των αρχών που θα πρέπει να διέπουν την ΤΝ, στην οποία προέβη ο ΟΟΣΑ στις 3 Μαΐου 2024.

Ακολουθεί, κλείνοντας, μια συγκέντρωση όλων των δεικτών που αναφέρθηκαν για την παρακολούθηση και αξιολόγηση της χρήσης τεχνολογιών τεχνητής νοημοσύνης (ΤΝ) σε διάφορους τομείς και κατηγορίες, όπως οικονομική και τεχνολογική ανάπτυξη, ανθρωποκεντρικές και κοινωνικές επιπτώσεις, και δείκτες που σχετίζονται με τη βιώσιμη ανάπτυξη. Αυτοί οι δείκτες περιλαμβάνουν ποσοτικά στοιχεία σχετικά με την υιοθέτηση και χρήση τεχνολογιών ΤΝ, καθώς και ανθρωποκεντρικούς-κοινωνικούς δείκτες που εστιάζουν στην αξιοπιστία, την υπευθυνότητα, την αλληλεγγύη και την περιβαλλοντική βιωσιμότητα.

⁸¹<https://www.tortoisemedia.com/intelligence/tortoiseairesponsibility/>

Συγκεντρωτικός πίνακας Βασικών Δεικτών Επίδοσης για την ΤΝ (KPIs for AI)

Κεφ.	A/A	Πηγή	Κατηγορία	Διάσταση	Δείκτης
4.5.3	1	Οι Δείκτες για την ΤΝ που προκύπτουν από τον DESI σε συνδυασμό με τη EUROSTAT	Οικονομικός-Τεχνολογικός	Εταιρείες που χρησιμοποιούν τεχνολογίες ΤΝ ανά τύπο τεχνολογίας ΤΝ	ΑΙ τεχνολογίες που αυτοματοποιούν διάφορες ροές εργασίας ή βοηθούν στη λήψη αποφάσεων (π.χ. ρομποτική αυτοματοποίηση διαδικασιών λογισμικού με βάση την τεχνητή νοημοσύνη)
	2		Οικονομικός-Τεχνολογικός		Μηχανική μάθηση (π.χ. βαθιά μάθηση) για την ανάλυση δεδομένων
	3		Οικονομικός-Τεχνολογικός		Τεχνολογίες ανάλυσης του γραπτού λόγου (π.χ. εξόρυξη κειμένου)
	4		Οικονομικός-Τεχνολογικός		Τεχνολογίες τεχνητής νοημοσύνης που αναγνωρίζουν αντικείμενα ή πρόσωπα με βάση εικόνες (αναγνώριση εικόνας, επεξεργασία εικόνας)
	5		Οικονομικός-Τεχνολογικός		Τεχνολογίες τεχνητής νοημοσύνης που μετατρέπουν τον προφορικό λόγο σε μορφή αναγνώσιμη από μηχανήματα (αναγνώριση ομιλίας)
	6		Οικονομικός-Τεχνολογικός		Τεχνολογίες που επιτρέπουν τη φυσική μετακίνηση μηχανών μέσω αυτόνομων αποφάσεων που βασίζονται στην παρατήρηση του περιβάλλοντος (π.χ. αυτοκινούμενα οχήματα)
	7		Οικονομικός-Τεχνολογικός		Τεχνολογίες τεχνητής νοημοσύνης που παράγουν γραπτό ή προφορικό λόγο (παραγωγή φυσικής γλώσσας)
	8		Οικονομικός-Τεχνολογικός	Μέγεθος εταιρείας	Μικρή - μεσαία - μεγάλη
	9		Οικονομικός-Τεχνολογικός	Είδος σκοπού για τη χρήση τεχνολογιών ΤΝ	Για την ασφάλεια ΤΠΕ
	10		Οικονομικός-Τεχνολογικός		Για διαδικασίες παραγωγής
	11		Οικονομικός-Τεχνολογικός		Για την οργάνωση των διαδικασιών διοίκησης επιχειρήσεων
	12		Οικονομικός-Τεχνολογικός		Για μάρκετινγκ ή πωλήσεις
	13		Οικονομικός-Τεχνολογικός		Για τη διαχείριση των επιχειρήσεων
	14		Οικονομικός-Τεχνολογικός		Για την εφοδιαστική αλυσίδα
	15		Οικονομικός-Τεχνολογικός		Για τη διαχείριση ανθρώπινου δυναμικού ή την πρόσληψη προσωπικού

Κεφ.	A/A	Πηγή	Κατηγορία	Διάσταση	Δείκτης
	16		Οικονομικός-Τεχνολογικός	Οικονομική δραστηριότητα της εταιρείας	Πληροφορίες και επικοινωνία
	17		Οικονομικός-Τεχνολογικός		Δραστηριότητες παροχής επαγγελματικών, επιστημονικών και τεχνικών υπηρεσιών
	18		Οικονομικός-Τεχνολογικός		Ηλεκτρικό ρεύμα, φυσικό αέριο, ατμός, κλιματισμός και ύδρευση
	19		Οικονομικός-Τεχνολογικός		Διοικητικές και υποστηρικτικές δραστηριότητες
	20		Οικονομικός-Τεχνολογικός		Βιοτεχνία
	21		Οικονομικός-Τεχνολογικός		Κτηματομεσιτικές δραστηριότητες
	22		Οικονομικός-Τεχνολογικός		Διαμονή
	23		Οικονομικός-Τεχνολογικός		Λιανικό εμπόριο
	24		Οικονομικός-Τεχνολογικός		Χονδρικό και λιανικό εμπόριο-επισκευή μηχανοκίνητων οχημάτων και μοτοσυκλετών
	25		Οικονομικός-Τεχνολογικός		Κατασκευές
	26		Οικονομικός-Τεχνολογικός		Μεταφορά και αποθήκευση
	27		Οικονομικός-Τεχνολογικός	Τρόπος απόκτησης τεχνολογιών TN	Οι τεχνολογίες TN ήταν εμπορικό λογισμικό ή συστήματα έτοιμα προς χρήση
	28		Οικονομικός-Τεχνολογικός		Οι τεχνολογίες TN αναπτύχθηκαν ή τροποποιήθηκαν από εξωτερικούς παρόχους
	29		Οικονομικός-Τεχνολογικός		Οι τεχνολογίες TN ήταν εμπορικό λογισμικό ή συστήματα που τροποποιήθηκαν από τους ίδιους τους υπαλλήλους.
	30		Οικονομικός-Τεχνολογικός		Οι τεχνολογίες TN αναπτύχθηκαν από τους ίδιους τους υπαλλήλους
	31		Οικονομικός-Τεχνολογικός		Οι τεχνολογίες τεχνητής νοημοσύνης ήταν λογισμικό ή συστήματα ανοικτού κώδικα που τροποποιήθηκαν από τους ίδιους τους υπαλλήλους

Κεφ.	A/A	Πηγή	Κατηγορία	Διάσταση	Δείκτης
4.5.4	32	European Commission, Ipsos Survey, European enterprise survey on the use of technologies based on artificial intelligence, 2020	Οικονομικός-Τεχνολογικός	Ενημέρωση για την τεχνητή νοημοσύνη	
	33		Οικονομικός-Τεχνολογικός	Υιοθέτηση της TN	
	34		Οικονομικός-Τεχνολογικός	Προμήθεια TN	
	35		Οικονομικός-Τεχνολογικός	Εξωτερικά εμπόδια στην υιοθέτηση της TN	
	36		Οικονομικός-Τεχνολογικός	Εσωτερικά εμπόδια στην υιοθέτηση της TN	
4.5.5	37	Πρόταση Κανονισμού “Πράξη για την TN”, Κώδικες Δεοντολογίας	Ανθρωπ/κός - Κοινωνικός	Δείκτες επίδοσης που αφορούν τη χρήση της TN	Γραμματισμός (αλφαριθμητισμός) του προσωπικού στον τομέα της TN
	38		Ανθρωπ/κός-Κοινωνικός		Εκτίμηση επιπτώσεων σε ευάλωτα άτομα ή ομάδες και ενίσχυση της προσβασιμότητας
	39		Ανθρωπ/κός-Κοινωνικός		Εκτίμηση επιπτώσεων στη διαφορετικότητα και υποστήριξη ισότητας των φύλων
	40		Ανθρωπ/κός-Κοινωνικός		Πιθανότητα μεροληψίας
	41		Ανθρωπ/κός-Κοινωνικός		Συμπεριληπτικός σχεδιασμός υπέρ ποικιλόμορφων ομάδων
	42		Ανθρωπ/κός-Κοινωνικός		Επιπτώσεις στη δημοκρατία και τους πολιτικούς θεσμούς
	43		Ανθρωπ/κός-Κοινωνικός		Περιβαλλοντική βιωσιμότητα

Κεφ.	A/A	Πηγή	Κατηγορία	Διάσταση	Δείκτης
4.5.5	44	Πρόταση Κανονισμού “Πράξη για την ΤΝ”, Νομοθετικό Δημοσιονομικό Δελτίο (στο σημείο 1.4.4.)	Ανθρωπ/κός-Κοινωνικός	Δείκτες επίδοσης της υλοποίησης της “Πράξης για την ΤΝ”	Αριθμός σοβαρών περιστατικών ή επιδόσεων ΤΝ που συνιστούν σοβαρό περιστατικό ή παραβίαση των υποχρεώσεων στον τομέα των θεμελιωδών δικαιωμάτων (ανά εξάμηνο) ανά πεδίο εφαρμογών: α) σε απόλυτους αριθμούς, β) ως ποσοστό των εφαρμογών που αναπτύχθηκαν, γ) ως ποσοστό των ενδιαφερόμενων πολιτών
	45		Οικονομικός-Τεχνολογικός		Α) Σύνολο επενδύσεων στον τομέα της τεχνητής νοημοσύνης στην ΕΕ (ετησίως)
	46		Οικονομικός-Τεχνολογικός		β) Σύνολο επενδύσεων στον τομέα της τεχνητής νοημοσύνης ανά κράτος μέλος (ετησίως)
	47		Οικονομικός-Τεχνολογικός		γ) Ποσοστό εταιρειών που χρησιμοποιούν ΤΝ (ετησίως)
	48		Οικονομικός-Τεχνολογικός		δ) Ποσοστό ΜΜΕ που χρησιμοποιούν ΤΝ (ετησίως)
4.5.5	49	Πρόταση Κανονισμού “Πράξη για την ΤΝ” - Δείκτες ΣΒΑ	Ανθρωπ/κός - Κοινωνικός	Δείκτες ΣΒΑ Ηνωμένων Εθνών [ΔΕΝ ΣΥΜΠΕΡΙΛΗΦΘΗΚΕ ΣΤΟ ΤΕΛΙΚΟ ΚΕΙΜΕΝΟ]	Η Επιτροπή δικαιούται να αναπτύξει μεθοδολογία που θα συμβάλει στον καθορισμό βασικών δεικτών επίδοσης και αναφοράς για τους Στόχους Βιώσιμης Ανάπτυξης (ΣΒΑ)
4.5.6	50	AI Watch Index 2021	Οικονομικός - Τεχνολογικός	Παγκόσμια άποψη για το τοπίο της ΤΝ	Οικονομικοί δείκτες AI
	51		Οικονομικός-Τεχνολογικός		Ένταση παίκτη Τεχνητής Νοημοσύνης
	52		Οικονομικός-Τεχνολογικός		Τομείς εξειδίκευσης της ΤΝ: συγκριτικό πλεονέκτημα σε θεματικούς τομείς της ΤΝ
	53		Οικονομικός-Τεχνολογικός		Θεματικές εστίες ΤΝ
	54		Οικονομικός-Τεχνολογικός		Συγκριτικό πλεονέκτημα της ΕΕ στο εμπόριο βιομηχανικής ρομποτικής
	55		Οικονομικός-Τεχνολογικός		Επενδύσεις Τεχνητής Νοημοσύνης στην ΕΕ
	56		Οικονομικός-Τεχνολογικός	Βιομηχανία	Προφίλ των επιχειρήσεων ΤΝ
	57		Οικονομικός-Τεχνολογικός		Νεοσύστατες επιχειρήσεις ρομποτικής στην ΕΕ
	58		Οικονομικός-Τεχνολογικός	Έρευνα και Ανάπτυξη	Παίκτης ΤΝ στην Ε&Α ΤΝ
	59		Οικονομικός-Τεχνολογικός		Βαθμολογία δραστηριότητας Ε&Α ΤΝ

Κεφ.	A/A	Πηγή	Κατηγορία	Διάσταση	Δείκτης
	60		Οικονομικός-Τεχνολογικός		Συνεργασίες μεταξύ κρατών
	61		Οικονομικός-Τεχνολογικός		Συνεργασίες μεταξύ ομότιμων εταίρων
	62		Οικονομικός-Τεχνολογικός		Στρατηγική θέση στο δίκτυο συνεργασιών
	63		Οικονομικός-Τεχνολογικός	Τεχνολογία	Απόδοση της έρευνας για την τεχνητή νοημοσύνη
	64		Ανθρωπ/κός-Κοινωνικός		Δραστηριότητα προτυποποίησης
	65		Ανθρωπ/κός-Κοινωνικός	Κοινωνικές πτυχές	Δείκτης ποικιλομορφίας των φύλων
	66		Ανθρωπ/κός-Κοινωνικός		Δείκτης γεωγραφικής ποικιλότητας
	67		Ανθρωπ/κός-Κοινωνικός		Δείκτης επιχειρηματικής ποικιλομορφίας
	68		Ανθρωπ/κός-Κοινωνικός		Δείκτης ποικιλομορφίας συνεδρίων
	69		Ανθρωπ/κός-Κοινωνικός		TN σε πανεπιστημιακά προγράμματα στην ΕΕ
	70		Ανθρωπ/κός-Κοινωνικός		Πανεπιστημιακές έδρες με περιεχόμενο TN στην ΕΕ
	71		Ανθρωπ/κός-Κοινωνικός		Ένταση TN σε πανεπιστημιακές έδρες στην ΕΕ
4.5.7	72	STAFF WORKING PAPER: AI Act Impact Assessment	Ανθρωπ/κός-Κοινωνικός	Τα συστήματα τεχνητής νοημοσύνης είναι ασφαλή και σέβονται τα θεμελιώδη δικαιώματα και τις αξίες της ΕΕ	”Αριθμός σοβαρών περιστατικών ή επιδόσεων τεχνητής νοημοσύνης που συνιστούν σοβαρό περιστατικό ή παραβίαση υποχρεώσεων θεμελιωδών δικαιωμάτων (εξάμηνο) ανά πεδία εφαρμογών, υπολογιζόμενα α) σε απόλυτες τιμές, β) ως ποσοστό των εφαρμογών που έχουν αναπτυχθεί και γ) ως ποσοστό των ενδιαφερόμενων πολιτών”
	73		Ανθρωπ/κός-Κοινωνικός	Βελτίωση των μηχανισμών διακυβέρνησης και επιβολής	”Αριθμός ανακλήσεων ή αποσύρσεων συστημάτων τεχνητής νοημοσύνης από την αγορά (εξάμηνο). Ανά πεδία εφαρμογών, υπολογιζόμενος α) σε απόλυτες τιμές, β) ως ποσοστό των εφαρμογών που έχουν αναπτυχθεί και γ) ως ποσοστό των ενδιαφερόμενων πολιτών”
	74		Ανθρωπ/κός-Κοινωνικός		Επίπεδο εμπιστοσύνης στην τεχνητή νοημοσύνη (ετήσιο) (θετικός δείκτης)
	75		Οικονομικός-Τεχνολογικός		Αριθμός εθνικών νομοθεσιών που θα κατακερματίσουν την ενιαία αγορά (εξάμηνο) (αρνητικός δείκτης)

Κεφ.	A/A	Πηγή	Κατηγορία	Διάσταση	Δείκτης
	76		Οικονομικός-Τεχνολογικός		Συνολικές επενδύσεις τεχνητής νοημοσύνης στην ΕΕ (ετήσιο)
	77		Οικονομικός-Τεχνολογικός		Συνολική επένδυση τεχνητής νοημοσύνης ανά κράτος μέλος (ετήσια)
	78		Οικονομικός-Τεχνολογικός		Μερίδιο εταιρειών που χρησιμοποιούν τεχνητή νοημοσύνη (ετήσιο)
	79		Οικονομικός-Τεχνολογικός		Μερίδιο ΜΜΕ που χρησιμοποιούν τεχνητή νοημοσύνη (ετήσιο)
	80		Οικονομικός-Τεχνολογικός		Έργα που έχουν εγκριθεί μέσω ρυθμιστικών δοκιμαστηρίων (sandbox) και διατίθενται στην αγορά (ετήσιος)
	81		Οικονομικός-Τεχνολογικός		Αριθμός ΜΜΕ που συμβουλευόμαστε Κόμβους ψηφιακής καινοτομίας και Εγκαταστάσεις δοκιμών και πειραματισμού σχετικά με την ΤΝ
4.5.8	82	AI Index - Stanford HAI	Ανθρωπ/κός-Κοινωνικός		Αύξηση αριθμού των μετρικών για τη δικαιοσύνη και τη μεροληψία της ΤΝ
	83		Ανθρωπ/κός-Κοινωνικός		Αύξηση χρήσης των εργαλείων ανίχνευσης τοξικού λόγου
	84		Ανθρωπ/κός-Κοινωνικός		Μέτρηση εκπροσώπησης φύλου του chatbot (σύμφωνα με έρευνα η επιλογή γυναικείας φωνής αναπαράγει στερεότυπα)
	85		Ανθρωπ/κός-Κοινωνικός		Μοντέλα ΤΝ που βελτιστοποιούν τη χρήση ενέργειας υπέρ του περιβάλλοντος
4.5.9	86	Ελληνικός Νόμος 4961/2022, άρθρο 14	Οικονομικός-Τεχνολογικός		Ποικιλομορφία στην έρευνα και στα συνέδρια
	87		Ανθρωπ/κός-Κοινωνικός		Διαφορετικότητα των φύλων
	88		Οικονομικός-Τεχνολογικός		Περιφερειακή διείσδυση της ΤΝ
	89		Οικονομικός-Τεχνολογικός		Επιχειρηματική πολυμορφία
	90		Οικονομικός-Τεχνολογικός		Επίδραση στην απασχόληση και στις αμοιβές
	91		Ανθρωπ/κός-Κοινωνικός		Επίπτωση των δραστηριοτήτων ΤΝ στα θεμελιώδη δικαιώματα
	92		Ανθρωπ/κός-Κοινωνικός		Εκπαιδευτικές δραστηριότητες σχετικά με την ανάπτυξη και χρήση της ΤΝ σε όλα τα επίπεδα και τα είδη εκπαίδευσης, στον ιδιωτικό τομέα και σε στελέχη του δημόσιου τομέα
4.5.10	93	Ταξινόμηση συστημάτων ΤΝ του ΟΟΣΑ	Ανθρωπ/κός-Κοινωνικός	Χρήστες του συστήματος AI	Ποιο είναι το επίπεδο ικανοτήτων των χρηστών που αλληλεπιδρούν με το σύστημα;

Κεφ.	A/A	Πηγή	Κατηγορία	Διάσταση	Δείκτης
	94		Ανθρωπ/κός-Κοινωνικός	Επηρεαζόμενοι ενδιαφερόμενοι	Ποιος επηρεάζεται από το σύστημα (π.χ. καταναλωτές, εργαζόμενοι, κυβερνητικές υπηρεσίες);
	95		Ανθρωπ/κός-Κοινωνικός	Προαιρετικότητα και επανόρθωση	Μπορούν οι χρήστες να εξαιρεθούν, π.χ. να αλλάξουν σύστημα; Μπορούν οι χρήστες να αμφισβητήσουν ή να διορθώσουν τα αποτελέσματα;
	96		Ανθρωπ/κός-Κοινωνικός	Ανθρώπινα δικαιώματα και δημοκρατικές αξίες	Μπορούν οι εκροές του συστήματος να επηρεάσουν τα θεμελιώδη ανθρώπινα δικαιώματα (π.χ. ανθρωπινή αξιοπρέπεια, ιδιωτική ζωή, ελευθερία έκφρασης, μη διάκριση, δίκαιη δίκη, επανόρθωση, ασφάλεια);
	97		Ανθρωπ/κός-Κοινωνικός	Ενημερία, κοινωνία και περιβάλλον	Μπορούν οι εκροές του συστήματος να επηρεάσουν τομείς της ζωής που σχετίζονται με την ενημερία (π.χ. ποιότητα εργασίας, περιβάλλον, υγεία, κοινωνικές αλληλεπιδράσεις, συμμετοχή των πολιτών, εκπαίδευση);
	98		Ανθρωπ/κός-Κοινωνικός	Δυναμικό μετατόπισης	Θα μπορούσε το σύστημα να αυτοματοποιήσει εργασίες που εκτελούνται ή εκτελούνταν από ανθρώπους;
	99		Οικονομικός-Τεχνολογικός	Συλλογή δεδομένων	Τα δεδομένα και οι εισροές από εμπειρογνώμονες παρέχονται, παρατηρούνται, συντίθενται ή προκύπτουν;
	100		Οικονομικός-Τεχνολογικός		Είναι τα δεδομένα δυναμικά, στατικά, δυναμικά που ενημερώνονται από καιρό σε καιρό ή σε πραγματικό χρόνο;
	101		Ανθρωπ/κός-Κοινωνικός	Δικαιώματα	Είναι τα δεδομένα ιδιότητα, δημόσια ή προσωπικά δεδομένα (που σχετίζονται με αναγνωρίσιμο άτομο);
	102		Ανθρωπ/κός-Κοινωνικός		Εάν πρόκειται για δεδομένα προσωπικού χαρακτήρα, είναι ανώνυμα ή ψευδώνυμα;
	103		Οικονομικός-Τεχνολογικός	Δομή και μορφή	Είναι τα δεδομένα δομημένα, ημιδομημένα, σύνθετα δομημένα ή αδόμητα;
	104		Οικονομικός-Τεχνολογικός		Είναι η μορφή των δεδομένων και των μεταδεδομένων τυποποιημένη ή μη τυποποιημένη;
	105		Οικονομικός-Τεχνολογικός	Κλίμακα	Ποια είναι η κλίμακα του συνόλου δεδομένων;
	106		Οικονομικός-Τεχνολογικός	Ποιότητα και καταλληλότητα	Είναι το σύνολο δεδομένων κατάλληλο για το σκοπό του; Είναι επαρκές το μέγεθος του δείγματος; Είναι αρκετά αντιπροσωπευτικό και πλήρες; Πόσο θορυβώδη είναι τα δεδομένα;
	108		Ανθρωπ/κός-Κοινωνικός	Διαθεσιμότητα πληροφοριών μοντέλου	Είναι διαθέσιμες πληροφορίες σχετικά με το μοντέλο του συστήματος;

Κεφ.	A/A	Πηγή	Κατηγορία	Διάσταση	Δείκτης
	109		Οικονομικός-Τεχνολογικός	Δημιουργία μοντέλων από μη-χανική ή ανθρώπινη γνώση	Το σύστημα μαθαίνει με βάση κανόνες που έχουν συνταχθεί από τον άνθρωπο, από δεδομένα, μέσω εποπτευόμενης μάθησης, μέσω ενισχυτικής μάθησης;
	110		Ανθρωπ/κός-Κοινωνικός	Διαφάνεια και επεξηγηματικότητα	Εάν οι πληροφορίες είναι διαθέσιμες στους χρήστες ώστε να μπορούν να κατανοήσουν τα αποτελέσματα του μοντέλου;
	111	Επικαιροποίηση αρχών TN από τον ΟΟΣΑ	Ανθρωπ/κός-Κοινωνικός	Ανθρώπινη εποπτεία	Διασφαλίσεις απέναντι σε αδικαιολόγητη διακινδύνευση και ανεπιθύμητη συμπεριφορά που μπορούν να παραχθούν από ένα σύστημα TN
	112		Ανθρωπ/κός-Κοινωνικός	Παραπληροφόρηση	Αντιμετώπιση της παραπληροφόρησης
	113		Ανθρωπ/κός-Κοινωνικός	Υπεύθυνη επιχειρηματικότητα	Υπεύθυνη επιχειρηματική συμπεριφορά σε όλο τον κύκλο ζωής του συστήματος τεχνητής νοημοσύνης
	114		Ανθρωπ/κός-Κοινωνικός	Διαφάνεια	Διευκρίνιση των πληροφοριών και διαφάνεια
	115		Ανθρωπ/κός-Κοινωνικός	Περιβάλλον	Περιβαλλοντική βιωσιμότητα
	116		Ανθρωπ/κός-Κοινωνικός	Παγκόσμια διακυβερνησιμότητα	Συνεργασία των δικαιοδοσιών για την προώθηση της διαλειτουργικής διακυβέρνησης καθώς ο αριθμός των πρωτοβουλιών πολιτικής για την τεχνητή νοημοσύνη αυξάνεται παγκοσμίως

Πίνακας 4.4. Συγκεντρωτικός πίνακας Βασικών Δεικτών Επίδοσης για την TN (KPIs for AI)

Κεφάλαιο 5

Συμπεράσματα

Η Τεχνητή Νοημοσύνη ως σύνολο τεχνολογιών και εφαρμογών έχει προκαλέσει ραγδαίες εξελίξεις σε ποικίλες πτυχές της δημόσιας και ιδιωτικής ζωής. Η μέχρι τώρα εμπειρία αλλά και οι προβλέψεις για το μέλλον φανερώνουν τη δυναμική της να επιφέρει ριζικές μεταβολές στον τρόπο με τον οποίο δραστηριοποιούμαστε σε κάθε τομέα, μεταβολές οι οποίες έχουν διττό χαρακτήρα. Η θετική διάσταση δείχνει οφέλη σε όλο το φάσμα της κοινωνικής και οικονομικής ζωής, η αρνητική διάσταση όμως δείχνει πιθανές απειλές σε τομείς όπως οι ατομικές ελευθερίες, το κράτος δικαίου, η δημοκρατία, το περιβάλλον, η υγεία, η ασφάλεια και οι γεωπολιτικές ισορροπίες. Τους προβληματισμούς αυτούς επαληθεύει η αναδρομή στο ιστορικό της ΤΝ ως επιστήμης, από τις απαρχές της μέχρι το τέλος του 20ου αιώνα, όταν τέθηκαν οι επιστημονικές βάσεις για τη μετέπειτα εκθετική εξέλιξή της, που ευνοήθηκε και από άλλες συνθήκες όπως η αυξημένη υπολογιστική ισχύς και η αύξηση των δεδομένων.

Γίνεται πλέον αντιληπτό ότι επίμαχο θέμα είναι το μοντέλο διακυβέρνησης της ΤΝ προκειμένου να εξισορροπηθούν οι προαναφερθείσες αντικρουόμενες διαστάσεις, κάτι που δεν μπορεί να επιτευχθεί προτού προσδιοριστεί ένας κοινά αποδεκτός ορισμός για την ΤΝ προκειμένου να εδραιωθούν η ασφάλεια δικαίου, η συναντίληψη και ένα κοινό λεξιλόγιο για το σκοπό της ρύθμισής της. Μετά από μακρά πορεία εννοιολογικών αναζητήσεων από επιστημονικούς και πολιτικούς φορείς, επικράτησε ο ορισμός του ΟΟΣΑ, ειδικά στο πλαίσιο της ευρωπαϊκής Πράξης για την ΤΝ, στην οποία εστίασαμε το ενδιαφέρον μας.

Καθώς υπήρξε αυξανόμενη ανάγκη για αξιοποίηση των εφαρμογών της ΤΝ τα τελευταία χρόνια, πολλές χώρες οδηγήθηκαν στη διαμόρφωση εθνικών στρατηγικών για την τεχνητή νοημοσύνη. Πλέον το παγκόσμιο τοπίο βρίθει αρχών και κατευθυντηρίων γραμμών για την ηθική τεχνητή νοημοσύνη, που προβληματίζουν ως προς το τι συνιστά “ηθική τεχνητή νοημοσύνη” και τι απαιτείται πρακτικά για την υλοποίησή της. Παρά τις διαφορές, κατατείνουν σε μια παγκόσμια σύγκλιση γύρω από καθολικές αρχές όπως διαφάνεια, δικαιοσύνη και δίκαιη μεταχείριση, έλλειψη δόλου, υπευθυνότητα και προστασία της ιδιωτικής ζωής αλλά και σε έννοιες όπως αξιόπιστη και ανθρωποκεντρική ΤΝ που χαρακτηρίζεται από την ανθρώπινη εποπτεία. Ωστόσο, παρατηρείται ότι ο “Παγκόσμιος Νότος” συχνά υποεκπροσωπείται στις ρυθμιστικές πρωτοβουλίες, γεγονός που σημειώθηκε ότι εγείρει ανησυχίες για την παγκόσμια ισορροπία στην ανάπτυξη και χρήση της ΤΝ.

Ένα ιδιαίτερα κρίσιμο ζήτημα είναι η λογοδοσία των αλγορίθμων, που έχει οδηγήσει στην υιοθέτηση εργαλείων όπως η Εκτίμηση Αλγοριθμικού Αντικτύπου, τα Μητρώα Αλγορίθμων και οι έλεγχοι αλγοριθμικής λειτουργίας (algorithm auditing). Παράλληλα, αναδεικνύεται ο ρόλος της εκάστοτε κυβέρνησης ως χρήστη της ΤΝ, καθώς η αυτοματοποιημένη λήψη αποφάσεων, όπως στην περίπτωση των προνοιακών επιδομάτων στην Ολλανδία, καταδεικνύει τις προκλήσεις που προκύπτουν. Στον αντίποδα, παρουσιάζει ιδιαίτερο ενδιαφέρον η ενίσχυση της κοινοβουλευτικής λειτουργίας μέσω της χρήσης εργαλείων ΤΝ, που μπορούν να προσφέρουν νέες δυνατότητες για τη βελτίωση της διαφάνειας, της συμμετοχικότητας και της αποτελεσματικότητας των διαδικασιών λήψης αποφάσεων κατά τη λειτουργία των κοινοβουλίων.

Ωστόσο, ελλοχεύει ο κίνδυνος η πληθώρα ηθικών αρχών να οδηγήσει σε ένα καθεστώς “αγοράς αρχών”, όπου ιδιωτικοί και δημόσιοι φορείς επιλέγουν τις αρχές που τους εξυπηρετούν περισσότερο για να δικαιολογήσουν τις τακτικές τους. Παράλληλα, η ρύθμιση της παραπληροφόρησης φαίνεται να έχει ένα παράδοξο αποτέλεσμα: όσο περισσότερο ρυθμίζεται στη Δύση, τόσο παραμελείται σε παγκόσμιο επίπεδο, γεγονός που μπορεί να έχει σοβαρές επιπτώσεις, ειδικά σε περιόδους εκλογών, όπως ήταν το 2024 σε 50 χώρες ανά την υφήλιο καθώς και στην Ευρωπαϊκή Ένωση.

Από την πλευρά μιας υψηλότερου επιπέδου διακυβέρνησης, με όρους οικουμενικότητας και παγκόσμιας συνεργασίας, διαπιστώθηκε ότι οι ανεξάντλητες ευκαιρίες που προσφέρει η TN αλλά και οι πολυεπίπεδοι κίνδυνοι που απορρέουν από τη χρήση της, έχουν οδηγήσει τόσο εθνικούς όσο και υπερεθνικούς φορείς στη θέσπιση ρυθμιστικών πλαισίων για τη βέλτιστη εφαρμογή της. Η ΕΕ επιδιώκει να αναλάβει ηγετικό ρόλο, δημιουργώντας ένα πλήρες νομοθετικό πλαίσιο με παγκόσμιο αντίκτυπο, εμπνευσμένο από τη λογική του “Brussels Effect”, ένα φαινόμενο που αναλύεται στην επικράτηση των νομοθετικών προτύπων της ΕΕ σε παγκόσμια κλίμακα, όπως συνέβη και με τον Γενικό Κανονισμό για την προστασία δεδομένων (GDPR).

Εν προκειμένω η Πράξη για την TN εισάγει ένα σύστημα ταξινόμησης βάσει κινδύνου και συνδέεται με άλλες ευρωπαϊκές ρυθμιστικές πρωτοβουλίες, ενώ ρυθμίζει την TN με γνώμονα την ατομική, κοινωνική και οικονομική ευημερία, την ασφάλεια των πολιτών της, τον σεβασμό των θεμελιωδών δικαιωμάτων, του κράτους δικαίου, της δημοκρατίας και του περιβάλλοντος, αλλά και την προώθηση της καινοτομίας. Με την Πράξη για την TN και τη διαβάθμιση των συστημάτων TN βάσει κινδύνου, η Ευρωπαϊκή Ένωση εισάγει ένα ρυθμιστικό πλαίσιο που είναι αλληλένδετο με άλλες Πράξεις της ΕΕ για τη ρύθμιση συναφών ψηφιακών πολιτικών, όπως η ρύθμιση των δεδομένων του δημόσιου τομέα αλλά και των βιομηχανικών δεδομένων, η οριοθέτηση της λειτουργίας των μεγάλων πλατφορμών, των ψηφιακών υπηρεσιών και της ψηφιακής αγοράς και τέλος, αλλά με ύψιστη σπουδαιότητα, η κυβερνοασφάλεια.

Μέρος αυτής της προσέγγισης είναι η θεσμοθέτηση παρατηρητηρίων για την TN μέσα από θεσμικά και κατευθυντήρια κείμενα της ΕΕ, καθώς εξυπηρετούν την αναγκαιότητα παρακολούθησης και συλλογής πληροφοριών για τον αντίκτυπο της χρήσης της TN με σκοπό την οικονομική ανάπτυξη, την κοινωνική και ατομική ευημερία αλλά και την αποτύπωση της επίδρασής της στην ψηφιακή δημόσια διοίκηση και στην προώθηση της διαλειτουργικότητας. Στόχο έχουν να αποτελέσουν χρήσιμο σημείο αναφοράς για όλους τους ενδιαφερόμενους φορείς, τα στελέχη του δημόσιου τομέα, τους υπεύθυνους χάραξης πολιτικής, τις ιδιωτικές εταιρείες, το ακαδημαϊκό και ερευνητικό περιβάλλον, που ενδιαφέρονται για τις τελευταίες τεχνολογικές εξελίξεις υπό το πρίσμα της βελτίωσης των λειτουργιών και της αρτιότητας της παροχής υπηρεσιών εντός του δημόσιου τομέα.

Στα ευρωπαϊκά προγραμματικά κείμενα για την τεχνητή νοημοσύνη τονίζεται η ανάγκη για αξιοπιστία και ανθρωποκεντρικότητα και διατυπώνονται γενικές αρχές για τα συστήματα TN, όπου τα θεμελιώδη δικαιώματα αναδεικνύονται σε βασικό άξονα για τον έλεγχο συμμόρφωσης με τις ευρωπαϊκές αξιακές και νομικές επιταγές. Για την επίτευξη αυτού του στόχου εισάγονται η εκτίμηση επιπτώσεων στα θεμελιώδη δικαιώματα, τεχνικά πρότυπα και δείκτες επίδοσης, που λειτουργούν ως μέτρο του ελέγχου συμμόρφωσης με τις διατυπωμένες αρχές, είτε για την αποτελεσματικότερη αξιολόγηση των επιπτώσεων στα θεμελιώδη δικαιώματα, είτε για τον εκ των υστέρων έλεγχο και παρακολούθηση.

Σε συνέχεια των παραπάνω, η Πράξη για την TN αποτελεί μια πρώτη προσπάθεια ενσωμάτωσης των θεμελιωδών δικαιωμάτων σε μια προσέγγιση ασφάλειας των προϊόντων, χρησιμοποιώντας τα εναρμονισμένα πρότυπα. Πρόκειται για μια διαδικασία η οποία βρίσκεται σε εξέλιξη μετά το αίτημα της Ευρωπαϊκής Επιτροπής προς τους αρμόδιους φορείς ευρωπαϊκών προτύπων και για ορισμένους αυτό μετατοπίζει τα θέματα ασφάλειας στον ηθικό και νομικό χώρο των δικαιωμάτων, με τον κίνδυνο να μετατοπιστεί συνακόλουθα το ζήτημα της ηθικής και νομικής κρίσης σε μια τεχνική διαδικασία. Συνεπώς είναι σημαντικό οι φορείς τυποποίησης να αναπτύξουν πρότυπα που να συμβάλλουν στην προστασία των θεμελιωδών δικαιωμάτων μέσω της διάδοσης καλών πρακτικών, αλλά να μην επεκτείνονται σε κοινωνικές, νομικές και αξιακές κρίσεις.

Όσον αφορά την εκ των υστέρων αποτίμηση της χρήσης της ΤΝ, μετά από τη δειγματοληπτική περιήγηση σε διαφορετικές πηγές καθορισμού σχετικών δεικτών επίδοσης, καταλήξαμε ότι πρόκειται για ένα ζήτημα που σχετίζεται με τον προσανατολισμό, τις προτεραιότητες και τους στόχους του φορέα που τους θέτει και δεν αποτελεί μια μονοσήμαντη ή αναπόδραστη ακολουθία βημάτων. Ειδικά κατά την εφαρμογή δημοσίων πολιτικών, η επιλογή δεικτών απεικονίζει την επιδιωκόμενη αποτελεσματικότητα σε συσχετισμό με τους διαθέσιμους πόρους και το κόστος. Το βέβαιο είναι ότι αν και γνωρίζουμε ότι η χρήση της τεχνητής νοημοσύνης προσφέρει τεράστιες ευκαιρίες για δημιουργία αξίας, δεν γνωρίζουμε πάντα πώς ακριβώς να την χρησιμοποιούμε υπεύθυνα.

Καθιερώνοντας λοιπόν δείκτες ανθρωποκεντρικούς και κοινωνικούς, συμπεριλαμβανομένων των περιβαλλοντικών και ηθικών, μετατοπίζουμε το βάρος στον επόμενο κρίσιμο τομέα που επιβάλλεται να μετρηθεί και αφορά την υπευθυνότητα, την αξιοπιστία και την ανθρωποκεντρική διάσταση της ΤΝ. Μάλιστα για καλύτερη εποπτεία αυτής της ποικιλομορφίας δεικτών, οι οποίοι παραμένουν διασκορπισμένοι σε ανομοιογενείς πηγές, δημιουργήθηκε στο πλαίσιο της παρούσας εργασίας ένας πρωτότυπος πίνακας που επιχειρεί να κωδικοποιήσει όλους τους προαναφερθέντες δείκτες αφενός ανά δύο ζεύγη κατηγοριών, οικονομικούς-τεχνολογικούς και ανθρωποκεντρικούς-κοινωνικούς, και αφετέρου ανά πηγή και με επισήμανση του υποκεφαλαίου στο οποίο έχει γίνει αναφορά τους.

Η ψηφιακή στρατηγική της Ευρωπαϊκής Ένωσης έχει αποδείξει ότι μπορεί να επηρεάσει παγκόσμια πρότυπα. Αναμένεται ότι η καθιέρωση ρητών ανθρωποκεντρικών δεικτών επίδοσης των συστημάτων τεχνητής νοημοσύνης από την ΕΕ μπορεί να επηρεάσει το παγκόσμιο τοπίο ως προς την προοπτική μιας τεχνολογίας που θα υπηρετεί την ευημερία, τα θεμελιώδη δικαιώματα, το κράτος δικαίου και τη δημοκρατία. Στο σημείο αυτό γεννιούνται νέες ανάγκες για στενότερη διεπιστημονική συνεργασία, αμοιβαία κατανόηση και άρση των αυστηρών ορίων μεταξύ του τεχνικού και του νομικού κόσμου.

Ο δημόσιος διάλογος που πυροδότησαν οι αβεβαιότητες που προκλήθηκαν από την τεχνητή νοημοσύνη είχαν ένα αποτέλεσμα για το οποίο τουλάχιστον αξίζει να της είμαστε ευγνώμονες. Έφερε και πάλι στο προσκήνιο ζητούμενα υπαρξιακά και φιλοσοφικά, σχετικά με την ανθρώπινη φύση και την αυτονομία της και επαναπροσδιόρισε τον διάλογο και την αναγκαιότητα για διεθνή συνεργασία σε ένα νεοπαγές οικοσύστημα, όπου γίνονται δυσδιάκριτα τα όρια μεταξύ ιδιωτικού και δημόσιου, αυθεντικότητας και παραποίησης, θεωρητικού και εφαρμοσμένου. Και εφόσον δεχτούμε ότι τα διλήμματα που προκαλεί η τεχνητή νοημοσύνη έδωσαν την αφορμή, αναδύονται πλέον εναργέστερα κάποια πλατύτερα ερευνητικά ερωτήματα για το μέλλον, όπως το πόσο μπορεί η ψηφιακή τεχνολογία να υποστηρίξει τη χάραξη και εφαρμογή πολιτικής, με τι όρους φέρνει τα δώρα της προόδου και της ευημερίας και κατά πόσο η καινοτομία απειλείται από τη ρύθμιση ή εάν πρόκειται για ψευδοδίλημμα, όπως υποστηρίζεται από πολλούς.

