



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΤΜΗΜΑ ΨΗΦΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ
Πρόγραμμα Μεταπτυχιακών Σπουδών
«ΔΙΚΑΙΟ ΚΑΙ ΤΕΧΝΟΛΟΓΙΕΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ»
Ακαδημαϊκό έτος 2022-2023

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Αλεξόπουλος Γεώργιος (Α.Μ.: mdi2103)

Τίτλος Εργασίας:

“Τεχνητή Νοημοσύνη στο Δημόσιο Τομέα και η Διακυβέρνησή της”

Artificial Intelligence in Public Sector and its Governance

Επιβλέπουσα:

Λίλιαν Μήτρου

Πειραιάς, Φεβρουάριος 2023

ΠΕΡΙΕΧΟΜΕΝΑ

1	Εισαγωγή.....	4
2	Επισκόπηση Βιβλιογραφίας.....	5
3	Ορισμοί	6
4	Τεχνητή Νοημοσύνη στο Δημόσιο Τομέα και Διακυβέρνηση	11
4.1	Οφέλη και Προκλήσεις	13
4.1.1	Οφέλη από τη χρήση της Τεχνητής Νοημοσύνης στο Δημόσιο Τομέα	13
4.1.2	Κίνδυνοι και Προκλήσεις από τη χρήση της Τεχνητής Νοημοσύνης.....	17
5	Διακυβέρνηση της Τεχνητής Νοημοσύνης	24
5.1	Το Πρόβλημα της Διαχείρισης της Τεχνητής Νοημοσύνης	24
5.2	Οι Επιλογές Διακυβέρνησης της Τεχνητής Νοημοσύνης	26
5.3	Προτεινόμενη Προσέγγιση για τη Διακυβέρνηση της Τεχνητής Νοημοσύνης	29
6	Εθνικές Στρατηγικές για τη Τεχνητή Νοημοσύνη	33
7	Από την Στρατηγική στη Ρύθμιση της Τεχνητής Νοημοσύνης	35
8	Ρύθμιση της Τεχνητής Νοημοσύνης στην Ελλάδα - Ν. 4961/2022	39
9	Νομικοί Προβληματισμοί από τη χρήση της Τεχνητής Νοημοσύνης.....	41
10	Συμπεράσματα.....	43
11	Επίλογος.....	45

Abstract

Artificial Intelligence (AI) has become a buzzword in recent years and its potential impact on the public sector is immense. AI technologies are increasingly being used to improve service delivery, reduce costs, and enhance decision-making. However, AI governance in the public sector is a critical issue that needs to be addressed to ensure that the technology is used responsibly and ethically. In this essay, we will examine the challenges of AI governance in the public sector and discuss the ways in which it can be effectively managed.

One of the key challenges of AI governance in the public sector is the lack of clarity around ethical and legal frameworks. While the development of AI is moving at a rapid pace, the legal and ethical frameworks that govern its use are still evolving. This lack of clarity can create a significant risk for the public sector, particularly when it comes to the use of AI in decision-making. To address this challenge, governments must work together to develop a comprehensive framework that takes into account the ethical and legal implications of AI use.

Another challenge of AI governance in the public sector is the potential for bias and discrimination. AI systems can be biased based on the data they are trained on, which can lead to discriminatory outcomes. This is particularly problematic in the public sector, where decisions made by AI systems can have a significant impact on people's lives. To address this challenge, it is essential to ensure that AI systems are developed using diverse and representative data sets, and that they are regularly audited to identify and address biases.

Transparency is another critical aspect of AI governance in the public sector. AI systems are often seen as black boxes, which can make it difficult to understand how decisions are being made. This lack of transparency can erode public trust in the technology and the organizations that use it. To address this challenge, governments and public sector organizations must be transparent about their use of AI, including how decisions are made and the data used to train AI systems.

Finally, there is a need for a skilled workforce that understands AI and its governance. The public sector must invest in training programs to equip its workforce with the necessary skills to understand and manage AI systems. This will help ensure that AI systems are used responsibly and ethically and will help to address the challenges of AI governance in the public sector.

In conclusion, AI governance in the public sector is a critical issue that needs to be addressed to ensure that the technology is used responsibly and ethically. To effectively manage AI governance, there is a need for comprehensive ethical and legal frameworks, diverse and representative data sets, transparent decision-making processes, and a skilled workforce. With these measures in place, the public sector can harness the full potential of AI while minimizing its risks.

1 Εισαγωγή

Οι τεχνολογικές εξελίξεις των τελευταίων δεκαετιών έχουν αλλάξει δραματικά την κοινωνική και καθημερινή ζωή των ανθρώπων. Τα προσωπικά και μη δεδομένα που παράγονται, και σχετίζονται με κάθε πτυχή της ζωής, συνεχώς αυξάνονται και δημιουργούν νέες συνθήκες συνοχής και διαμόρφωσης των σύγχρονων κοινωνιών. Οι αλγόριθμοι δίνουν τεράστιες δυνατότητες αξιοποίησης αυτών των δεδομένων και ανοίγουν δρόμους για την διαμόρφωση μιας κοινωνίας της τεχνητής νοημοσύνης.

Αντίστοιχα, ο πιθανός αντίκτυπός της στον δημόσιο τομέα είναι τεράστιος. Οι εν λόγω τεχνολογίες χρησιμοποιούνται όλο και περισσότερο για τη βελτίωση της παροχής υπηρεσιών, τη μείωση του κόστους και τη βελτίωση της λήψης κρίσιμων και μη αποφάσεων. Ωστόσο, η διακυβέρνηση της τεχνητής νοημοσύνης στον δημόσιο τομέα είναι ένα κρίσιμο ζήτημα που πρέπει να αντιμετωπιστεί για να διασφαλιστεί ότι η τεχνολογία χρησιμοποιείται υπεύθυνα και ηθικά.

Σε αυτή τη διπλωματική εργασία, εξετάζονται τα οφέλη και οι προκλήσεις της διακυβέρνησης της τεχνητής νοημοσύνης στον δημόσιο τομέα, ενώ αναζητούνται οι τρόποι με τους οποίους μπορεί να αντιμετωπιστεί αποτελεσματικά, τόσο το πλαίσιο ευθύνης, όσο και η απαραίτητη λογοδοσία, ως βάση ενός κράτους δικαίου. Παράλληλα, γίνεται ενδεικτική επισκόπηση της σημερινής κατάστασης όσον αφορά τις ως τώρα πρωτοβουλίες των κυβερνήσεων, τόσο σε εθνικό, όσο και σε διεθνές επίπεδο, και αφορά τις αντίστοιχες εθνικές στρατηγικές και τα ρυθμιστικά πλαίσια.

Τέλος, εξάγονται συμπεράσματα από την αντίστοιχη επισκόπηση της βιβλιογραφίας που αφορούν τη βέλτιστη δυνατή προσέγγιση διακυβέρνησης της τεχνητής νοημοσύνης ώστε να επιτευχθεί η αποκόμιση του μέγιστου οφέλους και η αποφυγή ενός δυστοπικού μέλλοντος.

2 Επισκόπηση Βιβλιογραφίας

Η παρούσα διπλωματική εργασία έχει ως στόχο μια δομημένη επισκόπηση της βιβλιογραφίας σχετικά με την χρήση της Τεχνητής Νοημοσύνης στο δημόσιο τομέα και την δημόσια διακυβέρνηση. Σκοπός είναι η κατανόηση των κύριων συστατικών της τεχνητής νοημοσύνης ώστε να γίνουν αντιληπτά, στο βαθμό του εφικτού, τα οφέλη αλλά και οι κίνδυνοι των τεχνολογικών αυτών δυνατοτήτων κατά την χρήση στο δημόσιο τομέα και διοίκηση.

Επισημαίνεται δε, πως ενώ γενικά για την τεχνητή νοημοσύνη η βιβλιογραφία είναι σχετικά εκτενής, το πεδίο εφαρμογής της στον δημόσιο τομέα και τη διακυβέρνηση δεν έχει αναλυθεί επαρκώς. Παρόλα αυτά όμως, τα τελευταία χρόνια, υπάρχει αξιοσημείωτη πρόοδος στην έρευνα του τομέα αυτού (Aoki, 2020), κυρίως λόγω των τεχνολογικών εξελίξεων αλλά και της προσοχής που δίδεται σε πολιτικό επίπεδο στον τομέα αυτό (Kuziemski & Misuraca, 2020).

Η εργασία αυτή ακολούθησε μια προσέγγιση συστηματικής βιβλιογραφικής ανασκόπησης, όπως αυτή ορίζεται από τον Kitchenham (2004), και ως εκ τούτου πραγματοποιήθηκαν οι κάτωθι ενέργειες. Πρώτον, ορισμός του θέματος και στρατηγική ανάλυσης αυτού. Δεύτερον, επιλογή μελετών προς διερεύνηση και επισκόπηση. Τρίτον, ενδελεχή επισκόπηση, μελέτη αξιολόγησης της συνάφειας αλλά και της ποιότητας του υλικού και τέλος εξαγωγή δεδομένων και συμπερασμάτων.

Αναλυτικότερα, ξεκινώντας από την έρευνα για την τεχνητή νοημοσύνη, ως γενικός όρος, διαπιστώνεται πως ο αριθμός των δημοσιευμένων εργασιών τη τελευταία εικοσαετία είναι τεράστιος και παράγεται κυρίως από τις Η.Π.Α, την Ευρώπη (αθροιστικά) και την Κίνα (Statista, 2022). Εντούτοις, ο όγκος της βιβλιογραφίας συρρικνώνεται όταν η αναζήτηση του όρου της τεχνητής νοημοσύνης συνδυάζεται στοχευμένα με το δημόσιο τομέα και διοίκηση, τη διακυβέρνηση αλλά και τις σχετικές εφαρμογές. Θέτωντας λοιπόν τα επιπρόσθετα αυτά κριτήρια αναζήτησης, κυρίως μέσω του Google Scholar, ο όγκος βιβλιογραφίας μειώνεται σε διαχειρίσιμα επίπεδα επισκόπησης.

Επιπρόσθετα, εφαρμόζοντας φίλτρα με ποιοτικές παραμέτρους στη συνάφεια των αντίστοιχων εργασιών, όπως για παράδειγμα, η τεχνητή νοημοσύνη στη δημόσια διοίκηση ή διακυβέρνηση να αποτελεί κύριο συστατικό της μελέτης προς ανάλυση, αυτή η εργασία κατέληξε στην επισκόπηση τριάντα (30) δημοσιευμένων εργασιών σε συνδυασμό με υποστηρικτικό υλικό δημοσιευμένο είτε από κυβερνητικές δομές (π.χ. Ηνωμένα Έθνη, Η.Π.Α, Ευρωπαϊκή Ένωση), ή από άλλες, ιδιωτικού ενδιαφέροντος, δομές (π.χ. Οργανισμός Οικονομικής Συνεργασίας και Ανάπτυξης [ΟΟΣΑ], Μη Κυβερνητικές Οργανώσεις [ΜΚΟ], ελεγκτικές εταιρίες, κτλ.).

Ως αποτέλεσμα της επισκόπησης του προαναφερθέντος υλικού, καταγράφησαν τα συστατικά ανάλυσης της εφαρμογής της τεχνητής νοημοσύνης στο δημόσιο τομέα. Τέλος, έχοντας ολοκληρώσει την επισκόπηση και την καταγραφή των σημείων που χρίζουν επισήμανσης και ανάλυσης, αυτή η εργασία συνθέτει τα δεδομένα αυτά μέσα από μια περιγραφική ανάλυση με σκοπό να αναδείξει τις δυνατότητες, τις προκλήσεις αλλά και τους κινδύνους από την εφαρμογή των τεχνολογιών αυτών στη δημόσια διακυβέρνηση και το δημόσιο τομέα γενικότερα.

3 Ορισμοί

Ο σκοπός της εργασίας αυτής είναι η κατανόηση της τεχνητής νοημοσύνης ειδικά στο πλαίσιο των εφαρμογών της στη δημόσια διοίκηση και την διακυβέρνηση. Ως εκ τούτου, παρουσιάζονται κάτωθι, διάφοροι ορισμοί, συναφή με τη τεχνητή νοημοσύνη, ώστε να παρέχεται μια πρακτική καθοδήγηση στην προσπάθεια εφαρμογής των πολυεπίπεδων τεχνολογιών αυτών στο δημόσιο τομέα.

Πρέπει όμως να είναι σαφές, πως η τεχνητή νοημοσύνη δεν είναι μια μονοδιάστατη επιστήμη, ώστε πιθανά προβλήματα να μπορούν να αντιμετωπιστούν άμεσα ή τεχνικά, αλλά αποτελεί ένα συνδυαστικό διάλυμα πολλαπλών επιστημών όπως της πληροφορικής, των μαθηματικών, της ψυχολογίας, της γλωσσολογίας, της νευρολογίας, της μηχανικής αλλά και της φιλοσοφίας (Παπαστάμου, 2018). Ως εκ τούτου, δεν πρέπει να αντιμετωπίζεται ως μια αφηρημένη έννοια, αλλά ως μια έννοια και επιστήμη που απαιτεί ολιστική προσέγγιση. (Geburu, 2019).

Το παραπάνω, ίσως, δικαιολογεί και το γεγονός της ύπαρξης δεκάδων διαφορετικών διατυπώσεων ορισμού της τεχνητής νοημοσύνης και για αυτό το λόγο, είναι πανθομολογούμενο, πως δεν υπάρχει ένας σαφής και μοναδικός ορισμός της τεχνητής νοημοσύνης (Lennartson 2018, Dobrev 2004, Simmons & Chapell 1988, Vinnova 2018). Γενικά έχει οριστεί με πολλαπλούς τρόπους ανάλογα με το πεδίο εφαρμογής και τον κύκλο ζωής ενός συστήματος τεχνητής νοημοσύνης, συμπεριλαμβανομένης της έρευνας, του σχεδιασμού, της ανάπτυξης και της χρήσης (UNESCO, 2020).

Παρόλα αυτά, κάτωθι παρουσιάζονται ορισμοί ώστε να γίνει πιο κατανοητή η πολυεπίπεδη αυτή έννοια. Σύμφωνα με την UNESCO (2020), και ως μια γενικευμένη αντίληψη, συστήματα τεχνητής νοημοσύνης θεωρούνται τα “τεχνολογικά συστήματα που έχουν την ικανότητα να επεξεργάζονται πληροφορίες με τρόπο που μοιάζει με ευφυή συμπεριφορά”. Δηλαδή, μια μηχανή να παρουσιάζει ενδείξεις ευφυΐας τέτοιας μορφής που είναι δύσκολο έως αδύνατο να διαπιστωθεί σε μια πιθανή δοκιμή αν είναι άνθρωπος ή μηχανή αυτός που δίνει την απάντηση. Αυτή η δοκιμή είναι ευρέως γνωστή ως Turing Test και θεωρείται η γέννηση της έννοιας της τεχνητής νοημοσύνης.

Ένας πιο ολοκληρωμένος ορισμός είναι ο κάτωθι:

“Τεχνητή Νοημοσύνη είναι ο τομέας της επιστήμης των υπολογιστών ο οποίος ασχολείται με τη σχεδίαση και υλοποίηση προγραμμάτων που μπορούν να μιμηθούν τις ανθρώπινες γνωστικές ικανότητες, εμφανίζοντας έτσι χαρακτηριστικά που συνήθως αποδίδουμε σε ανθρώπινη συμπεριφορά, όπως η μάθηση, η επίλυση προβλημάτων, η κατανόηση της φυσικής γλώσσας, η επίλυση προβλημάτων”(Βλαχάβας, Κεφαλάς κ.α., 2006).

Μιας όμως και το θέμα της εργασίας αυτής είναι η κατανόηση της προσέγγισης της εφαρμογής της τεχνητής νοημοσύνης στο δημόσιο τομέα, παρακάτω, παρατίθεται και ο ορισμός που υπάρχει στο προσφάτως δημοσιευμένο κανονισμό για τη θέσπιση εναρμονισμένων κανόνων σχετικά με την τεχνητή νοημοσύνη από την κυβερνητική δομή της Ευρωπαϊκής Ένωσης.

“Σύστημα τεχνητής νοημοσύνης σημαίνει λογισμικό που αναπτύσσεται με μια ή περισσότερες τεχνικές και προσεγγίσεις 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” (Ευρωπαϊκή Επιτροπή, 2021).

Γενικά, λόγω της ευρείας έννοιας της τεχνητής νοημοσύνης ο ορισμός της εκ των πραγμάτων πρέπει να είναι γενικός αλλά και ευπροσάρμοστος στις τεχνολογικές εξελίξεις, ενώ παράλληλα, θα μπορεί να εξειδικεύεται ανάλογα με το πεδίο εφαρμογής ώστε να περιγράφεται πιο αποτελεσματικά.

Επιπρόσθετα, στη βιβλιογραφία η τεχνητή νοημοσύνη ταξινομείται με διαφορετικούς τρόπους ώστε να μπορεί να διακρίνεται το επίπεδο ευφυΐας ενός τέτοιου συστήματος και να αποδίδει με μεγαλύτερη ακρίβεια τις δυνατότητες και τις πιθανές εφαρμογές. Παρόλα αυτά, και στο πεδίο της ταξινόμησης της τεχνητής νοημοσύνης, δεν υπάρχει ένας συνολικά αποδεκτός ορισμός από τους μελετητές. Η πιο διαδεδομένη κατηγοριοποίηση της τεχνητής νοημοσύνης είναι οι κάτωθι τρεις κύριες κατηγορίες όπως αυτές προτάθηκαν από τους Gurkaynak, Yilmaz & Haksever.

Η πρώτη κατηγορία είναι η στενή (αδύναμη) τεχνητή νοημοσύνη (Artificial Narrow Intelligence - ANI) και αναφέρεται στα συστήματα τα οποία έχουν σχεδιαστεί να επιλύουν μια συγκεκριμένη εργασία σε ένα οριοθετημένο περιβάλλον. Για παράδειγμα, οι φωνητικοί βοηθοί Siri και Alexa από τις εταιρίες Apple και Amazon αντίστοιχα, οι μηχανές αναζήτησης, οι στοχευμένες διαφημίσεις μέσα από τα κοινωνικά δίκτυα, ή ακόμα και ο υπολογιστής σκακιού Deep Blue της IBM που κέρδισε τον παγκόσμιο πρωταθλητή Κασπάροφ, αποτελούν τέτοια παραδείγματα.

Κοινό στοιχείο των προαναφερθέντων περιπτώσεων είναι πως τα εν λόγω συστήματα δεν μπορούν να εκκινήσουν ή να επιτελέσουν άλλη λειτουργία πέρα από το προκαθορισμένο περιβάλλον τους. Σε αυτή τη κατηγορία δηλαδή, η τεχνητή νοημοσύνη ξεπερνά, κατά πολύ, σε δυνατότητες τους ειδικούς στο συγκεκριμένο πεδίο, αλλά μόνο σε αυτό (Lennartson, 2018).

Η δεύτερη κατηγορία είναι η γενική (ή ισχυρή) τεχνητή νοημοσύνη (Artificial General Intelligence – AGI) η οποία αναφέρεται στη δυνατότητα ενός συστήματος να υποστηρίξει πολλαπλές εργασίες παράλληλα, και ουσιαστικά, να εκτελεί αυτόνομα πνευματικές εργασίες, όπως ένας άνθρωπος. Σε αυτή τη κατηγορία, ένα σύστημα τεχνητής νοημοσύνης μπορεί να ξεκινήσει να επιλύει ένα ή περισσότερα πρόβληματα, όπως σχεδιάστηκε, όμως είναι ικανό να εκπαιδευτεί μόνο του και να εφαρμόσει τις αποκομισθείσες, κατά την λειτουργία του, δεξιότητες, σε νέες εργασίες χωρίς ανθρώπινη παρέμβαση (Adams et al., 2012). Ουσιαστικά, σε αυτή τη κατηγορία το σύστημα αποδεικνύει μια δυνατότητα να κατανοήσει ένα μη προκαθορισμένο πλαίσιο και να αντιμετωπίσει αποτελεσματικά την τυχαιότητα του πραγματικού κόσμου.

Η τρίτη κατηγορία είναι η τεχνητή υπερνοημοσύνη (Artificial Super Intelligence – ASI) και αναφέρεται στα συστήματα τα οποία είναι ευφυέστερα του ανθρώπου σε κάθε γνωστικό πεδίο συμπεριλαμβανομένων και των “ήπιων” δεξιοτήτων (soft skills), όπως η δημιουργικότητα. Σε αυτή τη περίπτωση, η τεχνητή υπερνοημοσύνη θα ξεπεράσει τη συλλογική νοημοσύνη της ανθρωπότητας (Bostrom 2014, Kaplan 2015, Thierer 2017).

Στις μέρες μας κυριαρχούν οι εφαρμογές της στενής / αδύναμης τεχνητής νοημοσύνης (ANI), ενώ προβλέπεται πως στο άμεσο μέλλον, και για τις επόμενες δεκαετίες, αυτά τα συστήματα θα αποτελούν το μεγαλύτερο κομμάτι εφαρμογής στην ψηφιακή μας κοινωνία. Μάλιστα εκτιμάται ότι η συγκεκριμένη αγορά της στενής / αδύναμης τεχνητής νοημοσύνης που αποτιμήθηκε από τους ειδικούς το 2015 στα 130 δισεκατομμύρια δολάρια αναμένεται να εκτοξευθεί πάνω από τα 3 τρισεκατομμύρια δολάρια μέχρι το 2024 (Market Transparency Research, 2016).

Παρόλα αυτά όμως, δίστανται οι απόψεις των ειδικών για το πότε η γενική (ή ισχυρή) τεχνητή νοημοσύνη (AGI) θα γίνει πραγματικότητα. Οι πιο μετριοπαθείς, όπως ο Bostrom, προβλέπουν πως κάτι τέτοιο ίσως είναι εφικτό μέχρι το 2050, ενώ πιθανολογούν πως η πιο ασφαλής πρόβλεψη είναι περί το 2075.

Επιπροσθέτως, ο Bostrom (2006) και άλλοι ερευνητές (π.χ. Urban, 2015), προβλέπουν πως η γενική (ή ισχυρή) τεχνητή νοημοσύνη (AGI) θα εξελιχθεί πολύ γρήγορα σε τεχνητή υπερνοημοσύνη (ASI) χαρακτηρίζοντας το σημείο αυτό ως έκρηξη νοημοσύνης (intelligence explosion) ή μοναδικότητα (singularity). Μάλιστα σε αυτό το χρονικό σημείο

θα είναι αρκετά δύσκολο για τη πλειοψηφία των ανθρώπων να κατανοήσουν τον ακριβή αντίκτυπο της τεχνητής νοημοσύνης, αλλά και την εκθετική ανάπτυξη αυτού του είδους τεχνητής νοημοσύνης από εκεί και πέρα (Gurkaynak, Yilmaz & Haksever, 2016).

Είναι προφανές, από τα παραπάνω, πως η συνεχώς εξελισσόμενη τεχνητή νοημοσύνη αναμένεται να διαδραματίσει καταλυτικό ρόλο παγκοσμίως τα επόμενα χρόνια γενικώς και ειδικότερα στην διακυβέρνηση, και ως αποτέλεσμα, οι κυβερνήσεις θα πρέπει να κατανοήσουν σε βάθος αυτές τις τεχνολογικές εξελίξεις ώστε να μπορούν να επιτελέσουν πρακτικά την αποστολή τους.

Πριν όμως εξετάσουμε την επίδραση της τεχνητής νοημοσύνης στο δημόσιο τομέα και τη διακυβέρνηση, είναι ορθό να εξετάσουμε πέρα από τους παραπάνω ορισμούς της, και τα κύρια συστατικά της τεχνολογίας αυτής. Αυτό θα βοηθήσει ώστε να γίνει καλύτερα αντιληπτός ο πυρήνας των εφαρμογών της τεχνητής νοημοσύνης.

Αναλυτικότερα, το κύριο συστατικό για τη δημιουργία ενός συστήματος τεχνητής νοημοσύνης, είναι η μηχανική μάθηση (Machine Learning), η οποία μέσω πολύπλοκων μαθηματικών υπολογισμών και αλγορίθμων δίνει τη δυνατότητα σε ένα σύστημα, να αποκτήσει δική του γνώση εξάγοντας μοτίβα, πρότυπα και συσχετίσεις από δεδομένα που συλλέγει κατά την αλληλεπίδραση με το περιβάλλον που δρα (Goodfellow, Bengio and Courville 2016). Δηλαδή, η διαδικασία με την οποία ένας υπολογιστής μπορεί να παραμετροποιήσει και να βελτιώσει έναν αλγόριθμο ή ένα μοντέλο χωρίς ανθρώπινη συμμετοχή συνολικά ή και σε ένα μέρος αυτής. Γενικά, τα περισσότερα συστήματα τεχνητής νοημοσύνης βοηθούν στην ανάλυση μεγάλου όγκου δεδομένων (Big Data) που ο άνθρωπος δεν μπορεί να διαχειριστεί αποκλειστικά μόνος του. Μια υποκατηγορία αυτής, είναι η βαθιά μάθηση (Deep Learning), όπου η ανάλυση των δεδομένων και η εξόρυξη μοτίβων, γίνεται σε πολλαπλά επίπεδα, και όχι μόνο στα ευκόλως ορατά από την ανθρώπινη αντίληψη. Αυτό επιτυγχάνεται σε πολλές περιπτώσεις με την υιοθέτηση συγκεκριμένων μεθόδων και αλγορίθμων, γνωστά ως νευρωνικά δίκτυα ([Artificial] Neural Networks- [A]NN) τα οποία αποτελούνται από πολλούς νευρώνες - κόμβους που ανταλλάσσουν και επεξεργάζονται πληθώρα πληροφοριών.

Τέλος, μιας και το αντικείμενο αυτής της εργασίας είναι η εφαρμογή της τεχνητής νοημοσύνης στο δημόσιο τομέα και διακυβέρνηση, θα πρέπει, κυρίως για λόγους πληρότητας, να οριστεί η έννοια της κυβέρνησης και της δημόσιας διακυβέρνησης.

Κυβέρνηση για τους σκοπούς της εργασίας αυτής, ορίζεται ως η νόμιμη εξουσία με κάποια γεωγραφική δικαιοδοσία, είτε εθνική (π.χ. Ελλάδα), ή υπερεθνική (π.χ., Ευρωπαϊκή Ένωση) (Cihon et. al., 2021). Η διακυβέρνηση, σύμφωνα με τον Fukuyama (2013), και όσον αφορά το δημόσιο τομέα, είναι η ικανότητα ενός οργανισμού να θεσπίζει και να επιβάλλει κανόνες λειτουργίας μιας κοινωνίας αλλά και να προσφέρει υπηρεσίες. Ως εκ τούτου, οι κυβερνήσεις στρατηγικά εκμεταλλεύονται και προσπαθούν να εφαρμόσουν τις αναδυόμενες τεχνολογίες για να διασφαλίσουν από την μια, την επιβολή των κανόνων και την εύρυθμη λειτουργία του κράτους δικαίου, και από την άλλη, την αποτελεσματικότερη προσφορά υπηρεσιών στους πολίτες. Δηλαδή, οι κυβερνήσεις προσπαθούν μέσω της καινοτομίας να προασπίσουν το δημόσιο συμφέρον αλλά και να δημιουργήσουν αξία που εξελίσσει θετικά την κοινωνία.

4 Τεχνητή Νοημοσύνη στο Δημόσιο Τομέα και Διακυβέρνηση

Τις τελευταίες δεκαετίες, διαφορετικές τεχνολογίες έχουν οδηγήσει σε πολλαπλές στρατηγικές δημιουργίας αξίας στο δημόσιο τομέα. Από τις αρχές του 21^{ου} αιώνα, οι τεχνολογίες πληροφορικής και επικοινωνιών, και κυρίως το διαδίκτυο, έχουν προσελκύσει όλο και περισσότερο το ενδιαφέρον του επιστημονικού κλάδου και των κυβερνήσεων ώστε να εξετάσουν τις πρακτικές εφαρμογές των τεχνολογιών στη παραγωγή δημόσιας αξίας.

Αρχικά, η μηχανογράφηση του δημοσίου τομέα και στη συνέχεια η διάδοση του διαδικτύου, οδήγησαν, και συνεχίζουν να συνεισφέρουν, στο τεχνολογικό μετασχηματισμό της κοινωνίας. Η ψηφιοποίηση των διαδικασιών του δημοσίου τομέα, η διαλειτουργικότητα μεταξύ των δημόσιων οργανισμών, αλλά και η ύπαρξη μεγάλου όγκου δεδομένων, καθιστούν πλέον δυνατή τη ψηφιακή διακυβέρνηση και την αποκόμιση δημοσίου οφέλους από τη χρήση των νέων τεχνολογιών.

Όμως, η παραγωγή δημόσιας αξίας προϋποθέτει την εις βάθος κατανόηση της όποιας τεχνολογίας από τις κυβερνήσεις, αλλά και την ανάλυση των πιθανών κοινωνικοπολιτικών επιπτώσεων, ιδιαίτερα κατά το νέο τεχνολογικό κύμα όπου η τεχνητή νοημοσύνη είναι στο επίκεντρο.

Σημειώνεται δε, πως οι δημόσιοι οργανισμοί διαφέρουν κατά πολύ από τον ιδιωτικό τομέα, καθώς δρουν σε ένα πιο πολύπλοκο περιβάλλον με εξαιρετικά διαφοροποιημένα πολιτικά, πολιτιστικά και κοινωνικοοικονομικά οικοσυστήματα να αλληλεπιδρούν ανά περίπτωση (Misuraca et. al., 2020). Οι θεμελιώδεις αρχές, οι κοινωνικές αξίες, οι περίπλοκες σχέσεις του δημοσίου τομέα με άλλους φορείς, δημόσιους ή ιδιωτικούς, αλλά κυρίως, η συνεχής αλληλεπίδραση με τους πολίτες, καθιστούν την υιοθέτηση μιας καινοτόμου τεχνολογίας αρκετά απαιτητική και χρονοβόρα. Πρακτικές εφαρμογές της τεχνητής νοημοσύνης δεν μπορούν απλά να αντιγραφούν από το ιδιωτικό τομέα και να αξιοποιηθούν αυτούσιες στο δημόσιο και τη διακυβέρνηση, κυρίως λόγω του ότι η παραγωγή δημόσιας αξίας είναι ο πρωταρχικός στόχος και όχι το άμεσο οικονομικό όφελος (Desouza et al., 2020).

Επιπρόσθετα, τα συστήματα τεχνητής νοημοσύνης γίνονται ολοένα και πιο πολύπλοκα, ενώ παράλληλα λιγότερο προβλέψιμα, γεγονός που καθιστά σχεδόν αδύνατη, προς το παρόν, τη πρόβλεψη για το πως επηρεάζεται η δημόσια διακυβέρνηση και η εύρυθμη λειτουργία του δημοσίου τομέα. Μάλιστα τα τελευταία χρόνια, ο εκθετικά αυξανόμενος ρυθμός εφαρμογής της τεχνητής νοημοσύνης στο δημόσιο τομέα είναι δυσανάλογος της ικανότητας διαχείρισης των κυβερνήσεων σε παγκόσμιο επίπεδο (IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, 2019). Το γεγονός αυτό ενδέχεται να μην επιφέρει τα προσδοκώμενα οφέλη και να επηρεάσει αρνητικά τη σχέση εμπιστοσύνης δημοσίου τομέα και πολιτών. Για το λόγο αυτό, επισημαίνεται εντόνως από τους ερευνητές το κενό γνώσης στη ρύθμιση και διακυβέρνηση της τεχνητής νοημοσύνης στο δημόσιο τομέα και η αναγκαιότητα αντιμετώπισης του θέματος ολιστικά και όχι απλά τεχνικά όπως συμβαίνει μέχρι σήμερα (Wang & Siau, 2018; Wirtz et al., 2020). Πόσο δε μάλλον, όταν οι δημόσιοι οργανισμοί και οι υπηρεσίες τους απευθύνονται στις πιο ευάλωτες κοινωνικές ομάδες οι οποίες αποτελούν σημαντικό, αν όχι το μεγαλύτερο, ποσοστό πληθυσμού.

Πρέπει να σημειωθεί πως αναφερόμαστε κυρίως στη στενή ή αδύναμη τεχνητή νοημοσύνη (ANI) μιας και σύμφωνα με τους Thierer et al. (2017), το μεγαλύτερο μέρος των εφαρμογών της, ειδικότερα στο δημόσιο τομέα, πιθανόν θα παραμείνει αδύναμη και εξαιρετικά εξειδικευμένη για τις επόμενες αρκετές δεκαετίες.

Τελευταία βέβαια, οι κυβερνητικές δομές παγκοσμίως αρχίζουν να εστιάζουν στη χάραξη πολιτικής και στρατηγικών για την εφαρμογή της τεχνητής νοημοσύνης και μελετάται η διακυβέρνηση και τα ρυθμιστικά ζητήματα της τεχνολογίας αυτής, όμως η συστηματική επισκόπηση των αποτελεσμάτων χρήσης της, βρίσκεται ακόμη σε πρώιμο στάδιο. Παρόλα αυτά, και παρά τη μερική κατανόηση του αντικειμένου και των πιθανών επιπτώσεων, ήδη δημόσιοι οργανισμοί και κυβερνήσεις προσπαθούν να εφαρμόσουν τεχνολογίες τεχνητής νοημοσύνης στις υπηρεσίες τους με σκοπό να παράγουν άμεσα δημόσια αξία, είτε αυτή είναι λειτουργική, ή πολιτική, ή ακόμα και κοινωνική (Manzoni et. al., 2022). Όμως τα οφέλη και η παραγωγή δημόσιας αξίας μέσω της χρήσης της τεχνητής νοημοσύνης απαιτούν πολυεπίπεδη ανάλυση των στόχων αλλά και των τεχνικών και ρυθμιστικών προκλήσεων. Οι προκλήσεις και τα οφέλη από τη χρήση της τεχνητής νοημοσύνης αναλύονται παρακάτω.

4.1 Οφέλη και Προκλήσεις

4.1.1 Οφέλη από τη χρήση της Τεχνητής Νοημοσύνης στο Δημόσιο Τομέα

Όπως αναφέρθηκε και παραπάνω, στόχος των κυβερνήσεων είναι να παράξουν δημόσια αξία από τη χρήση της τεχνητής νοημοσύνης στο δημόσιο τομέα. Η προσδοκώμενη δημόσια αξία μπορεί να είναι λειτουργική, δηλαδή αύξηση της αποτελεσματικότητας παροχής υπηρεσιών, η αποδοτικότητα των εργαζομένων, ή ακόμα και καθοδήγηση των πολιτών ώστε να επωφεληθούν από τις δημόσιες υπηρεσίες. Επίσης, η επιθυμητή δημόσια αξία μπορεί να είναι πολιτική, δηλαδή, η ενίσχυση της ισότητας, της δικαιοσύνης, της προσβασιμότητας, της διαφάνειας, της λογοδοσίας αλλά και της συμμετοχής των πολιτών στη δημόσια ζωή μέσα από εφαρμογές της τεχνητής νοημοσύνης.

Τέλος, συστήματα τεχνητής νοημοσύνης μπορούν να παράξουν και κοινωνική δημόσια αξία, διευκολύνοντας την ένταξη των ευάλωτων ομάδων στη σύγχρονη ψηφιακή κοινωνία, προωθώντας την εμπιστοσύνη, την ποιότητα ζωής αλλά και τη βιωσιμότητα της ανάπτυξης της κοινωνίας (Barker et al., 2021).

Γενικά, τα οφέλη από την χρήση της τεχνητής νοημοσύνης στο δημόσιο τομέα μπορούν να κατηγοριοποιηθούν περαιτέρω, όχι μόνο με βάση τη φύση της προσδοκώμενης δημόσιας αξίας αλλά και ειδικότερα με το τύπο εφαρμογής αυτής. Αναλυτικότερα και σύμφωνα με τη πρόσφατη βιβλιογραφική έρευνα των Zuiderwijk et. al. (2021) , τα οφέλη αποδοτικότητας της τεχνητής νοημοσύνης αναφέρονται στη ενίσχυση της αποτελεσματικότητας της παροχής δημόσιων υπηρεσιών αυτοματοποιώντας ή απλοποιώντας διαδικασίες και επαναλαμβανόμενες εργασίες με τη βοήθεια μηχανικής μάθησης, απελευθερώνοντας έτσι πολύτιμους επιχειρησιακούς πόρους (Alexopoulos et al., 2019). Επιπλέον, τα συστήματα ψηφιακής διακυβέρνησης με τη χρήση της τεχνητής νοημοσύνης αποφέρουν και οικονομικά οφέλη. Μέσω της πιο στοχευμένης παροχής υπηρεσιών και της απλοποίησης των διαδικασιών, αυτά τα “έξυπνα” συστήματα ηλεκτρονικής διακυβέρνησης μειώνουν τα λειτουργικά κόστη, αλλά και υποκαθιστούν το εργατικό δυναμικό του δημοσίου τομέα σε σωρεία υπηρεσιών. Για παράδειγμα, ήδη στις σκανδιναβικές χώρες, και συγκεκριμένα στο δήμο Umeå (Lennartson, 2018), χρησιμοποιείται σύστημα τεχνητής νοημοσύνης Chatbot AI για την άμεση εξυπηρέτηση των πολιτών το οποίο μπορεί να επεξεργάζεται τον προφορικό λόγο και να καθοδηγεί τους πολίτες στη λήψη δημοσίων υπηρεσιών (π.χ. απάντησεις σε συχνές ερωτήσεις, προγραμματισμός συναντήσεων, δρομολόγηση αιτημάτων στην κατάλληλη δημόσια υπηρεσία, συμπλήρωση εγγράφων, κτλ.). Αντίστοιχο σύστημα Chatbot AI, χρησιμοποιείται ήδη και στη Φιλανδία για υποστήριξη και καθοδήγηση των πολιτών στην εκπλήρωση των φορολογικών τους υποχρεώσεων (Susar & Aquaro, 2019). Ως αποτέλεσμα, παρατηρείται σημαντική αύξηση της αποτελεσματικότητας στις παρεχόμενες δημόσιες υπηρεσίες ενώ παράλληλα, σημειώνεται σημαντική μείωση του διοικητικού και λειτουργικού κόστους του δημοσίου, ενώ εμμέσως αντιμετωπίζεται και η έλλειψη εργατικού δυναμικού.

Επιπρόσθετα, με τη χρήση μηχανικής μάθησης, ο δημόσιος τομέας δύναται να βελτιώσει την ικανότητα του στη λήψη αποφάσεων. Πιο συγκεκριμένα, οι κυβερνήσεις για πρώτη φορά στην ιστορία έχουν πρόσβαση σε τόσο τεράστιο όγκο δεδομένων των πολιτών και συνάμα ευφυή συστήματα υποστήριξης λήψης αποφάσεων με ακριβή αποτελέσματα. Με τη βοήθεια αυτών των συστημάτων μπορούν να επεξεργαστούν τα δεδομένα αυτά και να εξάγουν μοτίβα, να προβλέψουν κινδύνους και να παρέχουν στοχευμένες υπηρεσίες σε ομάδες ή και στο σύνολο των πολιτών, είτε σε πραγματικό χρόνο, ή σε πολύ σύντομο χρονικό διάστημα. Ένα τέτοιο παράδειγμα είναι και η υιοθέτηση συστημάτων τεχνητής νοημοσύνης για τη διαχείριση της πανδημίας COVID-19 από αρκετές κυβερνήσεις, ώστε να παρακολουθούν την εξέλιξη αυτής και να λαμβάνουν στοχευμένες αποφάσεις για την αντιμετώπιση της εξάπλωσής της. Αυτή η ικανότητα του δημοσίου τομέα να παρέχει άμεσα υπηρεσίες, στοχευμένα και προσωπικά σε κάθε πολίτη, πέρα της αποτελεσματικότητας, μειώνει παράλληλα και το κόστος και το χρόνο εξυπηρέτησης της παροχής δημοσίων υπηρεσιών, ενώ παράλληλα, βελτιώνει τη ποιότητα ζωής των πολιτών. Η αποτελεσματικότερη αλληλεπίδραση των πολιτών με το δημόσιο τομέα, αναπόφευκτα δύναται να ενισχύσει την εμπιστοσύνη των πολιτών και τη ποιότητα της διακυβέρνησης.

Ακόμη, η δημόσια υγεία μπορεί να προσδοκά οφέλη από τη χρήση της τεχνητής νοημοσύνης. Ήδη, οι κυβερνήσεις εξετάζουν την εφαρμογή καινοτόμων συστημάτων στις δημόσιες δομές υγείας για την αποτελεσματικότερη περίθαλψη των νοσηλευόμενων, ακόμη και απομακρυσμένα. Έχοντας τη δυνατότητα ακριβούς εποπτείας ζωτικών ενδείξεων (καρδιακοί παλμοί, επίπεδα οξυγόνου, αρτηριακή πίεση, κτλ.), τα συστήματα τεχνητής νοημοσύνης μπορούν να συνεισφέρουν στην άμεση ανταπόκριση του υγεινομικού μηχανισμού, συμβάλλοντας σημαντικά, στη λήψη κρίσιμων αποφάσεων (ενημέρωση για άμεση παρέμβαση ιατρού) ή και ακόμα προχωρώντας σε ενέργειες (χρήση ρομποτικής, χορήγηση φαρμάκου, κτλ.) ή συστάσεις προς τον νοσηλευόμενο (παρότρυνση για ακινησία, κτλ.). Εκτιμάται δε, πως μόνο η χρήση της τεχνητής νοημοσύνης σε συνδυασμό με τη ρομποτική στο κλαδο της υγείας, θα εκτοξευθεί από τα είκοσι (20) δισεκατομμύρια το 2022, στα διακόσια δέκα (210) δισεκατομμύρια εντός της δεκαετίας που διανύουμε (Grand View Research, 2023).

Όμως, τα οφέλη όμως στη διακυβέρνηση και το δημόσιο τομέα λόγω της χρήσης της τεχνητής νοημοσύνης δεν περιορίζονται στη παροχή άμεσων υπηρεσιών προς τους πολίτες. Ο δημόσιος τομέας διαχειρίζεται τεράστιο όγκο δεδομένων φυσικών και άλλων πόρων. Αυτό σημαίνει πως ο συνεχώς εξελισσόμενος ψηφιακός μετασχηματισμός και η αυτοματοποίηση διαδικασιών δημιουργούν πρόσφορο έδαφος για τις κυβερνήσεις να εστιάσουν αποτελεσματικότερα τη διαχείριση των φυσικών και όχι μόνο πόρων (π.χ. ενέργεια, αποθέματα νερού, υγειονομικό υλικό και φάρμακα κτλ.). Ως αποτέλεσμα, οι κυβερνήσεις με τη χρήση τεχνητής νοημοσύνης μπορούν να αποτρέψουν τη σπατάλη πόρων, προωθώντας παράλληλα ένα σταθερά βιώσιμο περιβάλλον και την αειφόρο ανάπτυξη (Fatima et al., 2020).

Τέλος και όσον αφορά τα οφέλη, οι κυβερνήσεις πέρα της αποτελεσματικότερης οργανωτικής και οικονομικής διαχείρισης του δημοσίου τομέα, εξετάζουν και προσπαθούν να εφαρμόσουν τη τεχνητή νοημοσύνη και στο επίπεδο της εθνικής ασφάλειας και άμυνας ώστε να προασπίσουν τις εθνικές στρατηγικές τους, να διασφαλίσουν τη γεωπολιτική τους επιρροή και το όποιο στρατιωτικό πλεονέκτημα κατέχουν ή επιθυμούν. Ως εκ τούτου, τεράστια ποσά επενδύονται σε αυτό το τομέα κυρίως από τις ηγέτιδες χώρες στην αμυντική βιομηχανία (π.χ. Η.Π.Α, Κίνα, Ρωσία, Γαλλία) για την παραγωγή αυτόνομων οπλικών συστημάτων, drones αλλά και για την ενίσχυση των υπηρεσιών συλλογής πληροφοριών (π.χ. μυστικές υπηρεσίες, κτλ.).

Είναι γεγονός πως οι εφαρμογές της τεχνητής νοημοσύνης αφορούν το σύνολο της σύγχρονης ψηφιακής κοινωνίας και παραπάνω παρατέθηκαν τα κυριότερα οφέλη που μπορεί να αποκομίσουν οι κυβερνήσεις με τη χρήση των τεχνολογιών αυτών στο δημόσιο τομέα. Πρέπει όμως να σημειωθεί το πολυσχιδές περιβάλλον δράσης του δημοσίου τομέα σε όλο το κοινωνικό φάσμα. Άρα, εμμέσως και πέρα των παραπάνω, η χρήση της τεχνητής νοημοσύνης στο δημόσιο τομέα, θα μπορούσε να επεκταθεί σε όλη το φάσμα δράσης της κοινωνίας, είτε αυτό αφορά την εποπτεία της βιομηχανίας, της ναυτιλίας και το σύνολο της εφοδιαστικής αλυσίδας, καθώς ο ρόλος του δημοσίου δεν είναι άμεσος αλλά αναπόφευκτα δραστηριοποιείται παντού (π.χ. τελωνεία, δασμοί, μεταφορές, διασφάλιση ομαλής ροής τροφίμων, υγειονομικού υλικού, κτλ.).

4.1.2 Κίνδυνοι και Προκλήσεις από τη χρήση της Τεχνητής Νοημοσύνης

Πέραν όμως των πιθανών θετικών αποτελεσμάτων από τη χρήση της τεχνητής νοημοσύνης, οι προκλήσεις που προκύπτουν και καλείται ο δημόσιος τομέας, και κατ' επέκταση οι κυβερνήσεις να αντιμετωπίσουν, είναι αρκετές και δυσεπίλιντες. Οι προκλήσεις και οι κίνδυνοι προκύπτουν, είτε από τη πολυπαραγοντική φύση της τεχνητής νοημοσύνης, αλλά σε αρκετές περιπτώσεις λόγω του ιδιαίτερου περιβάλλοντος δράσης του δημοσίου τομέα και της διακυβέρνησης. Πρακτικά αυτό σημαίνει πως προκύπτουν προβληματισμοί από την ίδια την τεχνολογία ανεξαρτήτου πεδίου εφαρμογής, ενώ αυτοί διογκώνονται ποσοτικά και ποιοτικά όταν προσπαθούν να εφαρμοστούν συγκεκριμένα στο δημόσιο τομέα. Γενικά, και όπως αναφέρθηκε παραπάνω, τα πολλαπλά διανύσματα σύστασης της τεχνολογίας αυτής, σε συνδυασμό με το πολυπαραγοντικό πεδίο δράσης του δημοσίου, δημιουργούν πολυεπίπεδες προκλήσεις που δεν έχουν αναλυθεί επαρκώς.

Όσον αφορά τους κινδύνους και τις προκλήσεις της εφαρμογής της τεχνητής νοημοσύνης, μπορούν να κατηγοριοποιηθούν αναλόγως, κυρίως για πρακτικούς λόγους ανάγνωσης και πληρότητας. Σύμφωνα με την λεπτομερή βιβλιογραφική επισκόπηση των Zuiderwijk et. al. (2021), οι προκλήσεις μπορούν να ομαδοποιηθούν σε οκτώ συναφείς κατηγορίες οι οποίες είναι: i) οι προκλήσεις που σχετίζονται με τα δεδομένα, ii) οργανωτικές και διαχειριστικές προκλήσεις, iii) προκλήσεις που αφορούν τις δεξιότητες των ανθρώπων που καλούνται να διαχειριστούν συστήματα τεχνητής νοημοσύνης, iv) προκλήσεις σχετικές με την ερμηνεία των αποτελεσμάτων χρήσης της τεχνολογίας αυτής, v) προκλήσεις σχετικά με τη νομιμότητα αλλά και κατ' επέκταση με την ηθική διάσταση της χρήσης της, vi) πολιτικές, νομικές και πολιτιστικές προκλήσεις, vii) κοινωνικές και τέλος viii) οικονομικές προκλήσεις.

Αναλυτικότερα, ο μεγάλος όγκος δεδομένων είναι ένα από τα κύρια συστατικά της τεχνητής νοημοσύνης για να μπορεί να προσφέρει το επιθυμητό αποτέλεσμα. Όμως, η διαθεσιμότητα των δεδομένων και ο τρόπος κτήσης αυτών εγείρουν προβληματισμούς. Τα διαθέσιμα δεδομένα δεν αντιπροσωπεύουν απαραίτητα το σύνολο του πληθυσμού, ενώ από ποιοτική σκοπιά (Toll et al., 2019), ίσως να μην αντανακλούν τη πραγματικότητα σε αρκετές περιπτώσεις. Παρά την διασύνδεση των οργανισμών του δημοσίου τομέα, τα

δεδομένα τους παραμένουν αδόμητα, ανομοιογενή ενώ η ενοποίησή τους, με επιβεβαίωση της ποιότητάς τους, είναι αρκετά δύσκολη (Gupta, 2019). Επιπρόσθετα, σημειώνεται πως τα δεδομένα που (πολύ) πιθανόν να βασίζεται ένα σύστημα τεχνητής νοημοσύνης που εφαρμόζεται στο δημόσιο τομέα, δεν προέρχονται αποκλειστικά από το δημόσιο ή τη κυβέρνηση, αλλά και από άλλες πηγές και ο τρόπος κτήσης αυτών είναι αμφισβητούμενος. Αν επιπροσθέτως ληφθεί υπόψιν και πως τα συλλεχθέντα δεδομένα πιθανόν να αντιπροσωπεύουν απόψεις ή προκαταλήψεις, αντί για ρεαλιστική απεικόνιση, τότε, η απαραίτητη ποιότητα και ακρίβεια των δεδομένων είναι αίολη και τα αποτελέσματα της χρήσης της τεχνητής νοημοσύνης τελικώς να επιφέρουν περισσότερα προβλήματα από όσα προσπαθούν να επιλύσουν.

Επίσης και όσον αφορά τα δεδομένα, ήδη από το σχεδιασμό, επιλέγεται, εσκεμμένα ή μη, ένα σύνολο δεδομένων εκπαίδευσης του αλγορίθμου μηχανικής μάθησης που ίσως να μην καλύπτει επαρκώς τους σκοπούς χρήσης μιας εφαρμογής τεχνητής νοημοσύνης. Ως πιθανό αποτέλεσμα λοιπόν, ο αλγόριθμος θα συνεχίσει να ανατροφοδοτείται με ανακριβή στοιχεία κατά τη λειτουργία του και τελικώς να αποτύχει το αντίστοιχο εγχείριμα.

Από τα παραπάνω είναι μάλλον δεκτό να συμμεριστεί κάποιος την άποψη του Thierer (2017), πως τα “μεγάλα” δεδομένα (Big Data) και η ποιότητα αυτών, είναι θεμέλιος λίθος επιτυχίας εφαρμογής ενός συστήματος τεχνητής νοημοσύνης.

Δεύτερον, η όποια καινοτόμα τεχνολογία δεν δύναται να επιτύχει αν δεν συμβαδίζει με αντίστοιχες οργανωτικές και διαχειριστικές παρεμβάσεις. Είτε στο δημόσιο, ή ακόμα και στον ιδιωτικό τομέα, η αντίσταση στην αλλαγή και την αφομίωση ενός νέου τρόπου λειτουργίας είναι δεδομένη, αφού αυτή η αντίσταση πηγάζει από την ανθρώπινη φύση (Coch, 1948). Αυτό πρακτικά σημαίνει πως μια δυνητικά ανατρεπτική τεχνολογία όπως η τεχνητή νοημοσύνη, αντιμετωπίζεται με καχυποψία και η επιτυχή εφαρμογή της εξαρτάται από τη προετοιμασία σε επίπεδο οργάνωσης. Ως εκ τούτου, παράλληλα με το σχεδιασμό ενός συστήματος τεχνητής νοημοσύνης θα πρέπει να οργανώνεται ο μετασχηματισμός των διαδικασιών και η λειτουργία στο νέο πλαίσιο. Αυτό όμως απαιτεί, ισχυρή ηγεσία, καταρτισμό στρατηγικής, αποτελεσματική επικοινωνία του αντίστοιχου οράματος και παρεμβάσεις στο πυρήνα της κουλτούρας ενός οργανισμού, ειδικότερα στο

δημόσιο τομέα, όπου η τεχνητή νοημοσύνη αμφισβητεί την παραδοσιακά γραφειοκρατική μορφή διακυβέρνησης (Bullock, 2019). Επισημαίνεται δε, πως η εκθετική ανάπτυξη της τεχνητής νοημοσύνης είναι δυσανάλογη με τη δυνατότητα των οργανισμών να προσαρμοστούν στις αλλαγές, ενώ είναι παθολογούμενη, η ανεπαρκής ικανότητα διαχείρισης και διακυβέρνησής της.

Τρίτον, μια ακόμη πρόκληση είναι η έλλειψη τεχνογνωσίας και δεξιοτήτων σχετικές με τη τεχνητή νοημοσύνη (Wirtz et al., 2019). Για να διαχειριστεί επαρκώς ένας οργανισμός αυτοματοποιημένα συστήματα λήψης αποφάσεων ή γενικότερα εφαρμογές τεχνητής νοημοσύνης, απαιτείται η ύπαρξη ενός ικανού αριθμού στελεχών και ανθρώπινου δυναμικού που θα μπορεί να συμμετάσχει στο σχεδιασμό, στη παρακολούθηση, στη βελτιστοποίηση και ακόμα και στη χρήση ενός αντίστοιχου συστήματος. Όμως, τα εν λόγω συστήματα απαιτούν υψηλού επιπέδου κατάρτιση και συγκεκριμένες δεξιότητες ανά περίπτωση, και αυτό αποτελεί μια ακόμα πρόκληση (Gupta, 2019). Το σύγχρονο εκπαιδευτικό σύστημα έχει αργούς ρυθμούς αφομίωσης και παροχής υψηλών τεχνικών δεξιοτήτων στους νέους, ενώ και στον εργασιακό χώρο, οι προφορά αντίστοιχων εκπαιδεύσεων είναι περιορισμένη. Πόσο μάλλον δε στο δημόσιο τομέα, όπου οι συνθήκες συνήθως δεν ευνοούν τη εξέλιξη των τεχνικών ικανοτήτων των εργαζομένων.

Επίσης, και όσον αφορά τη τέταρτη πρόκληση που αναφέρεται παραπάνω, η οποία έχει να κάνει με τη ερμηνεία των αποτελεσμάτων της τεχνητής νοημοσύνης, η διαδικασία αυτή είναι αρκετά πολύπλοκη. Όπερ σημαίνει, πως ο συνδυασμός έλλειψης τεχνογνωσίας ή κατανόησης ενός συστήματος τεχνητής νοημοσύνης και η πολυπλοκότητα του, δυσχεραίνουν την αφομίωση και κατανόηση των αποτελεσμάτων πρώτα από τον ίδιο τον οργανισμό αλλά και από τους πολίτες ή τους τελικούς χρήστες. Συνήθως, τα αποτελέσματα χρήσης της τεχνητής νοημοσύνης είναι δύσκολο να αιτιολογηθούν, μιας και ο τρόπος σχεδιασμού αλλά και η λειτουργία ενός τέτοιου συστήματος παραμένουν δυσνόητοι. Εξού και η πρόταση χαρακτηρισμού του τρόπου επεξεργασίας και εξαγωγής αποτελεσμάτων των συστημάτων μηχανικής μάθησης και τεχνητής νοημοσύνης, ως μαύρου κουτιού (Frischmann και Selinger, 2018), καθώς σε αρκετές περιπτώσεις δε δύναται να δικαιολογηθεί επαρκώς το τι συνέβη και πως προέκυψε ένα αποτέλεσμα. Ειδικότερα, σε αντίστοιχα συστήματα υποστήριξης λήψης

αποφάσεων, το έργο των υπευθύνων να ερμηνεύσουν τα αποτελέσματα της τεχνητής νοημοσύνης και να ορίσουν τις αντίστοιχες διορθωτικές ενέργειες ή να χαράξουν νέες πολιτικές μπορεί να οδηγήσει σε εσφαλμένες αποφάσεις λόγω αδυναμίας ορθής ερμηνείας (Janssen et al., 2020).

Πέμπτον, οι προκλήσεις νομιμότητας αλλά και η ηθικές προκλήσεις που προκύπτουν από τη χρήση της τεχνητής νοημοσύνης είναι αρκετές. Ειδικότερα, και όσον αφορά τις ηθικές προκλήσεις, στη βιβλιογραφία αναφέρονται τα ηθικά διλήμματα που προκύπτουν, όπως η μη ηθική κτήση και χρήση δεδομένων, προσωπικών και μη, οι διακρίσεις που προκαλεί η χρήση της τεχνητής νοημοσύνης και η πιθανώς άδικη παροχή δημοσίων υπηρεσιών (Chen et al., 2019), αλλά και η ηθική των μηχανών - ρομπότ.

Αρχικά, τα συστήματα τεχνητής νοημοσύνης σχεδιάζονται και τροφοδοτούνται με δεδομένα εκπαίδευσης ήδη από τα πρώτα στάδια δημιουργίας τους. Ένα αντίστοιχο σύστημα λοιπόν δύναται να αντανακλά, να αναπράγει και να ενισχύει τις αξίες αλλά και τις προκαταλήψεις και απόψεις των σχεδιαστών του και του περιβάλλοντος στο οποίο δρα, λόγω των δεδομένων που το αναδοτροφούν συνεχώς. Θα πρέπει οι σχεδιαστές να αναλύσουν και να αποφασίσουν ποια δεδομένα και ποιας ποιότητας, θα πρέπει να χρησιμοποιηθούν. Επιπρόσθετα, τα συστήματα τεχνητής νοημοσύνης αλληλεπιδρούν με το περιβάλλον συλλέγοντας συνεχώς πληροφορίες και δεδομένα. Ως αποτέλεσμα, θα πρέπει οι σχεδιαστές να λάβουν ιδιαίτερα υπόψιν και τη συλλογή δεδομένων κατά τη διάρκεια λειτουργίας ενός τέτοιου συστήματος. Αυτό σημαίνει πως ακούσια ή εκκούσια, είναι πιθανή η συλλογή προσωπικών ακόμη και ευαίσθητων δεδομένων των υποκειμένων μέσα από αλληλεπίδραση με άλλα συστήματα ή εφαρμογές. Υπάρχει λοιπόν, πιθανότητα κατάχρησης προσωπικών δεδομένων ώστε ένα αλγόριθμος να μπορεί να παράσχει πιο εξατομικευμένες υπηρεσίες. Σύμφωνα με τον Mittelstadt et al. (2016), οι εξατομικευμένες υπηρεσίες με χρήση προσωπικών δεδομένων από αντίστοιχα συστήματα περιορίζουν την προσωπική αυτονομία του υποκειμένου και αυτό θα μπορούσε να ερμηνευτεί ως ανήθικο. Άρα και ως κίνδυνος που θα πρέπει να ληφθεί υπόψιν καθώς εγείρονται και θέματα νομιμότητας και παραβίασης ανθρωπίνων δικαιωμάτων, όπως το απόρρητο.

Επίσης, τα δεδομένα που συλλέγονται αντανακλούν και τις πεποιθήσεις των υποκειμένων και της κοινωνίας γενικότερα. Ως εκ τούτου, δύναται ένα σύστημα τεχνητής νοημοσύνης να αναπαράγει φυλετικές, θρησκευτικές ή μεροληψίες σχετικές με το φύλο και τα αποτελέσματα της χρήσης του συστήματος να αμβλύνουν τις προκαταλήψεις. Άρα, οι σχεδιαστές αλλά και μετέπειτα οι διαχειριστές θα πρέπει τόσο σε επίπεδο σχεδιασμού και σε επίπεδο λειτουργίας να λάβουν υπόψιν τη πιθανότητα αυτή. Συνάμα θα πρέπει να εξετάζουν ανά τακτά χρονικά διαστήματα τόσο τη ποιότητα αλλά και τη προέλευση των δεδομένων που επεξεργάζεται ένα σύστημα τεχνητής νοημοσύνης. Η σύννομη κτήση και χρήση δεδομένων είναι μια πρόκληση ειδικά στις περιπτώσεις όπου τα αποτελέσματα μιας αντίστοιχης εφαρμογής αγγίζουν ή ξεπερνούν τα όρια της νομιμότητας.

Γενικά, τα προαναφερθέντα εγείρουν προβληματισμούς και όσον αφορά τη νομιμότητα. Τόσο σε επίπεδο απορρήτου, λόγω μη νόμιμης απόκτησης και επεξεργασίας προσωπικών δεδομένων των υποκειμένων, όσο και σε επίπεδο ασφάλειας, καθώς πολύπλοκα συστήματα τεχνητής νοημοσύνης διαχειρίζονται, συνήθως με δυσνόητο ή / και αδιαφανή τρόπο, ευαίσθητες πληροφορίες. Όσον αφορά την ασφάλεια των εν λόγω συστημάτων, αναφέρεται στη βιβλιογραφία, τόσο η ασφάλεια της πληροφορίας και των δεδομένων, όσο και η εν γένει συστημική ασφάλεια μιας εφαρμογής σε επίπεδο προστασίας από απόκτηση μη εξουσιοδοτημένης πρόσβασης ή ελέγχου και δυσμενούς χειραγώγησης από κακόβουλους. Οι Bostrom και Yudkowsky (2014) υπογραμμίζουν τη σημασία και αναγκαιότητα η τεχνητή νοημοσύνη να είναι ανθεκτική σε κακόβουλη χειραγώγηση από τον άνθρωπο.

Όσον αφορά τις ηθικές προκλήσεις σχετικές με τις εφαρμογές της τεχνητής νοημοσύνης, στη βιβλιογραφία αναφέρονται δυο βασικοί προβληματισμοί (Wirtz et al., 2019). Αρχικά, ο πρώτος προβληματισμός αφορά την ενσωμάτωση των ηθικών αρχών σε ένα τέτοιο σύστημα ώστε να διασφαλιστεί πως αυτά ενεργεί σύννομα και βάση των αξιακών αρχών της κοινωνίας (Anderson & Anderson, 2011). Απο τη μια, αποτελεί ιδιαίτερη τεχνική πρόκληση η ενθυλάκωση τέτοιων κανόνων στα συστήματα τεχνητής νοημοσύνης, αλλά επίσης και ο καταρτισμός ενός κοινά αποδεκτού ορισμού τέτοιων προτύπων ηθικής σε επίπεδο σχεδιασμού δεν είναι εύκολος, ιδιαίτερα σε ένα συνεχώς μεταβαλλόμενο

περιβάλλον. Επίσης, ενδέχεται τα συστήματα τεχνητής νοημοσύνης να αναπτύξουν το δικό τους σύστημα αξιών και να δημιουργήσουν το δικό τους πλαίσιο ηθικής αναφοράς (Wirtz et al., 2019), λόγω της εξελισσόμενης φύσης τους και της αλληλεπίδρασης με το περιβάλλον τους, το οποίο τελικά μπορεί να είναι ασύμβατο με το ανθρώπινο σύστημα αξιών. Το παραπάνω μπορεί να οδηγήσει δυνητικά ένα σύστημα τεχνητής νοημοσύνης να εναντιωθεί ακόμη και στον ίδιο τον άνθρωπο (Mittelstadt et al., 2016). Ο δεύτερος προβληματισμός συνδέεται με το παραπάνω και έχει να κάνει κυριώς με τα αυτόνομα οπτικά συστήματα τα οποία δύναται να λειτουργούν δίχως καμιά ανθρώπινη παρέμβαση, ιδιαίτερα κατά τη λειτουργία τους και πιθανόν να λαμβάνουν θανατηφόρες αποφάσεις που να μην δικαιολογούνται ηθικά (Patrick., Abney, & Bekey, 2012). Εξού και η εντονότερη κριτική για άμεση μελέτη των ηθικών επιπτώσεων των ερευνητών μηχανικής μάθησης και τεχνητής νοημοσύνης στο στρατιωτικό τομέα σε σχέση με αντίστοιχους ερευνητές σε άλλους τομείς εφαρμογής της τεχνητής νοημοσύνης (Zhang et. al., 2021).

Έκτον, από τη χρήση της τεχνητής νοημοσύνης προκύπτουν και πολιτικές προκλήσεις. Η εν λόγω τεχνολογία αυτή, αναθεωρεί εκ βάθρων τη δομή και λειτουργία των σύγχρονων κοινωνιών. Υπάρχει κίνδυνος οι εφαρμογές της τεχνητής νοημοσύνης να υπονομεύσουν, εσκεμμένα ή μη, θεμελιώδεις αξίες ενός κράτους δικαίου. Οι δυστοπικές προβλέψεις της εξέλιξης της τεχνητής νοημοσύνης, όπως αυτές αποτυπώνονται στο κινηματογραφικό χώρο, πιθανόν να είναι υπερβολικές, τουλάχιστον προς το παρόν. Παρόλα αυτά όμως, ακόμη και εφαρμογές αδύναμης τεχνητής νοημοσύνης δύναται να επηρεάσουν άμεσα και ουσιαστικά όλες τις εκφάνσεις της σύγχρονης ζωής. Δίχως να είναι ξεκάθαρο ποιος είναι υπεύθυνος για τη αλγοριθμικές αποφάσεις, χωρίς διαφάνεια για το ποιος είναι υπόλογος, και ποιος κατέχει τον έλεγχο ενός συστήματος τεχνητής νοημοσύνης, δεν νοείται κράτος δικαίου. Ως αποτέλεσμα, τίθονται ζητήματα χάραξης πολιτικής, στρατηγικής και δημιουργίας αντίστοιχου νομικού πλαισίου. Εξίσου σημαντικό πολιτικό θέμα είναι και η συγκέντρωση της ανάπτυξης και του ελέγχου της τεχνητής νοημοσύνης σε ελάχιστο αριθμό εταιριών και προσώπων. Ειδικότερα, οι τεράστιοι τεχνολογικοί κολοσσοί, Apple, Amazon, Google, Facebook, Tesla κτλ., έχουν ισχυρό προβάδισμα στη κατανόηση και την ανάπτυξη εφαρμογών τεχνητής νοημοσύνης, στενές διασυνδέσεις μεταξύ τους, ενώ το μέγεθος τους, τουλάχιστον μέχρι σήμερα, τους επιτρέπει

μυστικιστικές και αυταρχικές συμπεριφορές. Επίσης στο αυτό το τομέα των μεγάλων εταιριών τεχνολογίας υπάρχει και έλλειψη λογοδοσίας. Συνδυάζοντας το παραπάνω με την εν γένει πολυπλοκότητα της τεχνολογίας αυτής, τα προαναφερθέντα ζητήματα αναβαθμίζονται σε κρίσιμες πολιτικές προκλήσεις που θα πρέπει να αντιμετωπιστούν άμεσα.

Έβδομον, από τη χρήση τεχνητής νοημοσύνης προκύπτουν και κοινωνικές προκλήσεις που θα πρέπει να αντιμετωπιστούν αντιστοίχως. Ειδικότερα, οι εφαρμογές της τεχνητής νοημοσύνης και μηχανικής μάθησης δύναται να μετασχηματίσουν την αγορά εργασίας με ιδιαίτερα γρήγορους ρυθμούς. Η αντικατάσταση του ανθρώπινου δυναμικού με αυτόνομες μηχανές ή με συστήματα σε απλοποιημένες και επαναλαμβανόμενες εργασίες δύναται να ενισχύσει τις κοινωνικές ανισότητες, ειδικά για την εργατική τάξη, υποσκάπτωντας και τις δυνατότητες καινοτομίας της κοινωνίας (Gupta, 2019). Η πρόκληση για τις κυβερνήσεις είναι η ομαλή μετάβαση από την υπάρχουσα κατάσταση σε μια κοινωνία ευημερίας, όπου η εργασία είναι αναπόφευκτο δικαίωμα και όπου η τεχνητή νοημοσύνη δεν θα αντικαταστήσει τον άνθρωπο, αλλά θα συνδράμει θετικά στην ποιότητα της παραγωγής και της εργασίας. Επίσης, ο κοινωνικός αντίκτυπος της χρήσης της τεχνητής νοημοσύνης στο πλήρες φάσμα των ανθρώπινων δραστηριοτήτων δύναται να προκαλέσει ένα απρόσωπο, αριθμητικό περιβάλλον που υπονομοεύει τη κοινωνική διάσταση των υποκειμένων. Ως εκ τούτου, η κοινωνική συνοχή και η εμπιστοσύνη των πολιτών στις κυβερνήσεις μπορεί να διατηρηθεί με δυσμενή αποτελέσματα για τους ευάλωτους κυρίως πολίτες. Η κοινωνική αποδοχή της τεχνητής νοημοσύνης είναι βασικός πυλώνας επιτυχίας της τεχνολογικής μετάβασης που είναι ήδη σε εξέλιξη (Wirtz et al., 2019).

Τέλος, η όγδοη βασική πρόκληση από τη χρήση της τεχνητής νοημοσύνης είναι οι οικονομικοί κίνδυνοι που πιθανόν προκύπτουν. Δεν είναι ακόμη ξεκάθαρο στους ερευνητές ποιο θα είναι το οικονομικό αντίκτυπο από τη χρήση της τεχνολογίας αυτής. Ο υπερενθουσιασμός που επικρατεί για την άμεση χρήση της μπορεί να οδηγήσει σε υπερβολικές προσδοκίες και τεράστια χρήματα να δαπανηθούν σε περιπτώσεις όπου δεν υπάρχει αντίστοιχη ωριμότητα σχεδιασμού και αφομίωσης της. Για παράδειγμα, μονάχα το ενεργειακό αποτύπωμα και κόστος, αποκλειστικά για την εκπαίδευση ενός

συστήματος επεξεργασίας φυσικής γλώσσας με χρήση τεχνητής νοημοσύνης, ισοδυναμεί με το κόστος και τη ρύπανση εκατόν είκοσι πέντε (125) αεροπορικών ταξιδίων μετ' επιστροφής από τη Νέα Υόρκη στο Πεκίνο (Dhar, 2020). Επίσης, και στις περιπτώσεις επιτυχούς εφαρμογής, το συνολικό κόστος δεν δύναται να αποτιμηθεί εύκολα. Για παράδειγμα τα ασφαλιστικά και φορολογικά μοντέλα που χρησιμοποιούνται σήμερα μάλλον δεν μπορούν να καλύψουν επιτυχώς τον εργασιακό μετασχηματισμό και μετάβαση σε μια κοινωνία με ευρεία χρήση της τεχνητής νοημοσύνης.

Γενικά, η επένδυση για τη δημιουργία μιας σύγχρονης τεχνολογικής υποδομής για αποθήκευση, συλλογή και επεξεργασία μεγάλου όγκου δεδομένων είναι τεράστια (Roberts, 2017), ενώ παράλληλα, η μεγάλη ζήτηση για εμπειρογνώμονες του αντικειμένου εκτοξεύει, τόσο το μισθολογικό κόστος, όσο και το κόστος συνεχούς εκπαίδευσής τους (Bughin et al., 2017). Είναι σαφές πως τόσο σε επίπεδο σχεδιασμού όσο και σε επίπεδο λειτουργίας οι οικονομικές προκλήσεις από τη χρήση της τεχνητής νοημοσύνης είναι και ανεπαρκώς μελετημένες αλλά και δυσεπίλιτες.

5 Διακυβέρνηση της Τεχνητής Νοημοσύνης

5.1 Το Πρόβλημα της Διαχείρισης της Τεχνητής Νοημοσύνης

Δεδομένων των προαναφερθέντων προκλήσεων και κινδύνων από τη χρήση της τεχνητή νοημοσύνης, οι απειλές είναι αδιαμφισβήτητα πραγματικές. Ως αποτέλεσμα, οι κυβερνήσεις θα πρέπει άμεσα να επιλέξουν το τρόπο διακυβέρνησης, τις απαιτούμενες ρυθμίσεις του σχεδιασμού και λειτουργίας αυτής της τεχνολογίας, αλλά και να προβούν στη διαμόρφωση ενός αποτελεσματικού νομικού πλαισίου.

Ειδικότερα, για το δημόσιο τομέα και για τις κυβερνήσεις, οι προκλήσεις διακυβέρνησης της τεχνητής νοημοσύνης διογκώνονται σε σχέση με τον ιδιωτικό τομέα, καθώς είναι υπεύθυνοι για τη διατήρηση και την εξέλιξη μιας έννομης και ευημερίζουσας κοινωνίας. Αρχικά, οι δημόσιοι οργανισμοί δεν έχουν, ακόμα και είναι ψηφιακά ενοποιημένοι, τον απαιτούμενο όγκο και ποιότητα δεδομένων για την αποτελεσματική παροχή υπηρεσιών που βασίζονται στη τεχνητή νοημοσύνη, ειδικότερα στις μικρές και αναπτυσσόμενες χώρες. Τίθονται ακόμα και νομικά ερωτήματα με ποιο τρόπο οι δημοσίοι οργανισμοί μπορούν να εκμεταλευτούν δεδομένα των πολιτών τους που πιθανόν υπάρχουν σε άλλους τομείς πέρα του στενού δημοσίου τομέα.

Πέραν όμως της ικανότητας υιοθέτησης εφαρμογών τεχνητής νοημοσύνης και παροχής δημοσίων υπηρεσιών από το δημόσιο τομέα, οι κυβερνήσεις οφείλουν να ρυθμίζουν τις κοινωνικοοικονομικές δραστηριότητες των πολιτών σχεδιάζοντας και εφαρμόζοντας κανόνες και νόμους. Οι νόμοι δε, θα πρέπει να είναι προβλέψιμοι, ενώ όλοι θα πρέπει να είναι ίσοι έναντι αυτών (Marsh, 1959). Όμως, στη περίπτωση της τεχνητής νοημοσύνης, η οποία εκ της φύσεώς της αποτελεί μια δυναμική και συνεχώς μεταβαλλόμενη έννοια, η ρύθμιση της είναι πραγματικά μια τεράστια πρόκληση. Η ευθύνη και η λογοδοσία, τα οποία είναι βασικά συστατικά του κράτους δικαίου, στη περίπτωση της τεχνητής νοημοσύνης είναι τουλάχιστον θολά, σε υπερβολικό αριθμό περιπτώσεων. Μέχρι σήμερα, τα συστήματα αλγοριθμικών αποφάσεων και γενικά οι εφαρμογές τεχνητής νοημοσύνης, αντιμετωπίζονται, ανεπιτυχώς, με τα υφιστάμενα νομικά πλαίσια. Δηλαδή, αντιμετωπίζονται, συνήθως, ως απλές μηχανές και εργαλεία τα οποία ελέγχονται μερικώς ή πλήρως από τον άνθρωπο (Matthias, 2004), το οποίο όμως δεν συνάδει με τη πραγματικότητα. Υπάρχουν περιπτώσεις όπου ένα σύστημα τεχνητής νοημοσύνης καλείται να αντιμετωπίσει τυχαίες ή απροσδόκητες καταστάσεις για τις οποίες δεν έχει εκπαιδευτεί κατά τον σχεδιασμό του, και παρόλα αυτά προβαίνει σε μια ενέργεια. Στη βιβλιογραφία, αυτές οι περιπτώσεις αναφέρονται ως “ γωνιακές περιπτώσεις” (corner cases) (Bostrom & Ludkowsky 2014), και αποτελεί άγνωστο ποια θα είναι η ενέργεια που θα επιλέξει να προβεί το αντίστοιχο σύστημα. Μάλιστα, σε παρόμοια περιπτώσεις δεν είναι δεδομένο πως το εν λόγω σύστημα θα επιλέξει την ίδια ενέργεια. Εξετάζοντας την πολυπλοκότητα των συστημάτων μηχανικής μάθησης, αλλά και την συνεχή αυτοεκπαίδευση τους κατά τη διάρκεια λειτουργίας τους, είναι δεδομένη η δυσκολία κατανόησης το πως και γιατί ελήφθη μια απόφαση από ένα τέτοιο σύστημα. Η αυτονομία των εν λόγω συστημάτων στη λήψη αποφάσεων χωρίς ανθρώπινη παρέμβαση γεννά ερωτηματικά σχετικά με την απόδοση ευθύνης από νομική σκοπιά. Λαμβάνοντας υπόψιν την απουσία ελέγχου από τον άνθρωπο οι προβληματισμοί για την αυστηρή ευθύνη δεν μπορεί να βαραίνει πρακτικά και εξ ολοκλήρου τους σχεδιαστές τους συστήματος. Ως αποτέλεσμα, οι προβληματισμοί ευθύνης θα πρέπει να αντιστοιχισθούν σε όλο το κύκλο ζωής ενός συστήματος τεχνητής νοημοσύνης ώστε να επιτευχθεί η απαιτούμενη ευθύνη και η λογοδοσία. Όμως, και σε αυτή τη περίπτωση, μπορεί τελικώς ούτε οι σχεδιαστές, ούτε οι υπεύθυνοι λειτουργίας να μπορούν να γνωρίζουν ή να προβλέψουν τη

συμπεριφορά ενός τέτοιου συτήματος, δεδομένης της συνεχούς διάδρασης με το περιβάλλον και της αυτονομίας στη λήψη αποφάσεων. Δημιουργείται ανοπόφευκτα ένα κενό ευθύνης (Matthias, 2004) λόγω της έλλειψης ελέγχου στο σύστημα. Στις περιπτώσεις αυτές και συγκεκριμένα στο δημόσιο τομέα, είναι κρίσιμο να δημιουργηθεί ένα πλαίσιο πρακτικής διαχείρισης και από νομική σκοπιά, όπερ σημαίνει και την αντίστοιχη επικαιροποίηση του δημοσίου διοικητικού δικαίου, ώστε να αποφευχθούν κενά ευθύνης και λογοδοσίας. Παράλληλα δε, οι κυβερνήσεις και ο δημόσιος τομέας θα πρέπει να εξετάσουν και το νομικό πλαίσιο ευθύνης και λογοδοσίας των ιδιωτικών πρωτοβουλιών εφαρμογής τεχνητής νοημοσύνης σε ένα παγκοσμιοποιημένο περιβάλλον απέναντι σε μια ομάδα εταιριών - κολοσσών με τεράστιες δυνατότητες επιρροής στη διαμόρφωση του αντίστοιχου πλαισίου. Πιθανόν, μεμονομένες προσπάθειες σε τοπικό, ή ακόμα και εθνικό επίπεδο, να μπορούσαν να καλύψουν θεωρητικά τα όποια νομικά κενά της χρήσης της τεχνητής νοημοσύνης, όμως η πρακτική εφαρμογή και επιβολή των νόμων στη παγκοσμιοποιημένη κοινωνία να μην είναι εφικτή.

5.2 Οι Επιλογές Διακυβέρνησης της Τεχνητής Νοημοσύνης

Είναι κατανοητό πως αθροίζοντας τις προαναφερθείσες προκλήσεις συμπεραίνεται εύλογα πως η διακυβέρνηση και η ρύθμιση της τεχνητής νοημοσύνης αποτελούν το κλειδί επιτυχίας για την αφομίωση και τη χρηστή υιοθέτηση της τεχνολογίας αυτής, ειδικά στο δημόσιο τομέα. Η διαχείριση των κινδύνων είναι ζωτικής σημασίας για να καταφέρουν οι κυβερνήσεις, ο δημόσιος τομέας αλλά και ιδιωτικές πρωτοβουλίες να επωφεληθούν από τις δυνατότητες της τεχνητής νοημοσύνης.

Από τη μια, οι ιδιωτικές εταιρίες κολοσσοί (π.χ. Apple, Tesla, Google, Facebook, κτλ.) διαθέτουν τεράστιους οικονομικούς πόρους αλλά και τον απαιτούμενο όγκο δεδομένων. Παράλληλα, λόγω της συγκεκριμένης στόχευσης και κατανόησης του πεδίου επιχειρηματικής τους δράσης, μπορούν να διατυπώσουν με σχετική σαφήνεια τους στρατηγικούς στόχους χρήσης της τεχνητής νοημοσύνης. Επίσης, η ευελιξία σε επίπεδο οργάνωσης και η προσαρμοστικότητα σε αλλαγές τους δίνει ένα αξιοσημείωτο προβάδισμα αποτελεσματικής διακυβέρνησης της τεχνολογίας σε σχέση με τις κυβερνήσεις και τους δημόσιους οργανισμούς.

Αντίθετα, οι κυβερνήσεις και κατ' επέκταση οι δημόσιοι οργανισμοί στερούνται των παραπάνω δυνατοτήτων. Πάραυτα, προσπαθούν ήδη να δημιουργήσουν τις στρατηγικές υιοθέτησης αλλά και το απαιτούμενο ρυθμιστικό πλαίσιο ώστε να διαχειριστούν μια εν εξελίξη πραγματικότητα.

Πρακτικά, οι κυβερνήσεις καλούνται να επιλέξουν τη μορφή διακυβέρνησης και ρύθμισης μιας τεχνολογίας που δεν κατανοούν σε βάθος. Οι δυνατές επιλογές διακυβέρνησης εκτείνονται από την υιοθέτηση αυστηρών ρυθμιστικών πλαισίων έως τη πλήρη αυτορρύθμιση. Ξεκινώντας από την αυστηρή ρύθμιση, παρατηρείται αδυναμία εφαρμογής αυτής της επιλογής. Οι ραγδαίες εξελίξεις της τεχνολογίας αυτής, η ασύμμετρη υπερπληροφόρηση από παράγοντες που δρουν στο χώρο της τεχνητής νοημοσύνης (Taeihagh, 2021), αλλά και η έλλειψη κατανόησης και πρακτικής εφαρμογής ενός αυστηρού πλαισίου οδηγεί συχνά στη γενικευμένη διατύπωση νόμων που είναι ασαφείς και αδυνατούν πρακτικά να ρυθμίσουν τη εν λόγω τεχνολογία (Larsson, 2020). Η εκ των προτέρων υιοθέτηση απαγορευτικών μέτρων, ευρύτερα γνωστή στο νομικό κλάδο ως “ αρχή της προφύλαξης”, πιθανόν να προστατεύει από δυσμενή αποτελέσματα, όμως, περιορίζει τις δυνατότητες καινοτομίας και τα δυνητικά μακροπρόθεσμα οφέλη.

Στον αντίποδα, υπάρχει η επιλογή της πλήρους αυτορρύθμισης της τεχνητής νοημοσύνης που όμως εγκυμονεί κινδύνους. Το παθομολογούμενο προβάδισμα των ιδιωτικών εταιριών στη τεχνητή νοημοσύνη σε συνδυασμό με τη διαφορετική στόχευση τους έναντι των κυβερνήσεων (κέρδος έναντι δημόσιας αξίας), πιθανόν να δημιουργήσουν μεσομακροπρόθεσμα ένα περιβάλλον ανισορροπίας και διατάραξη της κοινωνικής συνοχής. Όταν οι πολιτικές και οι κανόνες εισάγονται αποκλειστικά από τον ιδιωτικό τομέα, είναι λογικό να αντικατροπτίζουν και να εξυπηρετούν τις στρατηγικές τους, ενώ παράλληλα, δεσμεύουν τις ρυθμιστικές αρχές εκ των προτέρων.

Επίσης, η πολυπλοκότητα των συστημάτων τεχνητής νοημοσύνης, και γενικά, η θεωρία του μαύρου κουτιού των αλγορίθμων μηχανικής μάθησης, θα μπορούσαν να θεμελιώσουν μια εξάρτηση των ρυθμιστών αρχών από τον ιδιωτικό τομέα στην ερμηνεία και την αξιολόγηση των αποτελεσμάτων (Guihot et al., 2017). Επίσης, οι αρχές και οι κανονισμοί που τίθενται από τον ιδιωτικό τομέα στο πλαίσιο της αυτορρύθμισης, συχνά

γίνονται μακριά από το δημόσιο έλεγχο με αποτέλεσμα το έλλειμα κατανόησης από τις κυβερνήσεις να γιγαντώνεται. Τα παραπάνω θα μπορούσαν να επιδεινώσουν σημαντικά τις όποιες δυνατότητες των κυβερνήσεων και του δημόσιου τομέα να παρακολουθήσει τις εξελίξεις στη παράξουν την επιθυμητή δημόσια αξία.

Όμως, πέρα των δυο άκρων της αυστηρής ρύθμισης έναντι της αυτορρύθμισης, υπάρχει και η μέση οδός ως επιλογή, δηλαδή, η επιβολή ηπιότερων μέτρων ρύθμισης, ευρέως γνωστά ως “soft law”. Η προσέγγιση αυτή αναφέρεται στην υιοθέτηση μη δεσμευτικών κανόνων, προτύπων, κατευθυντήριων γραμμών ακόμη και ενός κώδικα δεοντολογίας που καταρτίζονται από φορείς του κλάδου και τις ρυθμιστικές αρχές και δημιουργούν ένα ουσιαστικό πλαίσιο προσδοκιών χωρίς να δεσμεύει άμεσα τους εμπλεκόμενους (IEEE 2019). Επιπρόσθετα, οι κυβερνήσεις, με την υποστήριξη εμπειρογνομένων, καταρτίζουν αντίστοιχα στρατηγικές αλλά και οδηγίες ηθικής και διακυβέρνησης της τεχνητής νοημοσύνης ώστε να δημιουργήσουν ένα αρχικό πλαίσιο ρύθμισης και να επιτρέψουν την αποτελεσματική υιοθέτησή της τεχνολογίας. Σημειώνεται δε, πως οι παραπάνω πρακτικές ήπιας ρύθμισης ενισχύουν την αναγκαία ευελιξία καθώς μπορούν να τροποποιηθούν και να προσαρμοστούν ταχύτερα σε σχέση με αυστηρότερες μορφές ρυθμίσεις (νόμους, κτλ.) (Taeihagh, 2021), συμβαδίζοντας με τις όποιες τεχνολογικές εξελίξεις (Larsson, 2020). Επιπρόσθετα, στην επιλογή ηπιότερης ρύθμισης η αυστηρότητα δύναται να ποικίλει ανάλογα με τη στρατηγική που έχει καταρτιστεί, όμως το ουσιαστικό σε αυτή τη περίπτωση είναι ο χρόνος απόκρισης στις προκλήσεις και ο τρόπος δράσης στα αποτελέσματα χρήσης της τεχνολογίας. Παρόλα αυτά και στη περίπτωση ήπιας ρύθμισης υπάρχουν προκλήσεις που θα πρέπει να αντιμετωπιστούν. Ειδικότερα, οι κυβερνήσεις θα πρέπει να εποπτεύουν τη τήρηση των κατευθυντήριων γραμμών και κανόνων σε όλο το κύκλο ζωής ενός συστήματος τεχνητής νοημοσύνης. Από το σχεδιασμό έως τη λειτουργία του εν λόγω συστήματος, οι κυβερνήσεις θα πρέπει να δημιουργήσουν μηχανισμούς εποπτείας ώστε να διασφαλιστεί η συμμόρφωση με τις αρχές και αξίες που αυτές προσβέουν. Ακόμη, θα πρέπει να διασφαλιστεί πως το ηπιότερο πλαίσιο δεν εναντιώνεται σε υφιστάμενες νομοθεσίες ή κανόνες και κατευθυντήριες γραμμές σε άλλους εφαπτόμενους τομείς της διακυβέρνησης, καθώς αυτό θα δημιουργούσε τόσο θέματα συντονισμού, όσο και επιβολής των κανόνων (Guihot et al., 2017).

Γενικά, από τις παραπάνω προσεγγίσεις, η ηπιότερη ρύθμιση είναι αυτή που επιλέχθηκε σε παρόμοιες περιπτώσεις αναδυόμενων τεχνολογιών (π.χ. διάδοση του διαδικτύου) ώστε να προωθηθούν οι ηθικές αξίες αλλά και να αποφευχθούν μεροληπτικές και επιζήμιες για τη κοινωνία πρακτικές (Gasser & Almeida, 2017). Παρόλα αυτά, θα πρέπει να γίνουν συντονισμένες και διεπιστημονικές ενέργειες για την αξιολόγηση, τόσο των πρώιμων αποτελεσμάτων χρήσης της τεχνητής νοημοσύνης, όσο και για το καταρτισμό και την επικαιροποίηση στρατηγικών και ρυθμιστικών πλαισίων που θα διασφαλίζουν τις θεμελιώδεις αρχές και τα συμφέροντα της σύγχρονης κοινωνίας.

5.3 Προτεινόμενη Προσέγγιση για τη Διακυβέρνηση της Τεχνητής Νοημοσύνης

Είναι σαφές πως η επιλογή του τρόπου διακυβέρνησης της τεχνητής νοημοσύνης είναι καθοριστική ώστε να περιοριστούν οι υπαρκτοί κίνδυνοι και να μεγιστοποιηθούν τα πιθανά οφέλη. Για την αντιμετώπιση της πολυπλοκότητας αλλά κυρίως της αβεβαιότητας που προκαλεί η τεχνολογία αυτή, οι προαναφερθέντες προσεγγίσεις πιθανόν να μην είναι ικανές να διασφαλίσουν μια ομαλή μετάβαση σε μια ευημερίζουσα κοινωνία με ευρεία χρήση της τεχνητής νοημοσύνης. Ακόμη και η επιλογή της ήπιας ρύθμισης, όπως χρησιμοποιήθηκε σε παλαιότερες περιπτώσεις (π.χ. διαδίκτυο), ίσως να μην είναι αρκετή για τη διασφάλιση των θεμελιωδών αξιών της κοινωνίας. Εξού και οι πρόσφατες εκκλήσεις από τους ερευνητές για την υιοθέτηση ενός νέου μοντέλου διακυβέρνησης για την τεχνητή νοημοσύνη. Ειδικότερα, προτείνεται σθεναρά η υιοθέτηση ενός υβριδικού μοντέλου το οποίο έχει αφετηρία την δημιουργία ενός πλαισίου, αλλά παράλληλα έχει έντονα χαρακτηριστικά προσαρμογής στις εξελίξεις (Leiser & Murray 2016; Linkov et al., 2018; Tan & Taeihagh, 2021b). Προτείνεται δηλαδή, μια υβριδική διακυβέρνηση της τεχνολογίας με όλους τους κύριους παράγοντες να συμμετέχουν ενεργά, τόσο στην αρχική διαμόρφωση του ρυθμιστικού πλαισίου, όσο και στη συνέχεια κατά τον έλεγχο των αποτελεσμάτων και την επικαιροποίηση των στρατηγικών. Αυτό το καινοτόμο μοντέλο διακυβέρνησης θα δίνει τη δυνατότητα και ανάλογα με τη περίπτωση, να υιοθετούνται υβριδικές πρακτικές από όλο το φάσμα ρυθμιστικών δυνατοτήτων, από την αυτορρύθμιση και την από κοινού ρύθμιση έως την επιβολή αυστηρών κανόνων. Το κλειδί σε αυτό το μοντέλο είναι η σύμπραξη όλων των δυνάμεων που δραστηριοποιούνται στη τεχνητή νοημοσύνη είτε ως σχεδιαστές, ή ως ρυθμιστές ή ακόμα και ως παράγοντες που τελικώς επηρεάζονται από τη χρήση της

τεχνολογίας αυτής. Βασικό χαρακτηριστικό σε αυτό το υβριδικό μοντέλο διακυβέρνησης, είναι ο μειωμένος ρόλος των κυβερνήσεων δίνοντας χώρο σε μη κρατικούς φορείς να συνεισφέρουν στις εξελίξεις. Παρόλα αυτά, επιβάλεται μια δημόσια ρυθμιστική εποπτεία των δράσεων των ιδιωτικών φορέων με τη μορφή της ενδεδειγμένης παρακολούθησης των εξελίξεων αλλά παράλληλα, θα πρέπει να υιοθετηθεί μια επαναληπτική προσαρμογή και βελτίωση των κανονισμών, του πλαισίου και γενικά των στρατηγικών καθώς συνεχώς συλλέγονται νέες πληροφορίες (Linkon et al., 2018; Tan & Taeihagh, 2021b).

Σημειώνεται δε, πως οι παράγοντες που θα συμμετείχαν σε μια τέτοια σύμπραξη για την διακυβέρνηση της τεχνητής νοημοσύνης, δεν είναι μόνο κυβερνήσεις, φορείς του δημοσίου τομέα και ιδιωτικές εταιρίες τεχνολογίας. Δυστυχώς, και ανάλογα με την περίπτωση, οργανισμοί πιστοποίησης (π.χ. Διεθνής Οργανισμός Τυποποίησης – ISO, IEEE, WIPO, κτλ.), κοινωνικοί φορείς, υπερεθνικές οντότητες (π.χ. ΟΗΕ, ΟΟΣΑ) και μη κυβερνητικές οργανώσεις (ΜΚΟ), θα μπορούσαν να συμπράξουν με τη βοήθεια ειδικών του χώρου ώστε να συνεισφέρουν με την εμπειρία τους αλλά παράλληλα να ενισχύσουν και το απαιτούμενο πνεύμα συναίνεσης.

Η προαναφερθείσα προσέγγιση, θεωρητικά τουλάχιστον, αρχικά δημιουργεί πρόσφορο έδαφος για την υιοθέτηση κατευθυντήριων γραμμών που εξυπηρετούν όλους τους εμπλεκόμενους φορείς, αλλά δεύτερον και πιο σημαντικό, επιτρέπει τον άμεσο εντοπισμό κινδύνων και τη διαχείριση αυτών σε ένα πνεύμα συναίνεσης. Σημειώνεται επίσης, πως η επιλογή του υβριδικού μοντέλου διακυβέρνησης επιτρέπει στις κυβερνήσεις να εντοπίσουν σχετικά γρήγορα ακόμη και ζητήματα που δεν έχουν υποπέσει στην αντίληψή τους εκ των προτέρων, είτε λόγω της εξελικτικής δυνατότητας των αλγορίθμων μηχανικής μάθησης ή ακόμα και λόγω έλλειψης πληροφοριών σε κάποιο πεδίο εφαρμογής αυτών (Taeihagh, 2021). Η ευελιξία, η προσαρμοστικότητα και η σύμπραξη πολλαπλών φορέων είναι τα κύρια συστατικά επιτυχίας ενός τέτοιου μοντέλου διακυβέρνησης ώστε να επιτευχθεί η απαραίτητη συναίνεση για τον καταρτισμό προτύπων σχεδιασμού συστημάτων τεχνητής νοημοσύνης, αλλά, και να οικοδομηθεί ένα βιώσιμο μοντέλο διαχείρισης της τεχνολογίας αυτής.

Επίσης, με την επιλογή ενός τέτοιου μοντέλου διακυβέρνησης, θα μπορούσε να αναπτυχθούν και συγκεκριμένες τεχνικές διαχείρισης. Για παράδειγμα, πραγματοποιώντας καινοτόμους νομικούς πειραματισμούς και θέτοντας δοκιμαστικές ρυθμίσεις σε πολύ συγκεκριμένες περιπτώσεις ή τομείς και μέσω των επαναληπτικών ελέγχων, να προσαρμοστούν επιτυχώς οι γενικές κατευθυντήριες γραμμές, τα ρυθμιστικά πλαίσια ώστε να επιτευχθεί η απαραίτητη ευθύνη και λογοδοσία της τεχνητής νοημοσύνης (Taeihagh, 2021; Linkov et al., 2018).

Όμως, τα παραπάνω ίσως να μην είναι εφικτά σε τοπικό επίπεδο, και στις περισσότερες χώρες, ούτε καν σε εθνικό επίπεδο, για δυο βασικούς λόγους. Πρώτον, το διαδίκτυο έχει επιβάλλει ένα παγκοσμιοποιημένο πλαίσιο δραστηριοποίησης της κοινωνίας, άρα και αντίστοιχο περιβάλλον δεδομένων. Δεύτερον, οι ιδιωτικοί φορείς με πραγματικά βαθιά κατανόηση της τεχνητής νοημοσύνης που θα μπορούσαν να συνεισφέρουν ουσιαστικά στην διακυβέρνηση της τεχνητής νοημοσύνης, είναι συγκεκριμένοι, ελάχιστοι σε αριθμό, με παγκόσμια εμβέλεια δράσης και με τεράστια επιρροή σε πολλαπλούς τομείς. Ως εκ τούτου, η επιλογή της προαναφερθείσας υβριδικής διακυβέρνησης, προϋποθέτει μια εξωστρεφή προσέγγιση σε υπερεθνικό (π.χ. Ευρωπαϊκή Ένωση) ή σε παγκόσμιο επίπεδο, ειδικά για τις μικρές και αναπτυσσόμενες χώρες. Αυτό είναι αναγκαίο για τις μικρές και αναπτυσσόμενες χώρες, όμως ίσως αφήνει αδιάφορους τους ιδιωτικούς παράγοντες.

Όμως υπάρχουν τομείς όπως και οι χώρες αυτές συμμετέχουν στο παγκόσμιο στερέωμα, όπως οι αεροπορικές υπηρεσίες και μεταφορές, η ναυτιλία, η προμήθεια πρώτων υλών και προϊόντων. Άρα, θα μπορούσαν μέσω των δράσεων τους στη παγκόσμια εφοδιαστική αλυσίδα και την έκθεσή τους, να υιοθετήσουν τόσο αντίστοιχες τεχνολογίες, όσο και να αποκομίσουν τεχνογνωσία για την αντίστοιχη ρύθμιση του πλαισίου. Ιδανικά, θα μπορούσε να δημιουργηθεί ένα παγκόσμιο πλαίσιο διακυβέρνησης της τεχνητής νοημοσύνης ευθυγραμμισμένη με τα ανθρώπινα δικαιώματα και τις δημοκρατικές αξίες (Gasser & Almeida, 2017), ανάλογο των Ηνωμένων Εθνών. Μια τέτοια πρωτοβουλία θα μπορούσε να θέσει το βασικό πλαίσιο διακυβέρνησης της εν λόγω τεχνολογίας σεβόμενη τα κατά τόπους πολιτιστικά, πολιτικά και κοινωνικά χαρακτηριστικά.

Η υβριδική διακυβέρνηση που περιγράφεται παραπάνω, θα μπορούσε να εστιάσει σε τρία βασικά επίπεδα σύμφωνα με τους Gasser και Almeida (2017). Πρώτον, στο τεχνικό

επίπεδο, ώστε να εξεταστούν πρακτικά πως οι ηθικές αρχές, τα θεμελιώδη δικαιώματα των υποκειμένων και οι νομικές ρυθμίσεις μπορούν να μεταφραστούν σε τεχνικές προδιαγραφές. Για παράδειγμα, σε περιπτώσεις προληπτικής αστυνόμευσης με χρήση ενός συστήματος τεχνητής νοημοσύνης, θα μπορούσε να επιβληθεί να θολώνει τα πρόσωπα των ατόμων ώστε να αποφεύγονται μεροληπτικές κατηγοριοποιήσεις κοινωνικών ομάδων και μόνο σε εξαιρετικές περιπτώσεις να επιτρέπεται η εξουσιοδοτημένη πρόσβαση σε δεδομένα. Δεύτερον, αυτού του τύπου διακυβέρνηση, θα μπορούσε να εστιάσει στο ηθικό σχεδιασμό των εν λόγω συστημάτων. Όπως αναφέρθηκε και παραπάνω, το κλειδί είναι η στενή παρακολούθηση και η ερμηνεία των αποτελεσμάτων χρήσης ώστε να επιτυγχάνεται η επανηπτική προσαρμογή των κανόνων ηθικής αντίστοιχα.

Τρίτον, και δεδομένων των δυο παραπάνω επιπέδων (τεχνικό και ηθικού σχεδιασμού) και των αποτελεσμάτων χρήσης της τεχνητής νοημοσύνης, θα πρέπει να υπάρχει ένα επίπεδο συνεχούς ανάλυσης αυτών, κυρίως όσον αφορά τις κοινωνικές επιπτώσεις. Σύμφωνα με τον Rahwan (2018), στην ιδιάζουσα περίπτωση της τεχνητής νοημοσύνης, θα πρέπει η προσέγγιση “human in the loop” (εμπλοκή των υποκειμένων), να επεκταθεί σε ευρύτερο επίπεδο αγγίζοντας μια προσέγγιση τύπου “society in the loop” (ενεργή εμπλοκή της κοινωνίας). Δηλαδή, η κοινωνική συναίνεση και η σύμπραξη όλων των παραγόντων στον σχεδιασμό και τη διακυβέρνηση της εν λόγω τεχνολογίας είναι όχι απλά επιθυμητή αλλά αναγκαία, τόσο για το πλαίσιο σχεδιασμού και λειτουργίας, όσο και για την κατανομή οφέλους και ευθύνης.

Παρόλα αυτά, η εμπειρική γνώση είναι ελλιπής, και όπως επισημαίνεται σθεναρά από τους ερευνητές, θα πρέπει άμεσα να εξεταστούν σε μεγαλύτερο βάθος τα πλαίσια μιας υβριδικής διακυβέρνησης της τεχνητής νοημοσύνης με παγκόσμια και προσαρμοστικά χαρακτηριστικά αλλά και η δυνατότητα πρακτικής εφαρμογής αυτής (Wirtz et al., 2020).

6 Εθνικές Στρατηγικές για τη Τεχνητή Νοημοσύνη

Λαμβάνοντας υπόψιν τα παραπάνω, είναι αναγκαίο να περιγραφεί το που βρισκόμαστε σήμερα. Από τη μια, και σύμφωνα με την Ευρωπαϊκή επιτροπή (2018), η τεχνητή νοημοσύνη εφαρμόζεται ήδη σε πάρα πολλές δραστηριότητες και εκφάνσεις της κοινωνίας, και μάλιστα, σε όλα σχεδόν τα κυβερνητικά οικοσυστήματα παγκοσμίως. Από την άλλη, σε επίπεδο διακυβέρνησης, ήδη έχουν ξεκινήσει πρωτοβουλίες σε όλα τα επίπεδα χωρίς όμως να δύναται να εξαχθούν ασφαλή συμπεράσματα. Αναλυτικότερα, και σύμφωνα με τη μελέτη του Radu (2021), τα κράτη δεν έχουν επιβάλει ακόμα αυστηρούς περιορισμούς στην καινοτομία που σχετίζεται με τη τεχνητή νοημοσύνη. Παρόλα αυτά, ήδη αρκετές χώρες έχουν καταρτίσει τις εθνικές τους στρατηγικές ενώ εμπλέκονται και σε διεθνής προσπάθειες καταρτισμού υπερθνικών σχεδίων για την διακυβέρνηση της εν λόγω τεχνολογίας.

Έως το 2021, περισσότερα από τριάντα (30) χώρες βρίσκονταν σε διαδικασία καταρτισμού της εθνικής τους στρατηγικής ενώ σχεδόν είκοσι (20) είχαν δημοσιεύσει τη τελική τους πρόταση και η εφαρμογή τους βρίσκεται ήδη σε εξέλιξη. Η πρώτη χώρα που ανακοίνωσε την ολοκληρωμένη στρατηγική της για τη διακυβέρνηση της τεχνητής νοημοσύνης ήταν η Νότια Κορέα, ενώ αμέσως μετά ακολούθησε ο Καναδάς. Έκτοτε, αρκετές ευρωπαϊκές χώρες (Γαλλία, Γερμανία) και οι ηγέτιδες τεχνολογικά δυνάμεις (ΗΠΑ, Ηνωμένο Βασίλειο, Κίνα, Ιαπωνία, Σιγκαπούρη) δημοσίευσαν τις αντίστοιχες στρατηγικές τους οι οποίες αντικατροπτίζουν και τις φιλοδοξίες τους στον τομέα αυτό.

Σημειώνεται πως οι στρατηγικές αυτές δεν εμβαθύνουν στα κρίσιμα ζητήματα και της προκλήσεις της τεχνητής νοημοσύνης. Περισσότερο εστιάζουν στον καθορισμό της γενικής φιλοδοξίας τους, παρέχοντας κάποιες κατευθυντήριες γραμμές, ενώ άλλες, θέτουν συγκεκριμένες προτεραιότητες και καθορίζουν τη χρηματοδότηση για την επίτευξη των στόχων τους στο άμεσο και μεσοπρόθεσμο μέλλον (Dutton, 2018.) Όμως σε όλες τις περιπτώσεις, οι στρατηγικές που κατατέθηκαν επαναφέρουν τις κυβερνήσεις ως κεντρικό παράγοντα σε ένα τομέα που μέχρι πρόσφατα κυριαρχούσε η ιδιωτική πρωτοβουλία. Πλην όμως των περιπτώσεων της Φιλανδίας, της Ιαπωνίας και της Νοτίου Κορέας, οι υπόλοιπες χώρες δεν εξετάζουν επαρκώς τα όρια διάκρισης μεταξύ των συμφερόντων του δημοσίου και του ιδιωτικού τομέα.

Επίσης, η επισκόπηση των στρατηγικών αυτών αναδεικνύει, σύμφωνα με τον Radu (2018), τη ποικιλομορφία των προσεγγίσεων για τη διακυβέρνηση της τεχνητής νοημοσύνης, επιβεβαιώνει σε αρκετά σημεία την απροσδιοριστία της διαχείρισης της, ενώ σημειώνεται η έντονη τάση εμπορευματοποίησής της σε μια κούρσα ανταγωνισμού για την πρωτοκαθεδρία στο τομέα αυτό. Δυστυχώς, ο παγκόσμιος ανταγωνισμός για τον έλεγχο και την υπεροπλία στη τεχνητή νοημοσύνη ανάμεσα στις υπερδυνάμεις των ΗΠΑ, Κίνας και Ρωσίας αντανακλάται τόσο στις επιλεχθείσες από τις χώρες αυτές στρατηγικές, ενώ εύλογα διαπιστώνεται και η πολιτικοποίησή της διακυβέρνησής της. Μάλιστα, οι προαναφερθείσες υπερδυνάμεις είναι οι μόνες, μαζί με τη Γαλλία, που ξεκάθαρα θέτουν στρατιωτικούς στόχους συναφείς με τη τεχνητή νοημοσύνη στις στρατηγικές τους.

Βέβαια, υπάρχουν και αρκετά πιο ψύχραιμες τοποθετήσεις, όπως του Καναδά, για την οικοδόμηση μιας κοινής και παγκόσμιας συναίνεσης τουλάχιστον για τις βασικές πρακτικές και τα επιθυμητά πρότυπα που θα διέπουν τις εφαρμογές της μηχανικής μάθησης και της τεχνητής νοημοσύνης γενικότερα. Σε αυτή τη κατεύθυνση, ήδη η Ευρωπαϊκή Ένωση και ο ΟΑΣΑ έχουν προτείνει αντίστοιχες πρωτοβουλίες και κατευθυντήριες γραμμές, όμως προς το παρόν τα αποτελέσματα είναι πενιχρά.

Αντίστοιχα, η Γαλλία στην εθνική στρατηγική που δημοσίευσε στοχεύει συγκεκριμένα σε τέσσερις (4) σημαντικούς για αυτούς άξονες δημοσίου ενδιαφέροντος (υγεία, μεταφορές, περιβάλλον, άμυνα) μεταθέτοντας την δημιουργία πλαισίου για τις υπόλοιπους τομείς σε δεύτερο χρόνο ή στον ιδιωτικό τομέα να αναπτύξει πρωτοβουλίες. Επίσης, πρότείνει και την δημιουργία ενός CERN για την τεχνητή νοημοσύνη σε Ευρωπαϊκό επίπεδο τουλάχιστον με την αντίστοιχη υποστήριξη της Γερμανίας, δεδομένης της πρόθεσής τους για δημιουργία ενός ηθικού πλαισίου χρήσης της εν λόγω τεχνολογίας. Η Γερμανία με τη δική της εθνική στρατηγική στοχεύει στη διατήρηση της τεχνολογικής της υπεροπλίας, ενώ παράλληλα επιδιώκει μια προσέγγιση της ηθικής ήδη από το σχεδιασμό για τη τεχνητή νοημοσύνη.

Συνολικά όλες οι εθνικές στρατηγικές, πλην ΗΠΑ και Κίνας, θέτουν ως προτεραιότητα την ηθική στη τεχνητή νοημοσύνη και στη δημιουργία ενός πλαισίου καθοδήγησης των εξελίξεων μέσα από ένα ηθικό προσανατολισμό, χωρίς όμως να το εξειδικεύουν

ιδιαίτερα. Σε όλες όμως σχεδόν τις στρατηγικές, οι οργανισμοί για τα δικαιώματα, μη κυβερνητικοί οργανισμοί, οι τελικές χρήστες και γενικά οι αποδέκτες της τεχνολογίας δεν συμπεριλαμβάνονται ως ενεργοί συμμετέχοντες, τουλάχιστον προς το παρόν. Γενικά, η επισκόπηση των εθνικών στρατηγικών αντικατροπτίζει τις φιλοδοξίες των κρατών αυτών για να ηγηθούν στις εφαρμογές της τεχνητής νοημοσύνης παρά περιγράφουν τις απαντήσεις τους στις προκλήσεις διακυβέρνησης της. Είναι σαφές επίσης, πως οι έντονες διαφοροποιήσεις των εθνικών στρατηγικών συνδέονται άμεσα και με τη πολυπλοκότητα και την εξελικτική δυνατότητα της εν λόγω τεχνολογίας πέρα των πολιτικών επιλογών της κάθε χώρας.

7 Από την Στρατηγική στη Ρύθμιση της Τεχνητής Νοημοσύνης

Πιο πρόσφατα, και ως συνέχεια της δημοσίευσης των εθνικών τους στρατηγικών, τόσο η Ευρωπαϊκή Ένωση, όσο και οι ΗΠΑ, κατέθεσαν τα νομοσχέδια τους για τη διακυβέρνηση και ρύθμιση της τεχνητής νοημοσύνης. Ομοίως, προχωρούν και οι υπόλοιπες μη-Ευρωπαϊκές χώρες, μιας και τα κράτη μέλη της Ευρώπης καλύπτονται από τη νομοθεσία της ένωσης. Στην ενότητα αυτή, ο στόχος είναι η σύντομη και ενδεικτική επισκόπηση των νομοσχεδίων ώστε να παρέχεται βαθύτερη κατανόηση των προσπαθειών διακυβέρνησης και ρύθμισης της εν λόγω τεχνολογίας.

Αναλυτικότερα, τον Απρίλιο του 2021, η Ευρωπαϊκή Επιτροπή επικύρωσε το σχέδιο κανονισμού που επιχειρεί να θέσει τα πρότυπα αξιόπιστης και υπεύθυνης ανάπτυξης και λειτουργίας συστημάτων τεχνητής νοημοσύνης στην αγορά της ένωσης. Το προσχέδιο νόμου της ένωσης εισάγει ένα καινοτόμο καθεστώς ασφάλειας προϊόντων κατηγοριοποιώντας τα βάσει κινδύνου. Σκοπός είναι η διασφάλιση πως οι εφαρμογές / προϊόντα τεχνητής νοημοσύνης που είτε αναπτύσσονται, ή χρησιμοποιούνται, ή και τα δυο, είναι ασφαλή και σύμφωνα με τις αξίες που προσβέυει η Ευρωπαϊκή Ένωση.

Επιπρόσθετοι στόχοι που θέτει η Ένωση, είναι η διευκόλυνση των επενδύσεων και η δημιουργία ενός περιβάλλοντος που ενισχύει την καινοτομία, τη διακυβέρνηση της τεχνολογίας αυτής, ενώ παράλληλα θα διασφαλίζεται η δυνατότητα της επιβολής της αντίστοιχης νομοθεσίας, στο πλαίσιο μια ενιαίας και ασφαλούς αγοράς.

Σύμφωνα με το νομοσχέδιο, λαμβάνεται υπόψιν η κρισιμότητα της εφαρμογής του συστήματος της τεχνητής νοημοσύνης, κατηγοριοποιημένη σε τέσσερα επίπεδα, και αντιστοίχως προβλέπεται ένα πλαίσιο διαχείρισης. Πρακτικά αυτό σημαίνει πως οι εφαρμογές τεχνητής νοημοσύνης με αμελητέο ή χαμηλό κίνδυνο διέπονται από χαλαρότερο νομικό καθεστώς (π.χ. μη δεσμευτική συμμόρφωση και αυτοαξιολόγηση), ενώ σε περίπτωση υψηλού κινδύνου τα μέτρα ελέγχου και επιβολής είναι αυστηρότερα. Σε περιπτώσεις ύψιστου κινδύνου η σχετική νομοθεσία επιβάλλει ακόμα και απαγόρευση χρήσης αν κριθεί αναγκαίο. Γενικά, στη περίπτωση της Ένωσης και ανάλογα με τη κρισιμότητα, εφαρμόζεται πλέον ένα πλαίσιο κανονισμών που εκτείνεται από το φάσμα της αυτορρύθμισης με αξιολόγηση του αντικτύπου χρήσης της εν λόγω τεχνολογίας έως τη αυστηρή και ελεγχόμενη αξιολόγηση συμμόρφωσης σε όλη της διάρκεια του κύκλου ζωής ενός τέτοιου συστήματος από εξωτερικό οργανισμό.

Όσον αφορά τις εφαρμογές υψηλού κινδύνου και σύμφωνα με τον νέο Ευρωπαϊκό Κανονισμό, αυτές αναφέρονται σε κρίσιμες υποδομές (π.χ. υγεία), την εκπαίδευση και τις αυτοματοποιημένες αξιολογήσεις, τις δημόσιες υπηρεσίες, τα συστήματα επιτήρησης και επιβολής του νόμου, την απονομή δικαιοσύνης και τη διαχείριση του μετανεστευτικού. Ειδικότερα, σε αυτή τη κατηγορία υψηλού κινδύνου και σε περίπτωση που επιτραπεί η χρήση ενός τέτοιου συστήματος ύστερα από ενδεδειχτή έλεγχο, θα υπογράφεται και θα κοινοποιείται η αντίστοιχη ευρωπαϊκή σήμανση, ενώ μηχανισμοί διαφάνειας, συνεχούς παρακολούθησης και ελέγχου θα εφαρμόζονται καθόλη τη διάρκεια χρήσης με απαίτηση φυσικής επίβλεψης από φυσικά πρόσωπα.

Σημειώνεται όμως, πως τα πρότυπα συμμόρφωσης δεν έχουν αποσαφηνιστεί λόγω των ρευστών κοινωνικοπολιτικών δεδομένων και επαφίεται στα αρμόδια όργανα να ερμηνεύσουν κατά περίπτωση τον αντίστοιχο καταρτισμό προτύπων. Μάλιστα προβλέπεται η σύσταση ενός ενιαίου Ευρωπαϊκού Συμβουλίου Τεχνητής Νοημοσύνης (EAIB), το οποίο ως φορέας και σε συνεργασία με τις εθνικές αντίστοιχες εποπτικές αρχές που θα συσταθούν, και θα έχει την αρμοδιότητα ελέγχου και επιβολής των ρυθμίσεων ανά περίπτωση. Βέβαια, ο κανονισμός διασφαλίζει εν μέρει τη δυνατότητα καινοτομίας και ευελιξίας επιτρέποντας τη χρήση νομικών πειραματισμών σε ειδικές περιπτώσεις, όπως ερευνητικά κέντρα και μικρομεσαίες επιχειρήσεις.

Επιπρόσθετα, ο Ευρωπαϊκός Κανονισμός θέτει στόχους βιωσιμότητας της χρήσης της τεχνητής νοημοσύνης, επισημαίνει τη υποχρέωση παράλληλης συμμόρφωσης και με τα υφιστάμενα κανονιστικά πλαίσια (π.χ. προσωπικά δεδομένα – GDPR, ιατροτεχνολογία – MDR, κτλ.), ενώ εισάγει και τη κοινή και ασφαλή χρήση δεδομένων.

Τέλος, είναι σαφές πως ο κανονισμός έχει άμεση ισχύ για τα κράτη μέλη όμως επισημαίνεται σθεναρά η ανάγκη για μια παγκοσμιοποιημένη προσέγγιση στο τομέα της διακυβέρνησης της τεχνητής νοημοσύνης. Εξού και η αναφορά στην ανάγκη για εξωστρεφή και συνεργατική προσέγγιση της διακυβέρνησης μεταξύ κρατών με συναφείς δημοκρατικές καταβολές και αξίες αλλά και αντίστοιχες κοινές προσεγγίσεις στη ψηφιοποίηση της κοινωνίας. Είναι δεδομένη η δεσπίζουσα θέση της Ευρωπαϊκής Ένωσης στην παγκόσμια οικονομία και ως εκ τούτου, η συμμόρφωση με τον κανονισμό αποκτά και νομική ισχύ πέρα των στενών γεωγραφικών γραμμών της Ένωσης (Kop, 2021).

Στην αντίπερα όχθη του Ατλαντικού, στις ΗΠΑ, το νέο νομοσχέδιο για τη ρύθμιση της τεχνητής νοημοσύνης κατατέθηκε το Φεβρουάριο του 2022 προς κύρωση στη Γερουσία και τη Βουλή των Αντιπροσώπων χωρίς μέχρι σήμερα να έχει επικυρωθεί. Το νέο αυτό νομοσχέδιο είναι συνέχεια της υφιστάμενης νομοθεσίας του 2019 για την αλγοριθμική λογοδοσία και υιοθετεί μια προσέγγιση ρύθμισης σε υψηλό επίπεδο. Πάραυτα, τα κύρια σημεία του νομοσχεδίου δεν έχουν ως στόχο την απαγόρευση ή τον περιορισμό της εν λόγω τεχνολογίας. Αντιθέτως, στοχεύει στη δημιουργία μιας υποδομής διακυβέρνησης ώστε από τη μια πλευρά να αποτραπούν ή να λογοδοτήσουν κακοπροαίρετες προσπάθειες χρήσης αλλά από την άλλη, να ενθαρρυνθούν φορείς με καλές προθέσεις οι οποίοι αποδεικνύουν τις ηθικές, ασφαλείς και σύννομες εφαρμογές τέτοιου τύπου. Στη περίπτωση των ΗΠΑ, τόσο οι εκ των προτέρων αλλά και κατά τη διάρκεια του κύκλου ζωής, οι αξιολογήσεις των συστημάτων τεχνητής νοημοσύνης είναι ευπρόσδεκτη ώστε να δημιουργηθεί το πλαίσιο τεκμηρίωσης και μεταγενέστερης πιθανόν ρύθμισης ανά περίπτωση. Γενικά, και λαμβάνοντας υπόχιν την πολιτική και νομοθετική κουλτούρα των ΗΠΑ, το νομοσχέδιο απεικονίζει τη προσπάθεια εξισορρόπησης της ρύθμισης έναντι της καινοτομίας χωρίς να εμβαθύνει ιδιαίτερος σε αρκετά θολά σημεία που προκύπτουν από τη χρήση της εν λόγω τεχνολογίας.

Παρόλα αυτά, εισάγει μια πρωτοποριακή προσέγγιση σύγκρισης της υφιστάμενης διαδικασίας λήψης κρίσιμων αποφάσεων κατά αντιπαράσταση με την αλγοριθμική, ώστε να είναι δυνατή η κατανόηση των αποτελεσμάτων και να εντοπίζονται σημεία που χρίζουν περαιτέρω αξιολόγηση, εποπτεία ή και ρύθμιση.

Από την ανάλυση προκύπτει πως τα κύρια σημεία που διαφοροποιούνται ο κανονισμός της Ευρωπαϊκής Ένωσης και το αντίστοιχο νομοσχέδιο των ΗΠΑ, είναι πως στις ΗΠΑ χρησιμοποιείται ο όρος αυτοματοποιημένα συστήματα αποφάσεων (Automated Decision-Making Systems - ADS), σε σχέση με το γενικευμένο όρο της τεχνητής νοημοσύνης που χρησιμοποιείται στον Ευρωπαϊκό Κανονισμό, ενώ χρησιμοποιείται παρομοίως στη βιβλιογραφία. Αυτός ο ορισμός αντικατροπτίζει και συμπεριλαμβάνει όχι μόνο υφιστάμενες αντιλήψεις για το τι μπορεί να θεωρηθεί τεχνητή νοημοσύνη αλλά είναι τεχνολογικά πιο ουδέτερος με δυνατότητα μελλοντικά να συμπεριλαμβάνει καινοτόμες και άγνωστες προς το παρόν εφαρμογές αυτοματοποιημένης λήψης κρίσιμων αποφάσεων. Επίσης, στις ΗΠΑ, ο ορισμός όπως αναφέρθηκε παραπάνω επιτρέπει τη δυνατότητα ελέγχου και ρύθμισης γενικότερα, ακόμα και των περιπτώσεων όπου τέτοια συστήματα δεν λειτουργούν απολύτως αυτόνομα αλλά χρησιμοποιούνται για να υποστηρίξουν αποφάσεις. Το γεγονός αυτό ενισχύει τη δυνατότητα δημόσιας εποπτείας σε διαδικασίες λήψης κρίσιμων αποφάσεων που αντίστοιχα δεν προβλέπεται από τον Ευρωπαϊκό Κανονισμό.

Επιπρόσθετα, το νομοσχέδιο των ΗΠΑ, απαιτεί τη σύγκριση των αποτελεσμάτων από τη χρήση ενός τέτοιου συστήματος σε σχέση με τα αποτελέσματα στη προηγούμενη διαδικασία λήψης αποφάσεων, ώστε να γίνονται ευκολότερα κατανοητά τα οφέλη αλλά και οι κίνδυνοι ενώ ενισχύεται και η διαφάνεια. Βέβαια, οι ΗΠΑ, εξαιρούν τις αποφάσεις που λαμβάνονται από τη κυβέρνηση με την υποστήριξη ή εξ ολοκλήρου αλγορίθμων και εστιάζουν στις μεγάλες εταιρίες (π.χ. >\$50 εκ. τζίρος, 1 εκ. χρήστες, κτλ.), ενώ μάλλον ορθώς εξαιρεί τις μικρομεσαίες επιχειρήσεις λόγω υψηλού κόστους συμμόρφωσης και ενίσχυση της καινοτομίας.

Τέλος, είναι γεγονός πως η φιλοδοξία του νομοσχεδίου των ΗΠΑ είναι η διασφάλιση της ισότητας των υποκειμένων από την αυτοματοποιημένη λήψη αποφάσεων με συνεχή αξιολόγηση και παρακολούθηση, αλλά αυτό είναι αρκετά γενικό και πιθανόν μη

εφαρμοστέο. Επίσης, είναι λιγότερο συγκεκριμένο από τον κανονισμό της Ευρωπαϊκής Ένωσης, ενώ η επανάληψη του όρου “στο βαθμό του εφικτού” υπονομεύει την ίδια την προταθείσα νομοθεσία. Αντίθετα όμως, έχει μια πιο γενική και ρεαλιστική προσέγγιση στη νομική διαχείριση και αντιμετώπιση των ηθικών προκλήσεων σε σχέση με τον Ευρωπαϊκό Κανονισμό, ο οποίος έχει πιο μακροπρόθεσμη και ολιστική προσέγγιση για το νέο υπό διαμόρφωση πλαίσιο λειτουργίας της ψηφιακής και αλγοριθμικής διακυβέρνησης της Ένωσης (Mökander, 2022).

Η επισκόπηση τόσο του Ευρωπαϊκού Κανονισμού, όσο και του αντίστοιχου υπό επικύρωση νομοσχεδίου των ΗΠΑ, αναδεικνύει τη πρώτη προσπάθεια διακυβέρνησης και δημιουργίας ενός πλαισίου, είτε γενικού ή πιο στοχευμένου, σε ένα συνεχώς μεταβαλλόμενο και πολύπλοκο τομέα. Και στις δυο περιπτώσεις, αναφέρεται η ανάγκη για παγκόσμιες πρωτοβουλίες συνεργασίας και διακυβέρνησης μέσω της τεχνολογικής διπλωματίας, καθώς αναγνωρίζονται οι αντίστοιχες γεωγραφικές προκλήσεις, αλλά και η συνεργατική διαχείριση που απαιτείται για τη διασφάλιση της ελάχιστης λογοδοσίας.

8 Ρύθμιση της Τεχνητής Νοημοσύνης στην Ελλάδα - Ν. 4961/2022

Σε εθνικό επίπεδο και όσον αφορά τη προσπάθεια ρύθμισης της τεχνητής νοημοσύνης στην Ελλάδα, προσφάτως και συγκεκριμένα στον Ιούλιο του 2022, δημοσιεύθηκε στο αντίστοιχο ΦΕΚ ο νόμος Ν. 4961/2022 για τις αναδυόμενες τεχνολογίες πληροφορικής και επικοινωνιών, την ενίσχυση της ψηφιακής διακυβέρνησης και άλλες διατάξεις.

Ο πρόσφατος αυτός νόμος θεσπίζει ένα συνεκτικό νομοθετικό πλαίσιο για την τεχνητή νοημοσύνη, το διαδίκτυο των πραγμάτων, την παροχή ταχυδρομικών υπηρεσιών με τη χρήση Συστημάτων μη Επανδρωμένων Αεροσκαφών, τις τεχνολογίες καταναμημένου καθολικού (blockchain) και τα έξυπνα συμβόλαια καθώς και τις τρισδιάστατες εκτυπώσεις. Στην παρούσα εργασία μας απασχολεί το πρώτο σκέλος για τη τεχνητή νοημοσύνη. Οι ρυθμίσεις του νέου Νόμου έχουν άμεσο αντίκτυπο στον ψηφιακό μετασχηματισμό του δημόσιου και στους τομείς της τεχνολογίας και της ψηφιακής οικονομίας.

Αναλυτικότερα και αναμένοντας την υιοθέτηση της Πράξης για την Τεχνητή Νοημοσύνη από την Ευρωπαϊκή Ένωση, ο Ν. 4961/2022 εισάγει το εθνικό πλαίσιο για τη ρύθμιση της χρήσης τεχνολογιών τεχνητής νοημοσύνης στον δημόσιο και ιδιωτικό τομέα.

Όσον αφορά τους Δημοσίους Φορείς, και εξαιρώντας τα Υπουργεία Εθνικής Άμυνας και Προστασίας του Πολίτη για ευνόητους λόγους, η χρήση συστημάτων τεχνητής νοημοσύνης επιτρέπεται μόνο εφόσον προβλέπεται από ειδική διάταξη νόμου, που περιλαμβάνει κατάλληλες εγγυήσεις για την προστασία των δικαιωμάτων των φυσικών ή νομικών προσώπων, που επηρεάζονται από τα συστήματα αυτά. Ειδικότερα, θα πρέπει πριν τη χρήση ενός συστήματος τεχνητής νοημοσύνης, πέρα από την εκτέλεση εκτίμησης αντικτύπου του Κανονισμού (ΕΕ) 2016/679 («ΓΚΠΔ»), κάθε φορέας έχει την υποχρέωση να εκπονεί και αλγοριθμική εκτίμηση αντικτύπου για την αξιολόγηση των κινδύνων που ενδέχεται να προκύπτουν για τα δικαιώματα, τις ελευθερίες και τα έννομα συμφέροντα των προσώπων, που επηρεάζονται από το σύστημα αυτό. Επιπρόσθετα με Προεδρικό Διάταγμα εξειδικεύονται οι κατάλληλες εγγυήσεις για την προστασία των δικαιωμάτων των προσώπων, που επηρεάζονται από τη χρήση συστημάτων τεχνητής νοημοσύνης. Επίσης, κάθε δημόσιος φορέας θα πρέπει να παρέχει δημόσια πληροφορίες, μεταξύ άλλων, για τον χρόνο έναρξης και τις παραμέτρους λειτουργίας του συστήματος τεχνητής νοημοσύνης καθώς και για τις αποφάσεις που λαμβάνονται ή υποστηρίζονται μέσω αυτού. Τυχόν καταγγελίες για παραβιάσεις των υποχρεώσεων διαφάνειας εξετάζονται από την Εθνική Αρχή Διαφάνειας αντιστοίχως. Τέλος, κάθε δημόσιος φορέας οφείλει να τηρεί μητρώο με τα συστήματα τεχνητής νοημοσύνης που χρησιμοποιεί (*Μητρώο Συστημάτων ΤΝ*).

Όσον αφορά τους ιδιωτικούς φορείς, ο αντίστοιχος νόμος επιχειρεί να θωρακίσει την ορθή χρήση της τεχνητής νοημοσύνης στο πεδίο της απασχόλησης και της πρόσληψης / αξιολόγησης των εργαζομένων, επιτάσσει τη δεοντολογική χρήση δεδομένων, ενώ εισάγει και την υιοθέτηση αντίστοιχου μητρώου συστημάτων τεχνητής νοημοσύνης. Επιπρόσθετα, ο νόμος 4961/2022 ορίζει και ένα σύνολο από υποχρεώσεις των ιδιωτικών φορέων που αναλαμβάνουν να εκτελέσουν δημόσιες συμβάσεις που συμπεριλαμβάνουν την ανάπτυξη και προμήθεια συστημάτων τεχνητής νοημοσύνης.

Σημειώνεται δε, πως ο εν λόγω νόμος, επιτάσσει την σύσταση μιας συντονιστικής επιτροπής για την τεχνητή νοημοσύνη με κύριες αρμοδιότητες την κατάρτιση της Εθνικής Στρατηγικής για την τεχνητή νοημοσύνη και, γενικότερα, την χάραξη πολιτικής γύρω από αυτή, και, αφετέρου, συστήνει τη δημιουργία μια ειδικής επιτροπής για την εποπτεία της στρατηγικής, που μεριμνά για την υλοποίηση, τον συντονισμό των αρμοδίων φορέων και την μέριμνα για την εφαρμογή της. Για την επιτέλεση του έργου τους οι δύο επιτροπές τροφοδοτούνται με στοιχεία από το Παρατηρητήριο Τεχνητής Νοημοσύνης, που παρακολουθεί και αποτυπώνει σε εκθέσεις τις τεχνολογικές εξελίξεις και τις πολιτικές γύρω από την ΤΝ στην χώρα και σε διεθνές επίπεδο (Μήτσιος, 2022).

Όπως προκύπτει από τα παραπάνω, ο νόμος 4961/2022 αναμένεται να συνδράμει στην ενίσχυση του ψηφιακού μετασχηματισμού του δημόσιου και ιδιωτικού τομέα της χώρας. Τέλος, είναι σαφές πως και σε εθνικό επίπεδο γίνεται προσπάθεια εναρμόνισης με την Ευρωπαϊκή νομοθεσία και αναμένεται άμεσα να φανούν τα πρώτα αποτελέσματα της εν λόγω ρύθμισης καθώς και στην Ελλάδα, η χρήση της τεχνητής νοημοσύνης στη δημόσια διοίκηση αλλά κυρίως στον ιδιωτικό τομέα, αυξάνεται συνεχώς.

9 Νομικοί Προβληματισμοί από τη χρήση της Τεχνητής Νοημοσύνης

Είναι πασιφανές πως από τη χρήση της τεχνητής νοημοσύνης, ειδικότερα στη λήψη δημοσίων αποφάσεων, πέρα από τα πιθανά οφέλη, εγείρονται αρκετά ζητήματα τα οποία δεν δύναται να προβλεφθούν και ρυθμιστούν νομικά εκ των προτέρων. Αόριστες νομικές έννοιες όπως “ δημόσιο συμφέρον”, “σπουδαίος λόγος”, “σοβαρός κίνδυνος”, “αναγκαία μέτρα”, “επείγουσα και απρόβλεπτη ανάγκη”, “ταχεία ενέργεια”, κτλ., είναι έννοιες οι οποίες δεν μπορούν να εφαρμοστούν, εάν δεν καθοριστούν ή/και εξειδικευτούν σε συγκεκριμένη περίπτωση (Δαγτόγλου, 2014).

Ειδικότερα, και όπως αναφέρθηκε και παραπάνω, η νομική ερμηνεία αποτελεί το κλειδί στην αντιμετώπιση τυχόν προβλημάτων αυτοματοποιημένης λήψης αποφάσεων, όπου η εφαρμογή των σχετικών νόμων έρχεται αντιμέτωπη με τη “ζώσα διοικητική πρακτική” (Etscheid, 2019).

Ως εκ τούτου, προτείνεται σθεναρά, και από το υφιστάμενο νομικό πλαίσιο αλλά και από τους μελετητές, οι εφαρμογές της τεχνητής νοημοσύνης, ειδικά στη δημόσια σφαίρα, να μπορούν και να πρέπει να λειτουργήσουν ως συμπληρωματικό βοήθημα ως προς μια διοικητική απόφαση. Μακροπρόθεσμα κάτι τέτοιο πιθανόν να ανατραπεί καθώς το ερώτημα πλέον είναι, όχι το εάν τελικά θα χρησιμοποιείται η τεχνητή νοημοσύνη στην διακυβέρνηση, αλλά πόσο, που και πως. Προς το παρόν πάντως, υπάρχει μια γενική αίσθηση, πως η βέλτιστη επιλογή είναι η συμπληρωματική χρήση της τεχνητής νοημοσύνης στη διακυβέρνηση και όχι η αντικατάσταση της ανθρώπινης απόφασης.

Σημειώνεται δε, πως η δυνατότητα ανάθεσης διοικητικών εργασιών σε εργαλεία τεχνητής νοημοσύνης περιπλέκεται ακόμη περισσότερο όταν ένα τέτοιο σύστημα καλείται να παράξει μία απόφαση που δεν βασίζεται αποκλειστικά σε δεδομένα, αλλά και σε αξίες και αξιολογήσεις (Μήτρου, 2023). Ευτυχώς, σε αρκετές έννομες τάξεις, υπάρχει η νομική δυνατότητα της διακριτικής ευχέρειας στις διοικητικές αρχές ώστε η δημόσια διοίκηση να έχει ορισμένα περιθώρια ως προς την εφαρμογή αόριστων νομικών εννοιών σε μία συγκεκριμένη περίπτωση που καλείται να αντιμετωπίσει.

Είναι επίσης γεγονός πως οι άνθρωποι μπορούν να επιλέξουν να υποπέσουν σε σφάλματα ώστε να είναι πιο αποτελεσματικοί. Αντίθετα, τα συστήματα τεχνητής νοημοσύνης, μπορεί θεωρητικά να είναι αντικειμενικότερα, όμως αυτό δεν σημαίνει πως είναι και δίκαια στη πράξη (Bullock, 2019). Λαμβάνοντας υπόψιν τα παραπάνω, το Χάρτη των Θεμελιωδών Δικαιωμάτων και ειδικά το δικαίωμα στη χρηστή διοίκηση (άρθρο 41), είναι δεδομένη η ανάγκη για εποπτεία στη χρήση της τεχνητής νοημοσύνης στη δημόσια διοίκηση και λήψη αποφάσεων ώστε να δύναται η αιτιολόγηση με τρόπο λεπτομερή και εύληπτο στο ενδιαφερόμενο πρόσωπο-αποδέκτη μιας διοικητικής απόφασης για το πώς εφαρμόστηκε η διαδικασία στην ατομική περίπτωσή του (Hofmann, 2011).

Γενικά, και όπως υποστηρίζεται από τους Barth and Arnold (1999), η τεχνητή νοημοσύνη δύναται να ενισχύσει τη δημοκρατική διακυβέρνηση, καθώς μπορεί να προγραμματιστεί ώστε να εκτελεί με αποτελεσματικό και αποδοτικό τρόπο. Ως εκ τούτου, η δίκαιη τεχνητή νοημοσύνη, η οποία ως έννοια συσσωματώνει πολλές από τις αντίστοιχες αρχές και αξίες, ίσως πρέπει να είναι η κομβική επιδίωξη (Μήτρου, 2023).

10 Συμπεράσματα

Είναι γεγονός πως η τέταρτη βιομηχανική επανάσταση με κύρια συστατικά τη ψηφιοποίηση των δραστηριοτήτων της κοινωνίας και τη χρήση σύγχρονων τεχνολογιών, με τη τεχνητή νοημοσύνη σε εξέχουσα θέση, βρίσκεται ήδη εν εξελίξει. Ειδικότερα τη τελευταία δεκαετία, η πρόοδος στον τομέα της εν λόγω τεχνολογίας είναι ραγδαία, ενώ παράλληλα η εντατικοποίηση της υιοθέτησής της, τόσο στον ιδιωτικό όσο και το δημόσιο τομέα, επιβάλλει στις κυβερνήσεις να αντιμετωπίσουν ταχύτατες μεταβολές σε όλες τις εκφάνσεις της ανθρώπινης δραστηριότητας.

Αυτή η διπλωματική εργασία, προσπάθησε μέσω της αντίστοιχης βιβλιογραφικής επισκόπησης να αναδείξει, τόσο τα οφέλη από τη χρήση της τεχνητής νοημοσύνης στο δημόσιο τομέα όσο και τις προκλήσεις που καλούνται οι κυβερνήσεις να αντιμετωπίσουν. Αναγνωρίζοντας τις περιορισμένες δυνατότητες πρόβλεψης του μέλλοντος, θα πρέπει οι υπεύθυνοι χάραξης της πολιτικής να δημιουργήσουν ένα πλαίσιο διαχείρισης του μετασχηματισμού της κοινωνίας σε ένα περιβάλλον με έντονη χρήση της τεχνητής νοημοσύνης.

Τα δυνητικά οφέλη για την οικονομία, τις κοινωνικές αλληλεπιδράσεις, τη δημόσια διοίκηση και γενικότερα τη ποιότητα ζωής των πολιτών από τη χρήση της εν λόγω τεχνολογίας, είναι τεράστια, έχουν αναγνωριστεί ήδη και θα ήταν ανεπίτρεπτο να χαθούν από κακές πολιτικές επιλογές και πρακτικές. Ωστόσο, οι ραγδαίες τεχνολογικές εξελίξεις στον τομέα αυτό είναι συνεχείς και είναι αρκετά δύσκολο για τις κυβερνήσεις να συμβαδίσουν αντιστοίχως στη δημιουργία συνολικών πλαισίων υπεύθυνης διαχείρισης για την αντιμετώπιση των προκλήσεων που προκύπτουν. Μάλιστα, ο δημόσιος τομέας, σε σχέση με άλλους τομείς, έχει τον κυριότερο και κρισιμότερο ρόλο στην εξεύρεση αποτελεσματικών απαντήσεων στις ηθικές και όχι μόνο προκλήσεις που προκύπτουν από τη χρήση της τεχνητής νοημοσύνης.

Ήδη, οι κυβερνήσεις προσπαθούν μέσω αυτής της καινοτομίας να παράξουν δημόσια αξία και να επωφεληθούν, όμως οι προσπάθειες εφαρμογής αναδεικνύουν την ανάγκη για την υιοθέτηση μια νέας πολιτικής προσέγγισης καθώς τα υφιστάμενα διοικητικά και ρυθμιστικά πλαίσια δείχνουν να μην μπορούν να ανταποκριθούν στις προκλήσεις που προκύπτουν.

Απαιτείται λοιπόν, μια προσαρμοστική προσέγγιση διακυβέρνησης βασισμένη στις ηθικές αξίες και τις σύγχρονες ανάγκες της κοινωνίας. Όμως αυτό κάθε άλλο παρά εύκολο είναι. Όπερ σημαίνει πως απλά η επιλογή μιας πολιτικής δεν είναι αρκετή για να διασφαλίσει τα οφέλη και να μετριάσει του κινδύνους από τη χρήση της τεχνητής νοημοσύνης. Αυτό που απαιτείται είναι η μελέτη διαχείρισης των προηγούμενων ριζοκέλευθων τεχνολογιών (π.χ. Διαδίκτυο), εις βάθος κατανόηση της πολυπλοκότητας της τεχνητής νοημοσύνης, θετική και συμβιβαστική διάθεση δημιουργίας ενός πλαισίου σχεδιασμού και λειτουργίας τέτοιων εφαρμογών και έντονη προσαρμοστικότητα στις αναδυόμενες προκλήσεις.

Όπως αναφέρεται σε παραπάνω κεφάλαιο, οι επιλογές διακυβέρνησης της τεχνητής νοημοσύνης εκτείνονται από το φάσμα της αυτορρύθμισης μέχρι το καταρτισμό αυστηρού πλαισίου επιβολής. Σύμφωνα με τη βιβλιογραφική επισκόπηση δεν υπάρχουν αποδείξεις για τη βέλτιστη επιλογή, αλλά προκύπτει, πως μια υβριδική διακυβέρνηση με ενδελεχή εποπτεία και προσαρμοστικότητα, είναι ίσως η μοναδική και βιώσιμη επιλογή. Προκύπτει επίσης, πως η φύση της εν λόγω τεχνολογίας επιβάλλει τη σύμπραξη όλων των δυνάμεων (κυβερνήσεις, δημόσιος και ιδιωτικός τομέας, κοινωνικοί φορείς, κτλ.) ώστε να διασφαλιστεί η απαραίτητη συναίνεση για την αποτελεσματική κατανόηση και τη διαχείριση των κινδύνων και τη μεγιστοποίηση του οφέλους για τη κοινωνία. Μάλιστα η αναγκαία αυτή σύμπραξη δεν δύναται να περιορίζεται σε τοπικό και εθνικό επίπεδο, αλλά η παγκοσμιοποιημένη οικονομία και γενικά κοινωνία, μάλλον απαιτεί αντίστοιχη διεθνή αν όχι παγκόσμια δυνατότητα διακυβέρνησης, τουλάχιστον σε βασικές αρχές.

Στη παρούσα φάση, ήδη οι κυβερνήσεις έχουν κάνει βήματα προς το καταρτισμό του πλαισίου της διακυβέρνησης της τεχνητής νοημοσύνης. Παρόλα αυτά, η επισκόπηση των πολιτικών αυτών επιλογών και της αντίστοιχης βιβλιογραφίας οδηγεί στο συμπέρασμα πως υπάρχουν ήδη διαφοροποιήσεις μεταξύ των επιλεχθέντων εθνικών στρατηγικών όσο και της φιλοδοξίας των κυβερνήσεων για τη διακυβέρνηση και τη ρύθμιση της τεχνητής νοημοσύνης. Μελετώντας τις εθνικές στρατηγικές και τα νομοσχέδια ρύθμισης, σε άλλες περιπτώσεις δίδεται χώρος για περαταίρω τεχνολογική εξέλιξη και καινοτομία στο τομέα, ενώ σε άλλες περιπτώσεις επιλέγονται αυστηρότερα πλαίσια σύμφωνα με την αρχή της προφύλαξης.

Σε όλες όμως τις περιπτώσεις, είναι παθομολογούμενη η αδυναμία καταρτισμού ενός αποτελεσματικού πλαισίου διακυβέρνησης που θα αποδίδει τα οφέλη από τη χρήση της τεχνητής νοημοσύνης ενώ θα καλύπτει επαρκώς την απαιτούμενη ευθύνη και λογοδοσία όταν αυτό είναι απαραίτητο.

Τέλος, και όπως αναφέρθηκε και σε προηγούμενη ενότητα, η βιβλιογραφία και τα αποτελέσματα χρήσης της τεχνητής νοημοσύνης στο δημόσιο τομέα παραμένουν περιορισμένα και κυρίως προέρχονται από ανεπτυγμένες και ηγέτιδες στον τομέα χώρες. Ως εκ τούτου, υπάρχει ανάγκη να μελετηθούν αντιστοίχως από τους ερευνητές του τομέα, ώστε να ενισχύσουν με νέες πληροφορίες και με γνώση τις κυβερνήσεις, οι οποίες με τη σειρά τους θα πρέπει να ελέγξουν και να προσαρμόσουν τα πλαίσια διακυβέρνησης της τεχνητής νοημοσύνης.

11 Επίλογος

Όπως επισημαίνεται από τους στοχαστές του προηγούμενου αιώνα που ασχολήθηκαν με τις επιπτώσεις της αυτοματοποίησης της βιομηχανίας και του εμπορίου, υπάρχει μόνο ένας δρόμος, και αυτός είναι μόνο μπροστά. Το ίδιο ισχύει και για τη τεχνητή νοημοσύνη. Πράγματι, τα παραδοσιακά πλαίσια διακυβέρνησης δεν μπορούν να προσαρμοστούν στις τεχνολογικές εξελίξεις και ως αποτέλεσμα, νέες καινοτόμες, πολυμετοχικές και προσαρμοστικές προσεγγίσεις ίσως να είναι ικανές να διασφαλίσουν το μέγιστο όφελος και την εξάλειψη των κινδύνων από τη χρήση της τεχνητής νοημοσύνης. Η συνέναιση όλων των φορέων είναι προϋπόθεση για τον ομαλό μετασχηματισμό της κοινωνίας που έπεται λόγω της τεχνολογίας αυτής, ενώ η θετική και συμβιβαστική διάθεση κρίνεται αναγκαία για τις προκλήσεις ευθύνης και λογοδοσίας που ήδη προκύπτουν και θα πρέπει να αντιμετωπιστούν.

Αν και όπως υποστηρίζει ο Supiot (2015), με την ευρεία χρήση της τεχνητής νοημοσύνης απομακρυνόμαστε ολοένα και περισσότερο από τη διακυβέρνηση μέσω δίκαιων νόμων και οδεύουμε στη διακυβέρνηση από αριθμούς, υπάρχουν δυνατότητες μετριασμού αυτής της εξέλιξης. Με συμμετοχή όλων των παραγόντων της κοινωνίας, και με τη συνεχή ανθρώπινη εποπτεία, πιθανόν να είναι εφικτό να αποφύγουμε δυστοπικές καταστάσεις.

Αν η τεχνητή νοημοσύνη είναι κοινωνικοκεντρική και έχει ως πρωταρχικό στόχο να ενισχύσει τις θετικές ενέργειες των ανθρώπων αντί να τους αντικαταστήσει, τότε αυξάνονται οι πιθανότητες επιτυχούς αντιμετώπισης των προκλήσεων. Σε αυτό το πλαίσιο, οι κυβερνήσεις θα πρέπει να επιλέξουν μια προσέγγιση διακυβέρνησης όπου η τεχνητή νοημοσύνη θα πρέπει να προσαρμοστεί στις ηθικές αξίες και θεμελιώδεις αρχές της κοινωνίας παρά το ανάποδο. Σε κάθε περίπτωση όμως, το βέβαιο είναι πως βαδίζουμε ήδη προς την καθιέρωση της κοινωνία της τεχνητής νοημοσύνης (Wirtz, 2019).

Βιβλιογραφία

- Aoki, N. (2020). An experimental study of public trust in AI chatbots in the public sector. *Government Information Quarterly*, 37(4), 101490. <https://doi.org/10.1016/j.giq.2020.101490>
- Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications Policy*, 44(6), 101976. <https://doi.org/10.1016/j.telpol.2020.101976>
- Kitchenham, B. (2004). Procedures for performing systematic reviews. Keele, UK, Keele University, 33(2004), 1–26
- Statista Search Department (2023), AI Papers Publication, <https://www.statista.com/statistics/941037/ai-paper-publications-worldwide-by-country/+>
- Παπαστάμου Π.(2018), "Τεχνητή Νοημοσύνη και η επίδραση της στις ανθρώπινες σχέσεις", Τεχνολογικό Εκπαιδευτικό Ίδρυμα Δυτ. Ελλάδας
- Gebru, T. (2019). Oxford handbook on AI ethics book chapter on race and gender. arXiv preprint arXiv:1908.06165
- Dobrev, D. (2012) 'A Definition of Artificial Intelligence'. <http://proxy.lib.chalmers.se/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsarx&AN=edsarx.1210.1568&lang=sv&site=eds-live&scope=site>
- Simmons, A. B., & Chappell, S. G. (1988). Artificial intelligence-definition and practice. *IEEE Journal of Oceanic Engineering*. doi:10.1109/48.551
- Vinnova (2018) Artificiell intelligens i svenskt näringsliv och samhälle, Vinnova. <https://www.vinnova.se/globalassets/dokument/ai-analys-slutrapport-180430-final.pdf>
- Lennartson, J. (2018). Demystifying Artificial Intelligence -Exploring how public sector organizations can approach Artificial Intelligence
- UNESCO. (2020). First version of a draft text of a recommendation on the ethics of artificial intelligence. Retrieved from <https://unesdoc.unesco.org/ark:/48223/pf0000373434>

- Βλαχάβας Ι. , Κεφαλάς Π. , Βασιλειάδης Ν. , Κόκκορας Φ. , Σακελλαρίου Η., (2006), «Τεχνητή Νοημοσύνη» , Εκδόσεις Πανεπιστημίου Μακεδονίας , Γ' έκδοση
- European Commission (2021). Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). <https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-european-approach-artificial-intelligence>
- Gurkaynak, G., Yilmaz, I. and Haksever, G. (2016) 'Stifling artificial intelligence: Human perils', Computer law & security review, (5), p. 749. <http://proxy.lib.chalmers.se/login?url=http://search.ebscohost.com.proxy.lib.chalmers.se/login.aspx?direct=true&db=edsbl&AN=vdc.100053839046.0x000001&lang=sv&site=eds-live&scope=site>
- Adams, S., Arel, I., Bach, J., Coop, R., Furlan, R., Goertzel, B., Hall, J. S., Samsonovich, A., Scheutz, M., Schlesinger, M., Shapiro, S. C., Sowa, J. (2012). Mapping the landscape of human-level artificial general intelligence. AI Magazine, 33 (1), 25–42. doi:10.1609/aimag.v33i1.2322
- Nick Bostrom, Superintelligence: Paths, Dangers, Strategies (Oxford: Oxford University Press, 2014)
- Jerry Kaplan, Humans Need Not Apply: A Guide to Wealth and Work in the Age of Artificial Intelligence (New Haven, CT: Yale University Press, 2015)
- Thierer, A. D., Castillo O'Sullivan, A., & Russell, R. Artificial intelligence and public policy. Mercatus Research Paper, (2017)
- “Artificial Intelligence Market—Global Industry Analysis, Size, Share, Growth, Trends and Forecast 2016–2024” (Market Transparency Research, Albany, NY, March 2, 2016)
- Goodfellow, I., Bengio, Y., & Courville, A. Deep learning. MIT press. (2016)
- Cihon, P., Schuett, J., & Baum, S. D. (2021). Corporate governance of artificial intelligence in the public interest. Information.
- Fukuyama F. (2013), “What Is Governance?.” CGD Working Paper 314, Washington, DC: Center for Global Development, <http://www.cgdev.org/content/publications/detail/1426906>
- Brown, M. A. (2017). Global Governance and National Governance. In A. Farazmand (Ed.), Global Encyclopedia of Public Administration, Public Policy, and Governance

(pp. 1–9). Springer International Publishing. https://doi.org/10.1007/978-3-319-31816-5_1159-1

- Misuraca, G. and Van Noordt, C., *AI Watch - Artificial Intelligence in public services*, EUR 30255 EN, Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-19540-5, doi:10.2760/039619, JRC120399.
- IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. (2019). *Ethically aligned design. A vision for prioritizing human well-being with autonomous and intelligent systems* https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e.pdf?utm_medium=PR&utm_source=Web&utm_campaign=EAD1e&utm_content=geias&utm_term=undefined
- Wang, W., & Siau, K. (2018). *Artificial intelligence: a study on governance, policies, and regulations*.
- Wirtz, B. W., Weyerer, J. C., & Sturm, B. J. (2020). The dark sides of artificial intelligence: An integrated AI governance framework for public administration. *International Journal of Public Administration*, 43(9), 818-829.
- Thierer, A. D., Castillo O'Sullivan, A., & Russell, R. (2017). *Artificial intelligence and public policy*. Mercatus Research Paper.
- Manzoni, M., Medaglia, R., Tangi, L., Van Noordt, C., Vaccari, L. and Gattwinkel, D., *AI Watch Road to the adoption of Artificial Intelligence by the Public Sector: A Handbook for Policymakers, Public Administrations and Relevant Stakeholders*, EUR 31054 EN, Publications Office of the European Union, Luxembourg, 2022, ISBN 978-92-76-52131-0, doi:10.2760/693531, JRC1291003.
- Barker, L., Claps, M., Stevens, R., Crompvoets, J., Nasi, G., & Vandembroucke, D. (2021). *Leveraging the Power of Location Information and Technologies to Improve Public Services at the Local Level-State of the Art Report*. Publications Office of the European Union, Luxembourg, 2021, ISBN 978-92-76-41900-6, doi:10.2760/158709, JRC126562.
- Zuiderwijk, A., Chen, Y. C., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly*, 38(3), 101577.

- Alexopoulos, C., Lachana, Z., Androutsopoulou, A., Diamantopoulou, V., Charalabidis, Y., & Loutsaris, M. A. (2019). How machine learning is changing e government. In Paper presented at the 12th international conference on theory and practice of electronic governance, Melbourne, Australia.
- Susar, D., & Aquaro, V. (2019, April). Artificial intelligence: Opportunities and challenges for the public sector. In Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance (pp. 418-426).
- Grand View Research, Artificial Intelligence In Healthcare Grand View Research, Inc., <https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-healthcare-market>
- Fatima, S., Desouza, K. C., & Dawson, G. S. (2020). National strategic artificial intelligence plans: A multi-dimensional analysis. *Economic Analysis and Policy*, 67, 178–194. <https://doi.org/10.1016/j.eap.2020.07.008>.
- Gupta, K. P. K. K. P. (2019). Artificial intelligence for governance in India: Prioritizing the challenges using analytic hierarchy process (AHP). *International Journal of Recent Technology and Engineering*, 8(2), 3756–3762
- Toll, D., Lindgren, I., Melin, U., & Madsen, C.Ø. (2019). Artificial intelligence in Swedish policies: Values, benefits, considerations and risks. In Paper presented at the international conference on electronic government, San Benedetto del Tronto, Italy.
- Bullock, J. B. (2019). Artificial intelligence, discretion, and bureaucracy. *The American Review of Public Administration*, 49, 751–761. <https://doi.org/10.1177/0275074019856123>
- Coch, L., & French Jr, J. R. (1948). Overcoming resistance to change. *Human relations*, 1(4), 512-532.
- Frischmann, B. and Selinger, E. (2018) *Re-engineering Humanity*. Cambridge, UK: CUP.
- Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organising data for trustworthy artificial intelligence. *Government Information Quarterly*, 37 (3), 101493.

- Chen, T., Ran, L., & Gao, X. (2019, June). AI innovation for advancing public service: The case of China's first Administrative Approval Bureau. In Proceedings of the 20th Annual International Conference on Digital Government Research (pp. 100-108).
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 2053951716679679.
- Bernd W. Wirtz, Jan C. Weyerer & Carolin Geyer (2019) Artificial Intelligence and the Public Sector—Applications and Challenges, *International Journal of Public Administration*, 42:7, 596-615, DOI: 10.1080/01900692.2018.1498103
- Patrick, L., Abney, K., & Bekey, G. A. (Eds.). (2012). *Robot ethics: The ethical and social implications of robotics*. Cambridge, MA: MIT Press.
- Bostrom, N., & Yudkowsky, E. (2014). The ethics of artificial intelligence. *The Cambridge Handbook of Artificial Intelligence*, 1, 316–334.
- Anderson, M., & Anderson, S. L. (Eds.). (2011). *Machine ethics*. New York, NY: Cambridge University Press.
- Zhang, B., Anderljung, M., Kahn, L., Dreksler, N., Horowitz, M. C., & Dafoe, A. (2021). Ethics and governance of artificial intelligence: Evidence from a survey of machine learning researchers. *Journal of Artificial Intelligence Research*, 71, 591-666.
- Dhar, P. (2020). The carbon impact of artificial intelligence. *Nat. Mach. Intell.*, 2(8), 423-425.
- Roberts, A. (2017). Five big challenges to AI adoption and success. ClickZ. Retrieved July, 2, 2018.
- Bughin, J., Hazan, E., Ramaswamy, S., Chui, M., Allas, T., Dahlström, P., Nicolaus., H., Monica, T. (2017). *Artificial intelligence: The next digital frontier?* Edited by McKinsey Global Institute.
- Marsh, N. S. (1959). International commission of Jurists, *The Rule of Law in a Free Society*. In New Delhi, India: A Report on International Congress of Jurists
- Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and information technology*, 6(3), 175–183.
- Taeihagh, A. (2021). Governance of artificial intelligence. *Policy and society*, 40(2), 137-157.

- Guihot, M., Matthew, A. F., & Suzor, N. P. (2017). Nudging robots: Innovative solutions to regulate artificial intelligence. *Vand. J. Ent. & Tech. L*, 20, 385
- Larsson, S. (2020). On the governance of artificial intelligence through ethics guidelines. *Journal of Law and Society*, 1, 23.
- Gasser, U., & Almeida, V. A. (2017). A layered model for AI governance. *IEEE Internet Computing*, 21(6), 58–62.
- Leiser, M., & Murray, A. (2016). The role of non-state actors and institutions in the governance of new and emerging digital technologies. In *The oxford handbook of law, regulation and technology*, Eds. R. Brownsword, E. Scotford, K. Yeung, & O. U. Press. (pp. 670–704).
- Linkov, I., Trump, B. D., Anklam, E., Berube, D., Boisseasu, P., Cummings, C., Ferson, S., Florin, M., Goldstein, B., Hristozov, D., Jensen, K.A., Katalagarianakis, G., Kuzma, J., Lambert, J.H., Malloy, T., Malsch, I., Marcomini, A., Merad, M., Palma-Oliveira, J., Perkins, E., Renn, O., Seager, T., Stone, V., Vallero, D., Vermeire, T. (2018b). Comparative, collaborative, and integrative risk governance for emerging technologies. *Environment Systems and Decisions*, 38 (2), 170–176. <https://doi.org/10.1007/s10669-018-9686-5>.
- Tan, S. Y., & Taeihagh, A. (2021b). Adaptive governance of autonomous vehicles: Accelerating the adoption of disruptive technologies in Singapore. *Government Information Quarterly*, 38(2), 101546.
- Gasser, U., & Almeida, V. A. (2017). A layered model for AI governance. *IEEE Internet Computing*, 21(6), 58–62.
- Rahwan, I. (2018). Society-in-the-loop: Programming the algorithmic social contract. *Ethics and Information Technology*, 20(1), 5–14.
- European Commission. (2018). Artificial intelligence for Europe. In *Communication from the commission to the European Parliament, the European council, the council, the European economic and social committee and the committee of the regions*. Retrieved from Brussels <https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe>.
- Radu, R. (2021). Steering the governance of artificial intelligence: national strategies in perspective. *Policy and society*, 40(2), 178-193.

- Dutton, T. (2018). Building an AI world: Report on national and regional strategies. CIFAR, 6 December, <https://www.cifar.ca/cifarnews/2018/12/06/building-an-ai-world-report-on-national-and-regional-ai-strategies>.
- Kop, M. (2021, September). EU artificial intelligence act: the European approach to AI. Stanford-Vienna Transatlantic Technology Law Forum, Transatlantic Antitrust and IPR Developments, Stanford University, Issue.
- Mökander, J., Juneja, P., Watson, D. S., & Floridi, L. (2022). The US Algorithmic Accountability Act of 2022 vs. The EU Artificial Intelligence Act: what can they learn from each other?. *Minds and Machines*, 1-8.
- Daniel Akst, (2013) "What Can We Learn from Past Anxiety over Automation?," *Wilson Quarterly*, Summer 2013.
- Μήτσιος Στέφανος (2022), " Ν. 4961/2022: Το Νέο Εθνικό Νομοθετικό Πλαίσιο για τις Αναδυόμενες Τεχνολογίες Πληροφορικής και Επικοινωνιών", https://www.ey.com/el_gr/tax/tax-alerts/l-4961-2022-the-greek-legal-framework-on-emerging-technologies
- ΝΟΜΟΣ ΥΠ' ΑΡΙΘΜ. 4961 ΦΕΚ Α 146/27.7.2022, " Αναδυόμενες τεχνολογίες πληροφορικής και επικοινωνιών, ενίσχυση της ψηφιακής διακυβέρνησης και άλλες διατάξεις".
- Π. Δαγτόγλου, "Διοικητικό Δίκαιο", Αθήνα-Θεσσαλονίκη: Εκδόσεις Σάκκουλα, 2014, σ. 343.
- J. Etscheid, «Artificial Intelligence in Public Administration — A Possible Framework for Partial and Full Automation», στο: I. Lindgren κ.ά. (επιμ.), *International Conference on Electronic Government*, Springer, 2019, 248-261.
- Λ. Μήτρου, "Μπορεί ο Αλγόριθμος... να είναι ηθικός, να είναι δίκαιος, να είναι διαφανής, να δικάζει & να διοικεί;", Πανεπιστημιακές Εκδόσεις Κρήτης, 2023
- H. Hofmann, G.C. Rowe & A. Türk, *Administrative Law and Policy of the European Union*, Oxford University Press, 2011, σ. 193-195.
- T.J. Barth & E. Arnold, «Artificial intelligence and administrative discretion: Implications for public administration», *The American Review of Public Administration*, 29 (1999) 332-351
- Supiot, A. (2015) *La Gouvernance par les Nombres*. Paris: Fayard.

