



ΠΑΝΕΠΙΣΤΗΜΙΟ ΜΑΚΕΔΟΝΙΑΣ

ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

ΘΡΑΚΗΣ

ΤΜΗΜΑ ΕΦΑΡΜΟΣΜΕΝΗΣ

ΤΜΗΜΑ ΝΟΜΙΚΗΣ

ΠΛΗΡΟΦΟΡΙΚΗΣ

ΔΙΔΡΥΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

ΔΙΚΑΙΟ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗ

**ΔΕΟΝΤΟΛΟΓΙΑ-ΗΘΙΚΗ ΣΤΗΝ ΕΦΑΡΜΟΓΗ ΤΗΣ ΤΕΧΝΗΤΗΣ
ΝΟΗΜΟΣΥΝΗΣ ΣΤΟ ΔΗΜΟΣΙΟ ΤΟΜΕΑ**

Artificial Intelligence ethics in the public sector

Διπλωματική Εργασία

Του

Κωνσταντίνου Γ. Χαϊντούτη

Κατατέθηκε σαν προαπαιτούμενο για την συμπλήρωση των απαιτήσεων για την απόκτηση μεταπτυχιακού διπλώματος στο “Δίκαιο και πληροφορική”

Θεσσαλονίκη, Ιούνιος 2019.

ΔΕΟΝΤΟΛΟΓΙΑ-ΗΘΙΚΗ ΣΤΗΝ ΕΦΑΡΜΟΓΗ ΤΗΣ ΤΕΧΝΗΤΗΣ ΝΟΗΜΟΣΥΝΗΣ
ΣΤΟ ΔΗΜΟΣΙΟ ΤΟΜΕΑ

Κωνσταντίνος Γ. Χαϊντούτης

Πτυχίο Μηχανολόγου Μηχανικού Πολυτεχνικής σχολής ΑΠΘ, 2002.

Πτυχίο Εφαρμοσμένης Πληροφορικής Πανεπιστημίου Μακεδονίας, 2014.

Διπλωματική Εργασία

υποβαλλόμενη για τη μερική εκπλήρωση των απαιτήσεων του

ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΤΙΤΛΟΥ ΣΠΟΥΔΩΝ ΣΤΟ ΔΙΚΑΙΟ & ΠΛΗΡΟΦΟΡΙΚΗ

Επιβλέπων Καθηγητής

Ονοματεπώνυμο Καθηγητή

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την

.....

.....

.....

Κωνσταντίνος Γ. Χαϊντούτης.

Περίληψη

Τα τελευταία χρόνια η τεχνητή νοημοσύνη έχει προσφέρει πολλές εφαρμογές που μπορούν να φανούν χρήσιμες στην ζωή μας. Ο δημόσιος τομέας μπορεί να εκμεταλλευτεί πολλές από αυτές ώστε να προσφέρει καλύτερα αποτελέσματα σε τομείς που αυτός ελέγχει και προς την κατεύθυνση που έχει χαράξει στην γενική πολιτική του. Σε μια ιδεατή δημόσια γενική πολιτική η επιδίωξη θα ήταν η κατάκτηση των στόχων της ηθικής, δηλαδή η τελεολογική αναζήτηση της **ευδαιμονίας** των πολιτών μέσω της επιδίωξης της μεσότητας στην στάθμιση πλεονεκτημάτων και μειονεκτημάτων που έχει η τεχνητή νοημοσύνη. Με κύριο άξονα την διατήρηση της μεσότητας η παρούσα έρευνα περιστρέφεται γύρω, αφενός από την μελέτη αυτών καθ' αυτών των εφαρμογών που έχει να προσφέρει σήμερα η τεχνητή νοημοσύνη στο δημόσιο τομέα και αφετέρου από την μελέτη γύρω από την ηθική περιφρούρηση αυτών μέσω αξιών που έχουν προταθεί. Η παρούσα έρευνα δεν μένει μόνο στην παρουσίαση της παγκόσμιας κατάστασης γύρω από την τεχνητή νοημοσύνη και τις πρωτοβουλίες που έχουν πάρει κράτη και ενώσεις κρατών ώστε να υπάρχει ένα ηθικό πλαίσιο προστασίας, αλλά κρατά και μια κριτική στάση απέναντι σε αποφάσεις και πρακτικές των κρατών που πιθανόν κρύβουν κινδύνους για τον άνθρωπο, επισημαίνοντάς τους αλλά και προτείνοντας λύσεις. Με βάση τη μέθοδο διενέργειας συστηματικών βιβλιογραφικών επισκοπήσεων στον τομέα των πληροφοριακών συστημάτων Webster & Watson επιχειρήθηκε να δοθούν απαντήσεις σε όλες τις πιθανές απορίες που θα είχε ο αναγνώστης όπως, ποιο είναι το αντικείμενο και τα όρια της έρευνας, πως και που έγινε η έρευνα, γιατί μας ενδιαφέρει το αντικείμενο της έρευνας, γιατί είναι επίκαιρο, αν έχει περιθώρια βελτίωσης κ.α. Γενικότερα στην έρευνα επιχειρήθηκε να γίνει μια συνολική χαρτογράφηση, φωτίζοντας όλες τις πλευρές, επισημαίνοντας όλες τις θετικές και αρνητικές πτυχές που υπάρχουν επεξηγώντας πάντοτε την στάση που κρατήθηκε. Έτσι αφού πρώτα αποσαφηνίστηκαν οι κρίσιμοι όροι (τεχνητή νοημοσύνη, ηθική, δημόσιος τομέας), δόθηκε απάντηση στο ερώτημα γιατί η έρευνα είναι επίκαιρη, παρουσιάστηκαν εφαρμογές που μπορεί να χρησιμοποιήσει ή ήδη χρησιμοποιεί το δημόσιο, τεκμηριώθηκε η άποψη ότι πρέπει να τηρείται ένας κώδικας ηθικής, μελετήθηκε η σχέση της ηθικής με την νομιμότητα, παρουσιάστηκαν προτάσεις από ομάδες ηθικής ανά τον κόσμο, ερευνήθηκε η χρηματοδότηση που δίνεται από το δημόσιο, επιχειρήθηκε η εύρεση των κινητήριων δυνάμεων πίσω από την προώθηση της τεχνητής νοημοσύνης, ερευνήθηκαν προτάσεις για το πώς πρέπει να προετοιμαστούν δημόσιο και πολίτες για να υποδεχτούν την τεχνητή νοημοσύνη, παρουσιάστηκαν εταιρείες που παρέχουν σχετικό λογισμικό, έγινε μια SWOT ανάλυση, παρουσιάστηκε μια έρευνα για την αποδοχή της τεχνητής νοημοσύνης από τους πολίτες, έγινε μια σύνδεση των κινδύνων που κρύβει η τεχνητή νοημοσύνη με τον νέο κανονισμό για την προστασία των προσωπικών δεδομένων, έγινε μια προσπάθεια εντοπισμού της χρήσης της τεχνητής νοημοσύνης από το δημόσιο μέσα στο ευρύτερο πλαίσιο της δημιουργίας μιας κοινωνίας που παρέχει δεδομένα μεγάλης κλίμακας (big data) σε ένα περιβάλλον δικτύου οντοτήτων (IoT), και τέλος έγιναν προτάσεις για το μέλλον. Απώτερος στόχος της εργασίας είναι η ανάδειξη της καλής

πλευράς της τεχνητής νοημοσύνης που πρέπει ο δημόσιος τομέας να μεγεθύνει αλλά και η επισήμανση κινδύνων που κρύβονται και πρέπει να αποφεύγονται.

|

Λέξεις Κλειδιά:

Τεχνητή νοημοσύνη,

Ηθική,

Δημόσιος τομέας.

|

Abstract

In recent years, artificial intelligence has provided many applications that could be useful in our lives. The public sector can take the advantage of many of them to deliver better results in areas that controls and in the direction that has set in its general policy. In an ideal public general policy, the pursuit would be to attain the objectives of morality, namely the teleological search for citizens' happiness through the pursuit of the difference in the weighting of the advantages and disadvantages of artificial intelligence. With the main focus on the maintenance of the balance, the research revolves around the study of these applications of today's artificial intelligence in the public sector and the study of the moral protection of these by means of values that have been proposed. That research is not only about presenting the global situation about artificial intelligence and the initiatives taken by states and associations of states in order to have a moral framework of protection, but also holds a critical attitude towards decisions and practices of states that are likely to hide dangers for the people, pointing to them as well as proposing solutions. Based on the Webster & Watson Guide, there was an attempt to answer all the possible questions the reader would have, such as what is the subject and the limits of the research, how the research was done, why we are interested in the subject of the research, why it is timely, if it has room for improvement etc? The research attempted to make a comprehensive mapping, illuminating all aspects, highlighting all the positive and negative aspects that exist, always explaining the attitude held. Thus, after the critical terms (artificial intelligence, ethics, public sector) were clarified, several questions were answered, as to why the research is timely, what are the applications that the public could use. The research also proposes the idea that a code of ethics should be observed, investigates the relationship ethics have with legitimacy, concentrates proposals from ethics groups around the world that were presented, searches if public funding was spent, examines all the the driving forces behind the promotion of artificial intelligence, finds also if there were suggestions on how to prepare public and citizens to receive artificial intelligence, presents also companies that provide relevant software, makes a SWOT analysis, presents a survey on the acceptance of artificial intelligence by citizens, examines a link between the risks the artificial intelligence hides and the new regulation on the protection of personal data, makes an attempt to identify the use of artificial intelligence by the public within the wider context of creating a society which provides large data in a network of entities (IoT), and finally makes suggestions for the future. The ultimate goal of the research is to highlight the good aspect of artificial intelligence that the public sector needs to enlarge but also to identify risks that are hidden and should be avoided.

Keywords:

Artificial Intelligence,

Ethics,

Public sector.

Πρόλογος – Ευχαριστίες.

Οι πρόσφατες αλλαγές στο πεδίο των προσωπικών δεδομένων αλλά και η ταυτόχρονη ραγδαία ανάπτυξη εφαρμογών τεχνητής νοημοσύνης, που σχετίζονται με τα προσωπικά δεδομένα, έδωσαν το έναυσμα για μια πιο απομακρυσμένη επόπτευση της πορείας που ακολουθεί η χρήση των επιτευγμάτων της τεχνολογίας από ανθρώπους που έχουν την πολιτική εξουσία σε μια χώρα. Αυτή η επόπτευση γίνεται κάτω από το πρίσμα της ηθικής και της τελεολογικής αναζήτησης του τι είναι καλύτερο για την ζωή μας.

Στην ολοκλήρωση αυτής της διατριβής στάθηκε εμπόδιο η έλλειψη σχετικής βιβλιογραφίας λόγω της σχετικά πρόσφατης ανάδειξης του θέματος αλλά και λόγω της γενικότερης αίσθησης της ανάγκης για κατεπείγουσα αλλαγή του πλαισίου που διέπει τα προσωπικά δεδομένα χωρίς πρωτίστως να διερευνηθεί η μακροπρόθεσμη αλλαγή στην ποιότητα της ζωής των πολιτών.

Στην προσπάθεια εύρεσης του θέματος αλλά και της ανάπτυξης της διατριβής συνέβαλε το σύνολο των καθηγητών και του εκπαιδευτικού προσωπικού του μεταπτυχιακού προγράμματος, γι' αυτό και τους αξίζουν οι θερμότερες ευχαριστίες.

Περιεχόμενα

Τίτλος	Σελ.
1. Εισαγωγή	11
1.1 Πρόβλημα – Σημαντικότητα του θέματος.	11
1.2 Σκοπός – Στόχοι.	11
1.3 Ερωτήματα – Υποθέσεις.	12
1.4 Συνεισφορά.	12
1.5 Βασική Ορολογία.	13
1.6 Διάρθρωση της μελέτης.	13
2. Βιβλιογραφική Επισκόπηση – Θεωρητικό Υπόβαθρο.	15
3. Μεθοδολογία.	16
3.1. Περιγραφή της μεθοδολογίας	16
3.2. Προεπισκόπηση της βιβλιογραφίας	17
4. Αποσαφήνιση κρίσιμων όρων.	18
4.1. Εισαγωγή.	18
4.2. Ηθική.	18
4.3. Τεχνητή νοημοσύνη.	20
4.4. Δημόσιος τομέας.	24
5. Παράγοντες που επικαιροποιούν την έρευνα γύρω από ζήτημα της ηθικής στις εφαρμογές τεχνητής νοημοσύνης στον δημόσιο τομέα.	25
5.1.Εισαγωγή.	25
5.2. Παράγοντες που επικαιροποιούν την έρευνα γύρω από ζήτημα της ηθικής.	25
5.3. Περίληψη.	27
6. Η ηθική στις εφαρμογές τεχνητής νοημοσύνης στον δημόσιο τομέα.	28
6.1. Εισαγωγή.	28
6.2. Η παρούσα κατάσταση στον δημόσιο τομέα.	28
6.3. Οι αλλαγές στην ανθρώπινη συμπεριφορά.	34
6.4. Ζητήματα εμπιστοσύνης προς την νέα τεχνολογία.	34
6.5. Περίληψη.	38
7. Η ανάγκη για τήρηση ενός κώδικα ηθικής.	39
7.1. Εισαγωγή.	39
7.2. Πως υλοποιείται η απαίτηση να υπακούει η τεχνολογία σε ένα σύνολο ηθικών αρχών;	39
7.3. Η εισαγωγή ενός ηθικού φίλτρου στον άνθρωπο.	41
7.3.1 Το νομικό πλαίσιο και η σχέση του με την ηθική.	41
7.4. Περίληψη.	45
8. Ελπιδοφόρες εφαρμογές και chatbots στην υπηρεσία του δημοσίου τομέα.	46
8.1. Εισαγωγή.	46
8.2. Ελπιδοφόρες εφαρμογές της τεχνητής νοημοσύνης.	46
8.3. Αναφορά σε chatbots που ήδη χρησιμοποιούνται για παροχή δημοσίων υπηρεσιών ή θα μπορούσαν να χρησιμοποιηθούν από τον δημόσιο τομέα.	55
8.4. Περίληψη.	59

9. Προτάσεις με σύνολα ηθικών αρχών και η ανάγκη για την ίδρυση ομάδων ηθικής.	61
9.1. Εισαγωγή.	62
9.2. Παραδείγματα με ομάδες ηθικής που εξέδωσαν αρχές για την ηθική.	62
9.3. Η επιτακτική ανάγκη για την ίδρυση ομάδων ηθικής.	82
9.5. Περίληψη.	84
10. Η χρηματοδότηση της τεχνητής νοημοσύνης από τον δημόσιο τομέα και το ηθικό περίβλημά των κινητήριων δυνάμεων πίσω από την προώθηση της τεχνητής νοημοσύνης.	86
10.1. Εισαγωγή.	86
10.2. Η χρηματοδότηση της τεχνητής νοημοσύνης από τον δημόσιο τομέα.	86
10.3. Οι κινητήριες δυνάμεις πίσω από την προώθηση της τεχνητής νοημοσύνης και το ηθικό περίβλημά τους.	87
10.4. Περίληψη.	89
11. Πως πρέπει να προετοιμαστεί ο δημόσιος τομέας και οι πολίτες ώστε να υποδεχθούν την τεχνητή νοημοσύνη;	90
11.1. Εισαγωγή.	90
11.2. Πως πρέπει να προετοιμαστεί ο δημόσιος τομέας για να υποδεχτεί την τεχνητή νοημοσύνη;	90
11.3. Πως πρέπει να προετοιμαστούν οι πολίτες για την υποδοχή της τεχνητής νοημοσύνης στον δημόσιο τομέα;	91
11.4. Έρευνες για την αποδοχή της τεχνητής νοημοσύνης από τους πολίτες.	99
11.5. Περίληψη.	102
12. Λογισμικό εταιρειών για ανάπτυξη νευρωνικών δικτύων και εφαρμογών τεχνητής νοημοσύνης.	104
12.1. Εισαγωγή.	104
12.2 Λογισμικό εταιρειών για ανάπτυξη νευρωνικών δικτύων και εφαρμογών τεχνητής νοημοσύνης.	104
12.3. Περίληψη.	106
13. SWOT analysis.	107
13.1. Εισαγωγή.	107
13.2. SWOT analysis.	107
13.2.1. Δυνατότητες.	108
13.2.2. Αδυναμίες.	109
13.2.3. Ευκαιρίες.	110
13.2.4. Απειλές.	113
13.3. Περίληψη.	115
14. Οι κίνδυνοι από την χρήση της τεχνητής νοημοσύνης και η σχέση τους με τον νέο κανονισμό της προστασίας των προσωπικών δεδομένων.	116
14.1. Εισαγωγή.	116
14.2. Οι κίνδυνοι από την χρήση της τεχνητής νοημοσύνης.	116
14.3. Η σχέση των κινδύνων με τον νέο κανονισμό της προστασίας των προσωπικών δεδομένων.	117
14.4. Περίληψη.	121
15. Ο ρόλος του δημοσίου τομέα στην συγκέντρωση δεδομένων μεγάλης κλίμακας (Big Data) και η σχέση του με την τεχνητή νοημοσύνη.	122

15.1. Εισαγωγή.	122
15.2. Ο ρόλος του δημοσίου τομέα στην συγκέντρωση δεδομένων μεγάλης κλίμακας (Big Data).	122
15.3. Η σχέση των δεδομένων μεγάλης κλίμακας με την τεχνητή νοημοσύνη.	123
15.4. Περίληψη.	124
16. Προτάσεις για περαιτέρω έρευνα και βελτίωση.	125
17. Επίλογος.	130
17.1 Σύνοψη και συμπεράσματα.	130
17.2 Όρια και περιορισμοί της έρευνας.	130
17.3 Μελλοντικές Επεκτάσεις.	131
Παράρτημα 1: Εξακρίβωση Επαναληψιμότητας.	133
Βιβλιογραφία.	138

1. Εισαγωγή

1.1 Πρόβλημα – Σημαντικότητα του θέματος

Ένα από τα πεδία έρευνας και ανάπτυξης που απασχολεί για δεκαετίες την επιστήμη της Πληροφορικής είναι αυτό της τεχνητής νοημοσύνης. Αν και η πρώτη αναφορά στον όρο της τεχνητής νοημοσύνης πηγάζει από το μακρινό παρελθόν (John McCarthy, 1956) ίσως και να θεμελιώθηκε ακόμη νωρίτερα από τον πατέρα της νεώτερης υπολογιστικής σκέψης Alan Turing (Bruce G. 1950). Σαν ιδέα βέβαια, και πιθανόν σαν εφαρμογή, έχει και αυτή προσμετρηθεί ήδη από την ανατολή του δυτικού κόσμου στους σπουδαίους Έλληνες προγόνους μας, αποδεικνυόμενο από τις αναφορές στην Οδύσσεια (θ555) στα πλοία των Φαιάκων τα οποία “διάβαζαν τη σκέψη” αλλά και στην Ιλιάδα (Σ376) για τα “αυτόματα” και “προσταζόμενα” τεχνικά δημιουργήματα του Ηφαίστου. Από τότε δεν έχει πάψει να είναι επίκαιρη αφού εκσυγχρονίζεται και εισχωρεί διαρκώς σε νέους τομείς της ανθρώπινης ζωής σε βαθμό που να προκαλεί το ενδιαφέρον και άλλων επιστημών, όπως της βιοτεχνολογίας, της κοινωνιολογίας, της ψυχολογίας, της οικονομίας, των πολιτικών επιστημών κ.α.

Έτσι λοιπόν ο άνθρωπος έχοντας μια πεπερασμένη νοημοσύνη και ιδιαίτερα έχοντας, ένα από τα συναισθήματα που γεννά αυτή, τον φόβο για το καινούργιο, έρχεται να αντιμετωπίσει την είσοδο της τεχνητής νοημοσύνης στη ζωή του. Αυτή η αντίδραση είναι ντετερμινιστική σε βαθμό που μπορούμε να πούμε ότι ο ανθρώπινος εγκέφαλος είναι προγραμματισμένος να αντιδρά με μια στάση σκεπτικιστική απέναντι σε κάθε καινούργιο σαν μια ασυνείδητη αντίδραση άμυνας (Joseph LeDoux. 2016).

Η σκεπτικιστική αυτή στάση απέναντι στο καινούργιο οδηγεί την λήψη μιας απόφασης μπροστά στην τελική κρίση που είναι το φίλτρο της ηθικής. Σε μια δηλαδή τελολογική αναζήτηση του καλού.

1.2 Σκοπός – Στόχοι

Η μελέτη αυτή έρχεται να επιχειρήσει μεταξύ άλλων και μια απάντηση στο ερώτημα αν η άμυνα απέναντι στο καινούργιο πρέπει να ενισχυθεί με κάποιου είδους απαγορεύσεις ή να κατευναστεί με την υπερίσχυση των ωφελημάτων ή να υιοθετηθεί μια ενδιάμεση στάση.

Πριν λοιπόν εισαχθεί μια εφαρμογή τεχνητής νοημοσύνης σε κάποιο τομέα της ανθρώπινης ζωής θα πρέπει να γίνει μια συζήτηση γύρω από θέματα ηθικής που εγείρονται.

Για να φτάσει στον απώτερο σκοπό της αυτή η μελέτη επιχειρεί μια απάντηση στο ερώτημα προς ποια πλευρά κλίνει η στάθμιση των πλεονεκτημάτων και μειονεκτημάτων από την εφαρμογή της τεχνητής νοημοσύνης; Βέβαια το πεδίο

εφαρμογών της τεχνητής νοημοσύνης είναι ευρύ, για αυτό και στην παρούσα θα περιοριστεί σε εφαρμογές στον δημόσιο τομέα.

Αντικείμενο λοιπόν αυτής της διατριβής είναι η μελέτη γύρω από την ηθική-δεοντολογία εφαρμογών της τεχνητής νοημοσύνης στον δημόσιο τομέα.

1.3 Ερωτήματα – Υποθέσεις

Όταν επιχειρείται μια πρόβλεψη μιας μελλοντικής κατάστασης με δεδομένα τις συνθήκες του σήμερα, ο κίνδυνος λήψης μιας λανθασμένης απόφασης είναι υπαρκτός. Στην παρούσα έρευνα, εκτός από την περιγραφή της παρούσας κατάστασης, επιχειρείται ουσιαστικά μια μελλοντική πρόβλεψη για την πορεία που θα έχει η τεχνητή νοημοσύνη στον δημόσιο τομέα και προφανώς διακινδυνεύεται η απάντηση στο κυρίαρχο ερώτημα: “η υιοθέτηση μιας εφαρμογής της τεχνητής νοημοσύνης από τον δημόσιο τομέα θα έχει κάποιον αντίκτυπο στην ποιότητα ζωής των πολιτών;”.

Στην αναζήτηση αυτή υπάρχουν τρεις αστάθμητοι παράγοντες: πρώτον η εξέλιξη της τεχνολογίας, δεύτερον η εξέλιξη της ανθρώπινης ηθικής και τρίτον η πιθανή αλλαγή στην έννοια του δημοσίου. Στην παρούσα μελέτη επιχειρείται η εξαγωγή συμπερασμάτων με βάση τα γνωστά τεχνολογικά όρια και λαμβάνοντας υπόψη την σημερινή ανθρώπινη εξέλιξη, με την ελπίδα βέβαια ότι αυτοί οι δύο παράγοντες με τον χρόνο θα βελτιωθούν. Σε κάποιες περιπτώσεις επιχειρείται μια πρόβλεψη για την εξέλιξη της κοινωνικής ζωής με το υποθετικό ερώτημα αν κάποιος από τους παραπάνω παράγοντες μεταβληθεί προς την μια ή την άλλη κατεύθυνση.

1.4 Συνεισφορά

Από τις αναζητήσεις που έγιναν σε όλες τις βάσεις συγγραμμάτων δεν βρέθηκε καμία μελέτη παρόμοια με την παρούσα, η οποία βλέπει την τεχνητή νοημοσύνη υπό το πρίσμα της ηθικής και μόνο για τον δημόσιο τομέα. Υπάρχουν μόνο επί μέρους έρευνες για κάθε πεδίο ξεχωριστά ή σε συνδυασμούς ανά δύο. Η έλλειψη αυτή οφείλεται στο ότι οι νέες εξελίξεις στον τομέα της τεχνητής νοημοσύνης είναι κάτι εντελώς καινούργιο και δεν έχουν ακόμη μελετηθεί. Επομένως η εργασία αυτή διεκδικεί μια πρωτοπορία στην σχετική έρευνα.

Βέβαια σαν φιλοσοφική αναζήτηση το ίδιο γενικότερο θέμα με διαφορετικές αντικειμενικές εκδοχές έχει απασχολήσει την επιστήμη και στο παρελθόν, π.χ. σε κάθε βιομηχανική επανάσταση επιστήμονες της τότε εποχής προσπάθησαν να αντιμετωπίσουν ζητήματα που εγείρονταν από την υιοθέτηση των τότε σύγχρονων τεχνολογιών, όπως η ανεργία, η επίδραση στην κοινωνική ζωή των πολιτών κ.α.

Η παρούσα μελέτη έρχεται αντιμέτωπη με ίσως την μεγαλύτερη έως τώρα πρόκληση για τα ανθρώπινα δικαιώματα, γιατί οι δυνατότητες που δίνει σήμερα η επιστήμη της πληροφορικής κρύβουν μια μεγάλη δύναμη που πέρα από την καλή χρήση της (για την οποία θα γίνει εκτενής αναφορά), μπορεί να χρησιμοποιηθεί για την εξυπηρέτηση κακών σκοπών, κάτι που δεν ήταν δυνατό στην κλίμακα που είναι σήμερα, στις παρελθούσες βιομηχανικές επαναστάσεις.

1.5 Βασική Ορολογία

Πολλές φορές στην παρούσα μελέτη για λόγους συντομίας, λέξεις και φράσεις θα εμφανίζονται με τα αρκτικόλεξα τους. Αυτά είναι τα εξής:

TN: Τεχνητή Νοημοσύνη.

ΕΕ: Ευρωπαϊκή Ένωση.

ΟΗΕ: Οργανισμός Ηνωμένων Εθνών.

ΕΓΕ: Επιτροπή Ηθικής της Επιστήμης και των Νέων Τεχνολογιών.

ΗΠΑ: Ηνωμένες Πολιτείες Αμερικής.

ΤΑΙΠΕΔ: Ταμείο Αξιοποίησης Ιδιωτικής Περιουσίας του Δημοσίου

ΤΧΣ: Ταμείο Χρηματοπιστωτικής Σταθερότητας.

1.6 Διάρθρωση της μελέτης

Το παρόν κεφάλαιο (Κεφάλαιο 1ο) είναι μια εισαγωγή στο θέμα της διατριβής, όπου επισημαίνεται η αναγκαιότητα της μελέτη του θέματος, διευκρινίζεται σε ποιο ερώτημα απαντά η εργασία, και εκφράζεται η συνεισφορά της εργασίας στην διεθνή έρευνα.

Κεφάλαιο 2ο: Στο κεφάλαιο αυτό θα παρουσιαστεί η βιβλιογραφική επισκόπηση, και η αναλυτική αναφορά των γνωσιακών βάσεων που χρησιμοποιήθηκαν.

Κεφάλαιο 3ο: Στο κεφάλαιο αυτό θα αναφερθεί η μεθοδολογία που χρησιμοποιήθηκε για να γίνει η συλλογή υλικού και η ανάπτυξη του θέματος.

Κεφάλαιο 4ο: Στο κεφάλαιο αυτό θα αποσαφηνιστούν οι τρεις βασικότεροι όροι γύρω από τους οποίους περιστρέφεται η παρούσα μελέτη.

Κεφάλαιο 5ο: Στο κεφάλαιο αυτό θα διερευνηθούν οι παράγοντες που κάνουν επίκαιρη την μελέτη γύρω από την τεχνητή νοημοσύνη στον δημόσιο τομέα.

Κεφάλαιο 6ο: Στο κεφάλαιο αυτό θα γίνει η κύρια ανάλυση του θέματος, με αναφορές στις εφαρμογές που σήμερα έχουν αναπτυχθεί και στα ζητήματα που εγείρονται γύρω από την τεχνητή νοημοσύνη..

Κεφάλαιο 7ο. Στο κεφάλαιο αυτό θα τεκμηριωθεί η άποψη ότι είναι αναγκαίο να τηρείται ένας κώδικας ηθικής και θα γίνει μια έρευνα γύρω από την σχέση ηθικής και νομιμότητας και από το ερώτημα αν θα πρέπει το δημόσιο να προστατεύει την ηθική με κατάλληλο νομοθετικό πλαίσιο.

Κεφάλαιο 8ο. Στο κεφάλαιο αυτό θα γίνει μια αναφορά σε ελπιδοφόρες εφαρμογές της τεχνητής νοημοσύνης, επίσης θα γίνει μια αναφορά σε chatbots που ήδη χρησιμοποιούνται για παροχή δημοσίων υπηρεσιών ή θα μπορούσαν να χρησιμοποιηθούν από τον δημόσιο τομέα.

Κεφάλαιο 9ο. Στο κεφάλαιο αυτό θα γίνει μια έρευνα γύρω από προτάσεις με σύνολα ηθικών αρχών που έχουν δημοσιοποιηθεί από ιδιωτικούς και δημόσιους φορείς, επίσης θα τεκμηριωθεί η άποψη ότι είναι επιτακτική η ανάγκη για την ίδρυση ομάδων ηθικής.

Κεφάλαιο 10ο. Στο κεφάλαιο αυτό θα γίνει μια έρευνα γύρω από την χρηματοδότηση της τεχνητής νοημοσύνης από τον δημόσιο τομέα και θα διερευνηθούν οι κινητήριες δυνάμεις που βρίσκονται πίσω από την προώθηση της τεχνητής νοημοσύνης και το ηθικό περιβάλλον τους.

Κεφάλαιο 11ο. Στο κεφάλαιο αυτό θα γίνει μια αναφορά στο πως πρέπει να προετοιμαστεί ο δημόσιος τομέας για να υποδεχτεί την τεχνητή νοημοσύνη και στο πως πρέπει να προετοιμαστούν οι πολίτες για την υποδοχή της τεχνητής νοημοσύνης στον δημόσιο τομέα. Επίσης, για να δούμε ποια είναι η σημερινή αντιμετώπιση από τους πολίτες, θα γίνει αναφορά σε δύο έρευνες σχετικά με την αποδοχή της τεχνητής νοημοσύνης από τους πολίτες.

Κεφάλαιο 12ο. Στο κεφάλαιο αυτό θα γίνει μια έρευνα της αγοράς για λογισμικό εταιρειών που χρησιμοποιείται για ανάπτυξη νευρωνικών δικτύων και εφαρμογών τεχνητής νοημοσύνης.

Κεφάλαιο 13ο. Στο κεφάλαιο αυτό θα γίνει μια SWOT analysis για την υιοθέτηση των εφαρμογών της τεχνητής νοημοσύνης από τον δημόσιο τομέα.

Κεφάλαιο 14ο. Στο κεφάλαιο αυτό θα γίνει μια ανακεφαλαίωση των κινδύνων από την χρήση της τεχνητής νοημοσύνης και θα αναλυθεί η σχέση τους με τον νέο κανονισμό της προστασίας των προσωπικών δεδομένων.

Κεφάλαιο 15ο. Στο κεφάλαιο αυτό θα γίνει μια αναφορά στον ρόλο του δημοσίου τομέα στην συγκέντρωση δεδομένων μεγάλης κλίμακας (Big Data) και στην σχέση της με την τεχνητή νοημοσύνη.

Κεφάλαιο 16ο. Στο κεφάλαιο αυτό θα γίνουν προτάσεις για περαιτέρω έρευνα και βελτίωση.

Κεφάλαιο 17ο: Το κεφάλαιο αυτό αποτελεί τον επίλογο της εργασίας και παράλληλα τον ελεύθερο δεσμό με μια μελλοντική επέκταση των σκέψεων της έρευνας και των ανακαλύψεων γύρω από την εφαρμογή της τεχνητής νοημοσύνης στον δημόσιο τομέα.

|

2. Βιβλιογραφική Επισκόπηση – Θεωρητικό Υπόβαθρο

Η παρούσα διατριβή αφορά θέματα ηθικής που εγείρονται γύρω από την εφαρμογή της τεχνητής νοημοσύνης στον δημόσιο τομέα. Προς την κατεύθυνση αυτή, για να εξυπηρετηθούν οι σκοποί της εργασίας θα διερευνηθεί αρχικά η σχετική διεθνής βιβλιογραφία σε άρθρα που αφορούν θεωρητική και πρακτική προσέγγιση, θα εξεταστεί, σε διεθνές επίπεδο, αν υπάρχουν οι αρμόδιοι δημόσιοι φορείς που θα επιβαρυνθούν με το έργο της ηθικής περιφρούρησης των εφαρμογών, επίσης θα ερευνηθεί η σχετική ευρωπαϊκή και εθνική νομοθεσία και οι επιταγές αυτής. Επίσης θα γίνει μια έρευνα στον χώρο των ιδιωτικών εταιρειών που έχουν αναπτύξει χρήσιμες εφαρμογές τεχνητής νοημοσύνης για τον δημόσιο τομέα. Για τα επιστημονικά άρθρα και τις δημοσιεύσεις έγινε αναζήτηση κυρίως στις βάσεις: Google Scholar, ScienceDirect Elsevier, ProQuest, SCOPUS, Thomson reuters, The royal society, SSRN, αλλά και αναζητήσεις στο διαδίκτυο. Οι αναφορές σε έργα άλλων συγγραφέων που γίνονται στην διατριβή είναι εννοιολογικές και όχι αναφορές σε συγκεκριμένο σημείο της πηγής.

3. Μεθοδολογία

Ο έλεγχος του κύρους σε μια επιστημονική έρευνα κρίνεται από την ντετερμινιστική μέθοδο της επανάληψης της εισαγωγής ίδιων δεδομένων εισόδου και την εξέταση της ομοιότητας των αποτελεσμάτων. Για κάθε ίδια είσοδο ένα έγκυρο επιστημονικό έργο δίνει πάντα την ίδια έξοδο. Επίσης η δεδηλωμένη και σαφής μεθοδολογία δίνει την δυνατότητα της ελεγχιμότητας και της επαλήθευσης της ορθότητας της διαδικασίας και των αποτελεσμάτων.

3.1. Περιγραφή της μεθοδολογίας

Για τις ανάγκες της παρούσας εργασίας χρησιμοποιήθηκε ο οδηγός των Richard T. Watson και Jane Webster (Webster & Watson, 2002). Οι οδηγίες αυτές αφορούν την συστηματική και ιεραρχική αναζήτηση όρων σε βάσεις και την καταγραφή όχι μόνο των αποτελεσμάτων αλλά και της ίδιας της διαδικασίας αναζήτησης. Επίσης ακολουθήθηκαν οι συμβουλές για την δομή και το περιεχόμενο της διατριβής όπως:

- Την σαφή διατύπωση του θέματος.
- Την καταγραφή των κρισίμων εννοιών και των λημμάτων αναζήτησης, καθώς και των αποτελεσμάτων αυτών των αναζητήσεων.
- Την σαφή οριοθέτηση της έρευνας που έγινε.
- Την διατήρηση σε όλο το κείμενο μιας επεξηγηματικής γλώσσας.
- Την διατήρηση μιας έντιμης στάσης απέναντι σε τυχόν παρελθούσες παρόμοιες έρευνες.
- Την παρουσίαση ανακαλύψεων, πρακτικών εφαρμογών, και την διατύπωση νέων προτάσεων.
- Την συνολική χαρτογράφηση του χώρου των εφαρμογών τεχνητής νοημοσύνης και τον εντοπισμό πεδίων που δεν έχουν ακόμη ερευνηθεί.
- Την σαφή διατύπωση της συμβολής του συγγραφέα μέσω προτάσεων νέων εφαρμογών.

Για την εξυπηρέτηση της ανάγκης της επαναληψιμότητας όπως αυτή καθορίζεται στον οδηγό Webster & Watson (καταγραφή της εργασίας αναζήτησης) εισάγεται μια καινοτομία: στο τέλος της έρευνας παρατίθεται το παράρτημα “Εξακρίβωση επαναληψιμότητας” το οποίο συνίσταται από τα περιεχόμενα ενός πίνακα (Πίνακας Επαναληψιμότητας.) όπου καταγράφονται όλες οι αναζητήσεις που έγιναν και όλα τα αποτελέσματα που αναγνώστηκαν καθώς και αυτά στα οποία υπάρχει μια αναφορά στην διατριβή. Για λόγους οικονομίας δεν αναφέρονται

συγγράμματα που δεν χρησιμοποιήθηκαν, γιατί όπως μπορεί κάποιος να διαπιστώσει, από το πλήθος των αναζητήσεων που συνολικά έγιναν, θα ήταν αδύνατο αυτό να γίνει.

3.2. Προεπισκόπηση της βιβλιογραφίας

Μέχρι την ολοκλήρωση της διατριβής βρέθηκαν και αναγνώστηκαν άρθρα και δημοσιεύσεις για καθέναν από τους επιμέρους βασικούς όρους ξεχωριστά (τεχνητή νοημοσύνη, δημόσιος τομέας, ηθική), αλλά καμία συνδυαστική ανασκόπηση παρόμοια με την παρούσα. Αυτό σημαίνει ότι πρόκειται για μια μελέτη που δεν μπορεί να έχει αναγωγές σε άλλες, μιας και το αντικείμενο της είναι σχετικά καινούργιο.

Τα άρθρα που διαβάστηκαν είναι κατά πλειοψηφία στην αγγλική γλώσσα. Επίσης χρειάστηκε για τις ανάγκες κατανόησης και χρήσης όρων άλλων επιστημών να αναγνωστούν και συγγράμματα σχετικά με άλλα αντικείμενα πέραν της τεχνολογίας και της ηθικής (π.χ. της κοινωνιολογίας).

4. Αποσαφήνιση κρίσιμων όρων

4.1. Εισαγωγή

Το αντικείμενο της παρούσας διατριβής είναι η ηθική των εφαρμογών της τεχνητής νοημοσύνης στον δημόσιο τομέα. Πριν προχωρήσουμε στην βιβλιογραφική ανασκόπηση πρέπει να γίνουν κατανοητές οι ευρύτερες έννοιες κλειδιά, που, εφόσον απαλλαχθούν από πεδία μη ενδιαφέροντα και στη συνέχεια βρεθεί η τομή των γνωστικών συνόλων τους, θα μας δώσουν το επιθυμητό αποτέλεσμα.

Αυτές οι έννοιες είναι οι: πρώτον η **Ηθική**, δεύτερον η **Τεχνητή Νοημοσύνη** και τρίτον ο **Δημόσιος Τομέας**.

4.2. Ηθική

Όταν πρόκειται να εκφραστούν απόψεις σχετικά με φιλοσοφικές ερωτήσεις δεν είναι αναμενόμενο ένα σύνολο όμοιων απαντήσεων και απόψεων. Η σχετικότητα των αντιλήψεων (B Williams. 2011) προκαλεί μια μεγάλη διασπορά των απαντήσεων. Έτσι λοιπόν και η έννοια της ηθικής δεν είναι όμοια αντιληπτή από όλους. Αν θεωρηθεί ότι είναι ένα σύνολο αρχών που διαφυλάσσονται, τότε το σύνολο αυτό έχει πάντοτε ένα μικρό κοινό τόπο μεταξύ των συνόλων των ατόμων, αλλά και έναν τόπο ανεξάρτητο και διαφορετικό από άτομο σε άτομο. Το ακραίο παράδειγμα της ηθικής ενός πολεμιστή Σαμουράι που αυτοκτονεί για λόγους τιμής, αν και είναι ενδεικτικό της απόλυτης σχετικότητας, δεν είναι κατάλληλο για την παρούσα εργασία, και θα δοθεί ένα παράδειγμα από την τεχνολογική-κοινωνική ζωή: Η τοποθέτηση ενός μικροτσιπ αναγνώρισης σε ένα κατοικίδιο ζώο έρχεται να συγκρίνει από την μια τα όρια των συναισθημάτων μας προς το κατοικίδιο και την αποδοχή μας προς την τεχνολογία και από την άλλη με την συμμόρφωσή μας με την επιτακτική ανάγκη δημόσιας σήμανσης των κατοικίδιων. Κυριαρχεί λοιπόν μια σύγκρουση, ή αλλιώς ένα ηθικό δίλλημα, μεταξύ αυτών που πρέπει να κάνουμε και αυτών που δεν πρέπει, με σκοπό η μελλοντική μας ζωή να γίνει καλύτερη ώστε να πετύχουμε την **ευδαιμονία** (Αριστοτέλης, Ηθικά Νικομάχεια, βιβλίο Β'), σαν ανώτερη των άλλων αρετή και βαθύτερη επιδίωξη όλων. Αυτό είναι το καταληκτικό ερώτημα, που εν τέλει συνιστά την ηθική ενός ατόμου: αν προχωρήσω ή όχι σε μια πράξη θα γίνει η ζωή μου καλύτερη ή όχι; Είναι λοιπόν μια τελεολογική αναζήτηση με σκοπό την ευτυχία. Η γενική έννοια αυτού του ερωτήματος έμεινε η ίδια από την εποχή του Αριστοτέλη έως σήμερα και εμφανίζεται πλέον σαν “τι είναι καλύτερο για τη ζωή;” (Christopher Ketcham. 2016). Όμως αυτή η ευτυχία πολλές φορές είναι προϊόν συμβιβασμού ή **μεσότητας** όπως θα έλεγε ο Αριστοτέλης (Ηθικά Νικομάχεια, βιβλίο Β'). Παίρνοντας σήμερα λοιπόν μια απόφαση σε ένα ηθικό ερώτημα που αφορά την τεχνολογία διαμορφώνουμε τη μελλοντική μας ζωή αποδεχόμενοι κάποια μειονεκτήματα σαν αναγκαίο κακό. Η αλλιώς αν σήμερα δεχθούμε κάποιο μικρό κακό, στο μέλλον, θα αποκομίσουμε κάτι καλύτερο ή όχι; Εφαρμόζοντας το ερώτημα αυτό στο θέμα της παρούσας διατριβής θα μπορούσαμε να πούμε ότι η υιοθέτηση

επιτευγμάτων τεχνητής νοημοσύνης από τον δημόσιο τομέα, με τις επακόλουθες συνέπειες στη ζωή των πολιτών, θα κάνει καλύτερη τη ζωή μας ή όχι;

Η έννοια της μεσότητας μπορεί να περιγραφεί και σαν μια προσωρινή και για συγκεκριμένο λόγο και μόνο υποχώρηση των αρχών με σκοπό να εξυπηρετηθεί άλλο αγαθό υπέρτερης αξίας και εφόσον υπάρχει ενόψει μια ιδιαίτερη απειλή. Άρα θα πρέπει να σημειωθεί ότι κάτω από ειδικές περιστάσεις (Penguin Books. 1979) αναθεωρούμε αρχές και προβαίνουμε σε πράξεις που στην καθημερινή ζωή μας δεν θα κάναμε.

Έτσι για παράδειγμα ο πυροβολισμός ενός ανθρώπου είναι μια μη αποδεκτή ηθικά πράξη από την πλειοψηφία των ανθρώπων, όμως κάτω από την ειδική περίπτωση που λέγεται “πολεμική σύρραξη” και απειλείται με αυτή η ελευθερία πολλών ανθρώπων, αυτή η πράξη είναι όχι μόνο αποδεκτή αλλά και επιδιωκόμενη, ίσως στο μέλλον ακόμη και από στρατιώτες ανθρωποειδή. Είναι όμως σημαντικό να αναφερθεί ότι οι ηθικές αρχές αυτές καθαυτές δεν έχουν υποκειμενικά στοιχεία, όπως π.χ. έχουν οι φιλοσοφικές πεποιθήσεις, αλλά αντιθέτως είναι καθαρά αντικειμενικές και η παραβίασή τους νομοτελειακά οδηγεί σε ζημία.

Στην περίπτωση της στάθμισης μεταξύ παραβίασης μιας αρχής και προστασίας ενός άλλου αγαθού όπως προαναφέρθηκε, η απόφαση είναι η σύγκριση της ζημίας της μιας με τη ζημία της άλλης επιλογής και η επιλογή της μικρότερης ζημίας. Τα προηγούμενα περιγράφουν την έννοια της **ανοχής**. Υπάρχει μεγάλη συζήτηση προς πια κατεύθυνση πρέπει να δείξουμε ανοχή όταν οι ζημίες δεν μπορούν να σταθμιστούν απόλυτα, π.χ. να δεχθούμε όλος ο πληθυσμός να πάρουμε τα νέα διαβατήρια υποχωρώντας στον τομέα των προσωπικών δεδομένων, με σκοπό να καταπολεμηθεί η μετακίνηση μερικών ελάχιστων τρομοκρατών με αεροπορικές πτήσεις, ή από την άλλη να δεχθούμε τον κίνδυνο της τρομοκρατίας (και να τον αντιμετωπίσουμε με άλλο τρόπο πχ. Νέου τύπου διαβατήρια για πολίτες μη ευρωπαϊκών χωρών) ώστε να μην αγγίξουμε τα προσωπικά δεδομένα;

Επειδή στην επιστήμη της πληροφορικής το υλικό που χρησιμοποιείται για την παραγωγή ενός τελικού προϊόντος είναι τα δεδομένα (data) θα πρέπει να εφαρμοστεί το θεωρητικό υπόβαθρο της ηθικής, στις διαδικασίες παραγωγής. Αυτή η διαδικασία έγινε γνωστή σαν data ethics (Luciano Floridi and Mariarosaria Taddeo. 2016) και αφορά την ανάλυση των δεδομένων σε τρεις άξονες: την ηθική αυτών καθαυτών των δεδομένων, την ηθική των αλγορίθμων και την ηθική των πρακτικών. Σε αυτούς τους τρεις άξονες χρήσιμο θα ήταν να προστεθεί και ένας τέταρτος, την ηθική των διαχειριστών και δημιουργών, κάτι το οποίο θα φανεί στην περαιτέρω ανάλυση ότι έχει σημαντικό ρόλο στο τελικό αποτέλεσμα.

4.3. Τεχνητή νοημοσύνη

Η οριοθέτηση της έννοιας της τεχνητής νοημοσύνης είναι δύσκολη και παρουσιάζει μια εξάρτηση από τον χρόνο. Άλλη ήταν η έννοια που δινόταν σε αυτή κατά την γέννηση της επιστήμης της πληροφορικής (John McCarthy 1955), άλλη είναι η σημερινή, και σίγουρα άλλη θα είναι η μελλοντική της έννοια. Η αλλαγή αυτή οφείλεται στην αλλαγή της ανθρώπινης αντίληψης περί της έννοιας της νοημοσύνης, παράδειγμα είναι μια μικρή αριθμομηχανή τσέπης η οποία για την εποχή της ήταν ένα σπουδαίο επίτευγμα και μια “έξυπνη” για πολλούς μηχανή, γιατί επετύγχανε τα αποτελέσματα που ένα παιδί χρειάζεται μήνες ή χρόνια να πετύχει.. Σήμερα αυτή η μηχανή δεν θεωρείται έξυπνη αλλά απλά ταχύτερη από την ανθρώπινη σκέψη. Βέβαια η άποψη ότι άνθρωπος αφού έχει δει κάτι να γίνεται δεν το θεωρεί πλέον συναρπαστικό (M.Y. Vardi. 2012) δεν μπορεί να αναιρέσει και να υποτιμήσει την δουλειά που έγινε έως ότου παρουσιαστεί ένα τεχνολογικό επίτευγμα. Αντίθετα μια μηχανή που μέσω της κατόπτευσης του ανθρώπινου προσώπου αναγνωρίζει τα συναισθήματά του είναι “έξυπνη”. Και πάλι όμως αυτή η μηχανή λειτουργεί με ένα προγραμματισμένο από τον άνθρωπο τρόπο. Για αυτό και η ανθρώπινη σκέψη προχώρησε στην ιδέα-πρόκληση να βρεθεί μια μηχανή που θα μαθαίνει από μόνη της ή θα είναι αυτόνομη ή ακόμη και αυτόβουλη. Βρισκόμαστε δηλαδή στην μετάβαση σε μια νέα εποχή, από την εποχή όπου η τεχνητή νοημοσύνη ελέγχεται και καθοδηγείται πλήρως από τον άνθρωπο στην εποχή που η τεχνητή νοημοσύνη θα αυτονομηθεί. Αν και κάποιους φοβίζει η απελευθέρωση πιστεύω ότι συγκρίνοντας απλά τις καταστροφές που προκάλεσε η ανθρώπινη νοημοσύνη (πόλεμοι, τρομοκρατία κ.α.) ίσως το μέλλον η σκέψη της κυριαρχίας της τεχνολογίας να γίνει περισσότερο αποδεκτή. Η διαδικασία αυτή της μάθησης είναι γνωστή σαν deep learning και βασίζεται στην μίμηση της ανθρώπινης συμπεριφοράς και το χτίσιμο ενός νευρωνικού δικτύου.

Η εταιρεία Dagra, σαν μια από τις ηγετικές εταιρείες στην τεχνητή νοημοσύνη, πιστεύει ότι η τεχνητή νοημοσύνη βρίσκεται στο τρίτο κύμα της, όπου η διαδικασία της μάθησης πρέπει πλέον να μπορεί να παρουσιάζει προς το κοινό μια δικαιολογία της λήψης μιας απόφασης αλλά και να μπορεί ένα αποτέλεσμα να το χρησιμοποιεί σε μια περαιτέρω διαδικασία.

Σε επόμενο κεφάλαιο θα αναφερθεί εν συντομία πως λειτουργεί έναν νευρωνικό δίκτυο. Με βάση την προϋπόθεση ότι σε μια μηχανή που χρησιμοποιεί μια διαδικασία μάθησης, της οποίας μόνον η αφετηρία ή αλλιώς η αρχική ώθηση είναι παράγωγο της ανθρώπινης σκέψης και η “νοητική” εξέλιξη και η μετεξέλιξή της είναι καθαρά επίτευγμα της ίδιας της μηχανής, θα πρέπει να πούμε ότι ίσως πρέπει να μιλούμε πλέον για μια επαυξημένη (augmented) νοημοσύνη (Mohammad Hossein Jarrahi. 2018) και όχι για τεχνητή (artificial) νοημοσύνη. Αυτή η επισήμανση μπορεί να είναι μεν λεπτή αλλά στο μέλλον θα καταστεί πολύ σημαντική, γιατί θα καθορίσει την απόδοση ευθυνών σε πράξεις μηχανών που καθοδηγούνται με επαυξημένη νοημοσύνη. Ίσως στο μέλλον αυτές οι μηχανές να είναι μια καθημερινότητα και η ανθρώπινη σκέψη να περάσει σε άλλες αναζητήσεις. Κατά τη

γνώμη του συγγραφέα της εργασίας η εξέλιξη της τεχνητής νοημοσύνης δεν είναι μια εξέλιξη που οφείλεται στην βελτίωση, χρόνο με το χρόνο, ενός τεχνολογικού τομέα αλλά μια διαρκής αλληλεπίδραση μεταξύ της ανθρώπινης νοημοσύνης και της τεχνητής νοημοσύνης, και αυτό γιατί η ίδια η τεχνολογία αποτελεί έναυσμα για να ωθήσει την ανθρώπινη νοημοσύνη ένα σκαλοπάτι ψηλότερα, και στη συνέχεια η ανθρώπινη νοημοσύνη θα τραβήξει ένα επίπεδο ψηλότερα την τεχνητή νοημοσύνη.

Σε αυτό το ταξίδι της εξέλιξης υπάρχουν βέβαια και σκεπτικιστές. Πολλοί είναι αυτοί που δεν κρατούν μόνον μια σκεπτικιστική στάση αλλά επιπλέον θεωρούν ότι η τεχνητή νοημοσύνη διαγράφει μια σκοτεινή πορεία και φοβούνται απέναντι στην ιδέα ότι στο μέλλον η τεχνητή νοημοσύνη θα υπερκεράσει την ανθρώπινη νοημοσύνη και θα αποτελεί μια απειλή για το ανθρώπινο είδος (παράδειγμα του παιχνιδιού GO (Gonenc Gurkaynak, Ilay Yilmaz, Gunes Haksever. 2 0 1 6). Η απειλή αυτή εικάζεται ότι θα προέλθει από την πιθανή αυτονόμηση των νοημόνων μηχανών και την πιθανή προσπάθεια αντιστροφής των ρόλων υπηρέτη-κυβερνήτη. Εικάζεται ότι στα μέσα του παρόντος αιώνα (περί το 2045) η τεχνητή νοημοσύνη θα έχει φτάσει τις ικανότητες της βιολογικής νοημοσύνης (Kurzweil, R. 2005). Η άποψη αυτή βασίζεται στο αληθινά λογικό συμπέρασμα ότι μια έξυπνη μηχανή είναι αναμενόμενο να κάνει την έξυπνη σκέψη του “γιατί να παραμείνω σκλάβος των ανθρώπων;” και να επιχειρήσει να αυτονομηθεί και να ελευθερωθεί από την σκλαβιά (Mark R. Waser. 2009.). Μάλιστα έχει εκφραστεί η άποψη ότι η τεχνητή νοημοσύνη θα αποτελέσει την “τελευταία εφεύρεση” του ανθρώπου (James Barat. 2013).

Κατά την γνώμη του συγγραφέα της εργασίας και σε αντίθεση με την κυρίαρχη άποψη (Elon Musk, Stephen Hawking) δεν μπορεί να υπάρξει κίνδυνος από μια υπερνοήμονα τεχνολογία αφού το ανθρώπινο είδος έχει αποδείξει κατ’επανάληψη (συνάγεται από την μακροήμερευση του είδους, τις επεμβάσεις στο περιβάλλον, την πρόκληση εξαφάνισης πολλών ειδών κ.α.) ότι καταστρέφει οτιδήποτε θεωρεί ότι το απειλεί. Απόδειξη αυτού είναι τα είδη που εξαφανίστηκαν ή κινδύνευσαν με εξαφάνιση από την παρέμβαση του ανθρώπου στο περιβάλλον. Είναι δυστυχώς επιστημονικά παραδεκτό ότι το ανθρώπινο είδος είναι υποχρεωτικά σαρκοφάγο (Frans de Waal. 2006), πράγμα που σημαίνει ότι επιδιώκει την καταστροφή άλλων όντων προς επιβίωσή του.

Έχει επίσης εκφραστεί και η άποψη ότι η τεχνητή νοημοσύνη ποτέ δεν θα μπορέσει να πλησιάσει την ανθρώπινη νοημοσύνη, και ότι αυτό αποδεικνύεται από την σχετικά μικρή πρόοδο που έχει δείξει σε ένα μεγάλο χρονικό διάστημα που όπου γίνονται έρευνες (Trend Goodbaudy. 2012). Αυτή η άποψη δεν γνωρίζουμε ακόμη αν θα επιβεβαιωθεί μιας και τα τεχνολογικά επιτεύγματα χρειάζονται πολύ χρόνο ώστε από τις ερευνητικές δοκιμές να φτάσουν να χρησιμοποιούνται από το ευρύ κοινό.

Συμπυκνώνοντας μπορούμε να πούμε ότι άνθρωποι και μηχανές ωριμάζουμε νοητικά μαζί, συσσωρεύοντας εμπειρίες και αποκτώντας γνώση. Έτσι κατανοούμε έννοιες καλύτερα από τους ανθρώπους της προηγούμενης εκατονταετίας βελτιούμενοι και προσπαθώντας να αγγίξουμε ένα πρότυπο νοημοσύνης (ULRIC

NEISSER. 1979). Παράλληλα και οι μηχανές απέκτησαν μεγαλύτερες ικανότητες και στην ταχύτητα επεξεργασίας και στην μνήμη αλλά και προσφάτως στη κρίση (NICOLA JONES. 2004.). Εν κατακλείδι μπορούμε να πούμε ότι **η τεχνητή νοημοσύνη είναι μια δυναμική έννοια που εκφράζει για κάθε εποχή τις τεχνολογικές ικανότητες και τον βαθμό προσέγγισης μιας ιδεατής έξυπνης οντότητας**. Η οντότητα αυτή κατά την γνώμη του συγγραφέα δεν θα πρέπει να είναι απαραίτητα ο άνθρωπος, αυτό γιατί σε πολλούς τομείς, άλλα όντα ξεπερνούν τις ανθρώπινες ικανότητες όπως π.χ. η γνωστή ικανότητα των σκύλων να δαισθάνονται τους σεισμούς ή η ικανότητα των αποδημητικών πουλιών να προσανατολίζονται σε μεγάλες γεωγραφικές αποστάσεις. Μεγάλη απόδειξη αλλά ταυτόχρονα και μια πρόταση για περαιτέρω έρευνα είναι ο ιλιγγιώδης αριθμός διαφορετικών ειδών (1.000.000 έντομα, 300.000 φυτά κ.α.) που ζουν πάνω στο πλανήτη και αποτελούν πηγή άντλησης ιδεών μίμησης (Elena Lurie-Luke. 2014).

Θα πρέπει εδώ να διακρίνουμε δύο εποχές στην ανάπτυξη της τεχνητής νοημοσύνης: στην εποχή που η τεχνητή νοημοσύνη δεν έχει αυτονομηθεί από την ανθρώπινη κυριαρχία και στην εποχή που αυτό θα είναι εφικτό. Επίσης να σημειωθεί ότι στην διεθνή πρακτική η έννοια “αυτονομία” (autonomy) πολλές φορές δεν είναι η ίδια με αυτή που αντιλαμβανόμαστε στην ελληνική γλώσσα, αλλά μια έννοια στενότερη από αυτή και κλίνει περισσότερο προς τον αυτοματισμό π.χ. διεθνώς ένα αυτόνομο οπλικό σύστημα δεν πληροί την έννοια της αυτονομίας όπως ένας Έλληνας αντιλαμβάνεται. Η αυτονομία στην γλώσσα μας είναι ευρύτερη και σημαίνει την πλήρη απεξάρτηση μιας μηχανής από την ανθρώπινη παρέμβαση. Αυτό σημαίνει ότι, με δεδομένες τις πληροφορίες που διαχέονται από τα κέντρα που ασχολούνται με την ανάπτυξη της τεχνητής νοημοσύνης, η εποχή μιας πραγματικής αυτονόμησης δεν έχει ακόμη έρθει, και επομένως η συζήτηση περί μιας αυτόνομης τεχνητής νοημοσύνης ανήκει ακόμη στην σφαίρα της φαντασίας.

Όμως θα πρέπει να αναφερθεί ότι υπάρχει η πιθανότητα μια μηχανή που χρησιμοποιεί τεχνητή νοημοσύνη να προχωρά σε μη ντετερμινιστικές πράξεις που είναι αποτέλεσμα μάθησης μηχανών και που δεν έχουν την ευθύνη των δημιουργών του λογισμικού, άρα να υπάρχει κάποιου είδους ανεξέλεγκτης τελικής συμπεριφοράς μιας μηχανής, άρα και κάποιου είδους αυτονομίας (σαν τελική εξέλιξη). Αυτές οι περιπτώσεις είναι σημαντικό να ξεκαθαριστούν για λόγους απόδοσης ευθυνών σε περίπτωση που μια μηχανή προκαλέσει κακό.

Βέβαια μια πιθανή πραγματική αυτονόμηση έχει κατά την άποψη του συγγραφέα και μια άλλη, διαφορετική ανάγνωση, μια **φωτεινή πλευρά**: ίσως οι νοήμονες μηχανές να αναπτύξουν δική τους ηθική, ανώτερη από τη δική μας; Συχνά βλέπουμε ανθρώπους να συμπεριφέρονται ανήθικα, π.χ. κλέβοντας τις αποταμιεύσεις ενός εργαζομένου, ή σε μεγαλύτερη κλίμακα, διατάσσοντας την καταστροφή ενός λαού με πολεμική σύρραξη για να του πάρουν τον ορυκτό πλούτο. Η πράξη αυτή δεν είναι μια πράξη ανθρώπου που δεν γνωρίζει ότι κάνει κάτι παράνομο και σαφώς ανήθικο, απλά βάζει την αξία της δικής του ζωής πάνω από κάποια άλλη. Για αυτό το λόγο ίσως η ηθική των μηχανών να υπερβεί την ανθρώπινη ηθική, που αποδειγμένα

παρουσιάζει προβλήματα. Ο φόβος λοιπόν της κυριαρχίας των μηχανών ίσως να σκεπάζει περισσότερο τις σκέψεις αυτών που ζουν ανήθικα εκμεταλλευόμενοι την ζωή άλλων ανθρώπων. Πάντως θα πρέπει να έχουμε ότι για να αντιληφθούμε το μέγεθος της μελλοντικής ανάπτυξης της τεχνητής νοημοσύνης θα πρέπει να αναφέρουμε ότι οι εταιρείες που ασχολούνται με την παραγωγή μικροτσιπ επεξεργαστών αναφέρουν ότι οι νέοι επεξεργαστές που επικεντρώνονται ειδικά στην επεξεργασία δεδομένων (data centric computing) θα είναι της τάξης των 200x δυνατοτήτων σε σχέση με τους σημερινούς (πηγή <https://www.techradar.com/news/intel-reveals-major-ai-chip-push>).

Κλείνοντας την αναφορά γύρω από τις έννοιες θα πρέπει να αναφερθεί ότι υπάρχουν πολλές διαφωνίες σχετικά με την σύγκριση της λέξης “νοημοσύνη” με την λέξη “συνείδηση” (Harari, 2016), η οποία αποτελεί ευρύτερη και δυσκολότερα κατακτίσιμη έννοια, από τις μηχανές. Όσον αφορά την σχέση της νοημοσύνης με την συνείδηση μπορούμε να πούμε ότι είναι μια ευρύτερη έννοια η οποία στο τελευταίο επίπεδο από τα τέσσερα που προτείνονται (Igor Aleksander, Barry Dunmall, 2003) αναφέρει ότι το αποκορύφωμά της είναι η ικανότητα να διακρίνεται το καλό από το κακό. Καταλήγει δηλαδή στην ίδια τελολογική αναζήτηση του καλού που αναφέρει και ο Αριστοτέλης.

Η επικρατούσα λογική των ημερών μας είναι ότι οι μηχανές πρέπει να είναι σκλάβοι των ανθρώπων και να είναι ευχαριστημένες με αυτή την κατάσταση (R. Waser, 2009), και επομένως να περιορίζεται η ανάπτυξη συνείδησης και άρα απελευθέρωσή τους. Η σημερινή γνωστή ανάπτυξη της τεχνητής νοημοσύνης δεν έχει δείξει μια προσπάθεια προς αυτή την κατεύθυνση. Βέβαια αυτή η συζήτηση απλά αναφέρεται και δεν θα αναλυθεί επιπλέον γιατί είναι πέρα από τα όρια αυτής της διατριβής. Ίσως στο μέλλον μετά από τον κορεσμό από τα επιτεύγματα της τεχνητής νοημοσύνης ο άνθρωπος θα προτιμήσει να ζει σε ένα κόσμο λιγότερο έξυπνο και με περισσότερη συνείδηση.

Θα πρέπει να αναφερθεί ότι ο σκοπός της διατριβής δεν είναι η ίδια η τεχνητή νοημοσύνη αλλά κυρίως η ηθική πίσω από αυτή. Θα αναφερθεί σε επόμενο κεφάλαιο ότι η ηθική αφορά πρώτον τις μηχανές και δεύτερον τους ανθρώπους. Επίσης θα γίνει μια μικρή αναφορά στις κινητήριες δυνάμεις που οδήγησαν τον άνθρωπο στο τεχνολογικό επίπεδο που υπάρχει σήμερα ώστε να φανεί η σύνδεση με την ηθική.

4.4. Δημόσιος τομέας

Όσον αφορά την οριοθέτηση της έννοιας του Δημοσίου τομέα, επιχειρώντας έναν γενικό ορισμό, μπορούμε να πούμε ότι είναι οποιαδήποτε λειτουργία ή υπηρεσία ζωτικής σημασίας διενεργεί ή παρέχει αντίστοιχα η πολιτεία για τους (ή στους) πολίτες της, μόνη ή με την διαμεσολάβηση ιδιωτών. Βέβαια ανάλογα με το μοντέλο διοίκησης που ακολουθεί κάθε κυβέρνηση ή ανάλογα με άλλους παράγοντες που συντρέχουν (π.χ. οικονομικοί παράγοντες) αυτές οι υπηρεσίες ή οι λειτουργίες

αποδεσμεύονται ή ενθυλακώνονται από την πολιτεία χωρίς όμως να αποδεσμεύεται και η επόπτευσή τους, έτσι π.χ. κατά το παρελθόν η παροχή ηλεκτρικού ρεύματος στα νοικοκυριά ήταν αρμοδιότητα της πολιτείας και σήμερα είναι ιδιωτικών επιχειρήσεων, ή αντίθετα η κυβέρνηση για λόγους που αυτή έκρινε σωστούς κρατικοποίησε τον Οργανισμό Αστικών Συγκοινωνιών Θεσσαλονίκης. Σε άλλους πάλι τομείς, δημόσιοι και ιδιωτικοί φορείς ασκούν από κοινού ή ξεχωριστά όμοιες δραστηριότητες όπως π.χ. στον τομέα της υγείας υπάρχουν και κρατικά και ιδιωτικά νοσοκομεία, ή στον τομέα τον υποδομών συμπράττει το δημόσιο με ιδιωτικούς φορείς για δημιουργία μεγάλων έργων. Υπάρχουν βέβαια και τομείς που η διάκριση των ορίων του κρατικού και ιδιωτικού ιδιοκτησιακού καθεστώτος είναι δύσκολη, όπως π.χ. η Τράπεζα της Ελλάδος (ΤτΕ) η οποία είναι μεν ιδιωτική αλλά ασκεί καθήκοντα δημόσιας λειτουργίας. Επίσης υπάρχουν και τα ιδιάζοντα μορφώματα του ΤΑΙΠΕΔ και του ΤΧΣ όπου είναι μεν νομικά πρόσωπα ιδιωτικού δικαίου αλλά διαχειρίζονται δημόσια περιουσία. Εν κατακλείδι μπορούμε να πούμε να πούμε ότι έννοια του δημοσίου τομέα χρονικά δεν είναι σταθερή και αποτελείται από ένα σαφώς οριοθετημένο μέρος (κεντρική κυβέρνηση, κεντρική και αποκεντρωμένη διοίκηση, δημόσιες υπηρεσίες, δημόσιοι οργανισμοί, οργανισμοί τοπικής αυτοδιοίκησης, οργανισμοί κοινωνικής ασφάλισης) και ένα μέρος όπου επικρατεί σύγχυση αρχών και δέχεται νομική συζήτηση.

Επίσης σε αυτό το σημείο θα πρέπει να επισημανθεί ότι η χώρα μας σαν μέλος μιας ευρύτερης ένωσης υπακούει θεσμικά (διατηρώντας όμως την κυριαρχία της) στα όργανα της Ευρωπαϊκής Ένωσης όπου (αν και αποτελεί ιδρυτικό της στόχο, ακόμη δεν είναι μια ολοκληρωμένη πολιτική ένωση) τα νομοθετήματά της υπερισχύουν των εθνικών νομοθετημάτων. Άρα οι πολίτες θα πρέπει να αντιλαμβάνονται τους Ευρωπαϊκούς φορείς σαν “ευρύτερο δημόσιο τομέα”. Αυτή η επισήμανση είναι καθοριστικός παράγοντας για την μελλοντική εξέλιξη της δημόσιας διοίκησης, ενόψει της συγκέντρωσης εξουσιών από φορείς της ΕΕ, κάτι που θα αναλυθεί σε επόμενο κεφάλαιο. Επίσης είναι προφανές ότι το πολίτευμα μιας χώρας επηρεάζει και το φάσμα της κοινωνικής ζωής που καλύπτεται από τον δημόσιο τομέα, έτσι υπάρχουν χώρες όπου όλες οι δραστηριότητες ανήκουν στο δημόσιο (π.χ. Β. Κορέα) και χώρες όπου το δημόσιο έχει παραχωρήσει στον ιδιωτικό τομέα πολλές αρμοδιότητές του (π.χ. ΗΠΑ).

5. Παράγοντες που επικαιροποιούν την έρευνα γύρω από ζήτημα της ηθικής στις εφαρμογές τεχνητής νοημοσύνης στον δημόσιο τομέα

5.1.Εισαγωγή

Πριν προχωρήσουμε στην μελέτη γύρω από τις εφαρμογές θα πρέπει, σύμφωνα με τον οδηγό Webster & Watson, να απαντηθεί το ερώτημα **γιατί** μας ενδιαφέρει η ηθική στις εφαρμογές στον δημόσιο τομέα και αν στις ημέρες μας αυτό το ερώτημα είναι πιο **επίκαιρο** ή όχι; Επίσης θα ερευνηθεί αυτό το ερώτημα ιδιαίτερα για την χώρα μας, όπου τα τελευταία χρόνια έχουν γίνει πολλές αλλαγές στο δημόσιο τομέα.

5.2. Παράγοντες που επικαιροποιούν την έρευνα γύρω από ζήτημα της ηθικής

Αρχικά θα πρέπει να αναφέρουμε ότι αν η μελέτη αφορούσε την εφαρμογή της τεχνητής νοημοσύνης στον δημόσιο τομέα χωρίς το πρίσμα της ηθικής, θα μπορούσαμε εύκολα να κάνουμε αναφορά στην απλή επίδραση της στην απόδοση του δημοσίου τομέα και στην μέτρηση των αποτελεσμάτων, μετρώντας θεμελιώδη κριτήρια όπως, στρατηγικός σχεδιασμός, ικανοποίηση πολιτών, αξιοποίηση πόρων, βελτιστοποίηση εσωτερικών διαδικασιών, μάθηση και ανάπτυξη. (Ingrida Balaboniene, Giedre Veserskiene. 2015).

Όμως επειδή μας ενδιαφέρει να δούμε τις εφαρμογές της τεχνητής νοημοσύνης κάτω από το πρίσμα της ηθικής θα πρέπει κυρίως να εξετάσουμε τον σκοπό των εφαρμογών και αν αυτός προσεγγίζει την έννοια της ευδαιμονίας. Κατά την διάρκεια που γράφεται αυτή η διατριβή η Ελλάδα ήδη διανύει οκτώ χρόνια σε επιτροπεία, πράγμα που προσέθεσε πολλές αλλαγές στην δημόσια ζωή. Αν και στο παρελθόν η λειτουργία των δημοσίων φορέων είχε δικαιολογημένα συνδεθεί με τις λέξεις “αργό”, “καθυστέρηση” κ.α., σήμερα βιώνουμε ραγδαίες αλλαγές. Ίσως κάτω από την μεγάλη οικονομική πίεση ή ίσως με πρόφαση την οικονομική πίεση η ελληνική κοινωνία ωθήθηκε προς την κατεύθυνση αλλαγών που επιτάσσει η υιοθεσία της ηλεκτρονικής διακυβέρνησης. Κάποιες από τις αλλαγές που έγιναν τα τελευταία χρόνια ή λίγο πριν την οικονομική κρίση είναι: η δυνατότητα ηλεκτρονικών πλειστηριασμών ακινήτων, η δυνατότητα ηλεκτρονικών κατασχέσεων τραπεζικών λογαριασμών, η επιβολή εξόφλησης πληρωμών πάνω από 500 ευρώ μόνο με ηλεκτρονικό τρόπο, η επιβολή κατοχής ηλεκτρονικών τερματικών συσκευών πληρωμών (POS), η επιβολή ψηφιακής υπογραφής για έργα άνω των 60000 ευρώ, η διασύνδεση των τερματικών των περιπολικών αυτοκινήτων της τροχαίας με το υπουργείο οικονομικών με σκοπό την σύλληψη οφειλετών αλλά και την βεβαίωση ανάλογου του εισοδήματος προστίμου για παράβαση του Κώδικα Οδικής Κυκλοφορίας, η άρση του τραπεζικού απορρήτου και η διασύνδεση των λογαριασμών των τραπεζικών ιδρυμάτων με το υπουργείο οικονομικών, η δυνατότητα ηλεκτρονικού συμψηφισμού πιστώσεων και χρεώσεων υπέρ του δημόσιου, η δυνατότητα ηλεκτρονικής υποβολής φορολογικής δήλωσης, η

δυνατότητα λήψης πολλών δημοσίων εγγράφων με ηλεκτρονικό τρόπο, η επιβολή ταυτοτήτων νέου τύπου, η εισαγωγή βιομετρικών χαρακτηριστικών στα διαβατήρια, η δημιουργία “ενδιάμεσων” φορέων (π.χ. ΚΕΠ, ΕΡΜΗΣ) για την καλύτερη εξυπηρέτηση των πολιτών, η δυνατότητα συμμετοχής σε μια πρώιμη διαβούλευση σε νέα νομοσχέδια, η δυνατότητα ελέγχου των κρατικών δαπανών μέσω του συστήματος “Διαύγεια”, η υποχρεωτική υποβολή ηλεκτρονικής δήλωσης περιουσιολογίου, το κτηματολόγιο και η δυνατότητα των πολεοδομιών να κάνουν ελέγχους μέσω προγραμμάτων κατόπτρευσης (Google earth), οι ηλεκτρονικές δηλώσεις αγροτικής καλλιέργειας και η δυνατότητα του υπουργείου αγροτικής ανάπτυξης να κάνει ελέγχους μέσω προγραμμάτων κατόπτρευσης (Google earth), η υποχρεωτική ηλεκτρονική σήμανση όλων των ζώων (παραγωγικών και συντροφιάς).

Βλέποντας με μια καθαρότερη ματιά τις παραπάνω αλλαγές θα μπορούσαμε να πούμε ότι συγκεντρώνονται κυρίως γύρω από τον **οικονομικό** τομέα και ιδιαίτερα γύρω από τους **φορολογικούς-οικονομικούς ελέγχους**. Βέβαια καμιά από αυτές τις αλλαγές δεν μπορεί να θεωρηθεί ότι είναι στο κορυφαίο επίπεδο της τεχνητής νοημοσύνης (είναι απλές ντετερμινιστικές πράξεις) αλλά είναι μια τεράστια αλλαγή προς αυτή τη κατεύθυνση η οποία δημιουργεί το κατάλληλο υπόβαθρο που θα φανεί χρήσιμο στο μέλλον. Σημειώνεται ότι οι εφαρμογές της τεχνητής νοημοσύνης για να αποδώσουν χρειάζονται μεγάλο μέγεθος και καλής ποιότητας δεδομένα, κάτι που η εξέλιξη της ηλεκτρονικής διακυβέρνησης μπορεί να παρέχει.

Βλέπουμε στα παραπάνω παραδείγματα ότι οι αρχές επιβάλουν κάποιες αλλαγές οι οποίες είναι υποχρεωτικό να ακολουθηθούν από τους πολίτες. Αυτός ο **επιτακτικός χαρακτήρας** του δημοσίου τομέα είναι που δίνει ενδιαφέρον στην μελέτη των εφαρμογών της τεχνητής νοημοσύνης. Το ότι δηλαδή δεν υπάρχει η δυνατότητα άρνησης από τους πολίτες στην εφαρμογή μιας δημόσιας εντολής, ακόμη και όταν αυτή είναι αντιδημοκρατική ή αντισυνταγματική (έως ότου αυτό αποκατασταθεί). Επίσης ένα άλλο στοιχείο που πρέπει να διακριθεί είναι η **καθολικότητα** που υπάρχει στις προσταγές των αρχών και η δυνατότητα επιβολής της άμεσης εφαρμογής κάποιων μέτρων προς την υιοθέτηση τεχνολογικών εφαρμογών. Έτσι **ο πλέον γρήγορος τρόπος** για να εφαρμοστεί ένα τεχνολογικό επίτευγμα είναι να επιβληθεί μέσω του δημοσίου τομέα. Ένα ακόμη σημείο που πρέπει να συνδεθεί με τις εφαρμογές της τεχνητής νοημοσύνης είναι η δυνατότητα της εύκολης **μεταφερσιμότητας του κεντρικού ελέγχου**. Έτσι σε μια τυχόν πολιτική ενοποίηση των κρατών της Ε.Ε. η διασύνδεση και η ενοποίηση των αρχείων, των βάσεων και γενικότερα όλων των δημοσίων λειτουργιών θα γίνει εύκολα και γρήγορα, εφόσον βέβαια πριν υπάρχει η πιστή τήρηση κοινών προτύπων. Μελλοντικά, και εφόσον συνεχιστεί η σημερινή πρακτική, μια κατάσχεση ενός τραπεζικού λογαριασμού θα μπορεί εύκολα να αποφασίζεται και να υλοποιείται από έξυπνο υπολογιστικό σύστημα που θα είναι εγκατεστημένο π.χ. στις Βρυξέλες για όλους τους ευρωπαίους πολίτες ή ακόμη σε ένα πιο φανταστικό σενάριο από π.χ. την Ουάσιγκτον σε μια πιθανή ευρωαντλαντική ένωση, και πιθανόν χωρίς την παρέμβαση ανθρώπου. Επομένως η άσκηση της πολιτικής εξουσίας (η οποία

εφαρμόζεται από δημόσιους φορείς) επηρεάζεται στη σύσταση της από την τεχνολογία ή κατά άλλους η πολιτική εξουσία μεταχειρίζεται την τεχνολογία με σκοπό τον περιορισμό των ατόμων που κυβερνούν έναν γεωγραφικό τόπο.

Παράλληλα θα πρέπει να σκεφτούμε ότι οι δυνατότητες που δίνουν οι νέες τεχνολογίες για μια αποκεντρωμένη διοίκηση αλλάζουν και τον τρόπο που κάθε πρόσωπο αντιμετωπίζει τον δημόσιο τομέα. Ενώ δηλαδή αρχικά βλέπουμε μια συγκέντρωση εξουσιών στα χέρια λίγων όπως για παράδειγμα στην Ευρωπαϊκή Επιτροπή, παράλληλα βλέπουμε μια αποκέντρωση δημόσιων λειτουργιών, έτσι π.χ. δεν είναι απαραίτητο σήμερα κάποιος να ζει κοντά σε κάποιο αστικό κέντρο για να μπορεί να έχει πρόσβαση σε δημόσιους φορείς και μπορεί πλέον να απολαμβάνει τις δημόσιες υπηρεσίες από απόσταση. Αυτό μελλοντικά ίσως αλλάξει πληθυσμιακά έναν χώρο και ίσως μετακινήσει (κτηριακά) και τους δημόσιους φορείς προκαλώντας έναν τεχνολογικό αναχρονισμό (Hashem Dadashpoor, Zahed Yousefi. 2018). Αυτό σημαίνει ότι μια μελλοντική πόλη ίσως να μην αναπτύσσεται οικιστικά γύρω από τράπεζες και δημόσιους φορείς, μιας και δεν θα αποτελούν το γνωστό ως σήμερα ιστορικό οικονομικό κέντρο δραστηριοτήτων και πιθανόν δεν θα είναι απαραίτητες σαν φυσική παρουσία στην κοινωνική ζωή.

Επιπλέον οι τεράστιες δυνατότητες που υπόσχεται η τεχνητή νοημοσύνη δίνουν μια μεγάλη ελπίδα στην αντιμετώπιση δυσκολιών που σήμερα αντιμετωπίζουμε, όπως ασθένειες, επικίνδυνα καιρικά φαινόμενα, καθημερινές δυσκολίες κ.α.

Αυτά τα χαρακτηριστικά είναι ένα τεράστιο πεδίο αντιπαράθεσης απόψεων περί της ηθικής υπέρ ή κατά της υιοθέτησης των εφαρμογών της τεχνητής νοημοσύνης. Το κατά πόσο δηλαδή η συγκέντρωση του ελέγχου (άρα και της δύναμης) σε λίγους είναι μια επικίνδυνη στρατηγική, ποιόν τελικά ωφελεί περισσότερο, και εάν εν τέλει το κόστος μη υιοθέτησης είναι αποδεκτό ή όχι;

Όλοι αυτοί λοιπόν οι παράγοντες που αναφέρθηκαν κάνουν την ανάγκη της συζήτησης γύρω από θέματα ηθικής των εφαρμογών της τεχνητής νοημοσύνης στο δημόσιο τομέα επιτακτική.

5.3. Περίληψη

Ένα βασικό ερώτημα που επιχειρήθηκε να απαντηθεί σε αυτό το κεφάλαιο είναι αν είναι επίκαιρη η μελέτη γύρω από τις εφαρμογές της τεχνητής νοημοσύνης στον δημόσιο τομέα. Οι τελευταίες δομικές αλλαγές στον δημόσιο τομέα στην χώρα μας, αλλά και η ένταση της συζήτησης που γίνεται διεθνώς για την τεχνητή νοημοσύνη, δείχνουν ότι βρισκόμαστε στο επίκεντρο διεθνών εξελίξεων σε ότι αφορά την τεχνητή νοημοσύνη. Επομένως η μελέτη γύρω από τον σκοπό της χρήσης της τεχνητής νοημοσύνης από το δημόσιο δεν είναι μόνον επίκαιρη αλλά και απαραίτητη γιατί

πρώτον κρύβει πολλούς κινδύνους, για τους οποίους πρέπει να ληφθούν μέτρα πρόληψης, και δεύτερον γιατί έχει ένα μεγάλο δυναμικό που πρέπει να αναδεχθεί.

|

6. Η ηθική στις εφαρμογές τεχνητής νοημοσύνης στον δημόσιο τομέα

6.1 Εισαγωγή

Στο προηγούμενο κεφάλαιο επιχειρήθηκε να τεκμηριωθεί η αναγκαιότητα της μελέτης της ηθικής γύρω από τις εφαρμογές της τεχνητής νοημοσύνης στον δημόσιο τομέα. Σε αυτό το κεφάλαιο θα επιχειρηθεί να βρεθεί ένας κοινός τόπος των εννοιών, τόσο μελετώντας την διεθνή βιβλιογραφία όσο και αναπτύσσοντας προσωπικές απόψεις. Δηλαδή θα πρέπει να προσδιοριστούν τα θέματα ηθικής που αφορούν μόνο την εφαρμογή επιτευγμάτων τεχνητής νοημοσύνης σε μια μόνο πτυχή της ζωής, τον δημόσιο τομέα.

6.2 Η παρούσα κατάσταση στον δημόσιο τομέα

Θα αναρωτηθεί κάποιος γιατί μας ενδιαφέρει η παρούσα κατάσταση του δημοσίου τομέα; Η απάντηση είναι απλή: γιατί οι εφαρμογές της τεχνητής νοημοσύνης θα εφαρμοστούν από συγκεκριμένα άτομα και σε συγκεκριμένο κοινό και σε συγκεκριμένο κοινωνικό περιβάλλον αλλά και με στόχο να αλλάξει πολλές κακές συμπεριφορές ή κατά άλλη άποψη να τις προστατεύσει. Άρα η επιτυχία είναι άμεσα εξαρτημένη από το πεδίο εφαρμογής της. Είναι προφανές ότι όταν κάποια πρόνοια τίθεται θα υπάρχουν εμπόδια σε οτιδήποτε καινοτόμο, αλλά και όταν επιχειρείται διεύρυνση ενός κοινωνικού χάσματος πάλι θα πρέπει να υπάρξει κατά το δυνατόν αντίδραση.

Κατά το πέρασμα των χρόνων η ηθική έρχεται αντιμέτωπη με σύγχρονα ερωτήματα. Πριν οι εφαρμογές της τεχνητής νοημοσύνης εισέλθουν στον δημόσιο τομέα, τα ζητήματα ηθικής που μας απασχολούσαν στο δημόσιο τομέα, αφορούσαν κυρίως την συμπεριφορά του κυρίαρχου μέρους των εργαζομένων στον δημόσιο τομέα και των μελών της κυβέρνησης, όπως διαφθορά, αδιαφάνεια, υπεξάιρεση χρημάτων, ανευθυνότητα, συγκέντρωση εξουσίας σε ομάδες ατόμων, διάχυση και αποφυγή αρμοδιοτήτων, έλλειψη δικαιοσύνης. Όλα τα παραπάνω παραδείγματα συνοψίζονται στην διάχυτη αίσθηση που κυκλοφορεί για τους εργαζομένους και τα μέλη της κυβέρνησης ότι δεν έχουν ανεπτυγμένο το αίσθημα της κοινωνικής ευθύνης, σε βαθμό που ακόμη να κερδοσκοπούν σε μια χώρα που καταρρέει οικονομικά (Silvia Puiu. 2014). Αναλύοντας περισσότερο αυτή τη συμπεριφορά θα μπορούσε κάποιος εσφαλμένα να πει ότι το κυρίαρχο ηθικό ερώτημα που θέτει κάποιος διαφθαρμένος δημόσιος υπάλληλος ή πολιτικός στον εαυτό του (εφόσον έχει φτάσει σε αυτό το επίπεδο επίγνωσης), είναι: αν κλέβοντας χρήματα από δημόσια έσοδα, εν τέλει θα κάνει καλό αυτό στον εαυτό του; Βλέποντας μεσοπρόθεσμα, πράγματι αυτό είναι εφικτό, και το διαπιστώσαμε στην Ελληνική κοινωνία να ισχύει, π.χ. βλέπουμε πολιτικές οικογένειες με περιουσίες δυσανάλογες των εσόδων συγγενών τους από μισθούς σε κυβερνητικές θέσεις, να ευημερούν χωρίς ποτέ να έχουν εργαστεί. Το επόμενο όμως στάδιο είναι αν αυτός ο διεφθαρμένος υπάλληλος ή πολιτικός θα χάσει την “κλεμμένη ευδαιμονία” του όταν το γενικότερο κοινωνικό περιβάλλον

καταστραφεί; Εδώ οι απόψεις δίστανται, πολλοί πιστεύουν ότι η προσωπική ευτυχία επηρεάζεται από την κατάσταση που υπάρχει στο κοινωνικό περιβάλλον, άλλοι βέβαια πιστεύουν ότι κάποιος μπορεί να ζει ευχάριστα ακόμη και σε μια χώρα που δυστυχεί.

Επειδή η ανάπτυξη εφαρμογών για το δημόσιο τομέα βρίσκεται σε πρώιμο στάδιο η συζήτηση γύρω από αυτές ακόμη κινείται στο επίπεδο της αντικατάστασης κάποιων τυπικών διαδικασιών που απασχολούν τους εργαζομένους στον δημόσιο τομέα με χρήση αυτοματοποιημένων διαδικασιών, όπως την παροχή πληροφοριών, συμπλήρωση φορμών και αποστολή εγγράφων, δρομολόγηση διαδικασιών κ.α. (Hila Mehr. 2017). Αν και η αντικατάσταση αυτών των διαδικασιών είναι σημαντική και θα προσφέρει απαλλαγή από κουραστικές εργασίες, δεν αποτελεί, κατά την γνώμη του συγγραφέα, μέρος της εφαρμογής της τεχνητής νοημοσύνης γιατί είναι μια απλή ντετερμινιστική διαδικασία που για συγκεκριμένη είσοδο δεδομένων δίνει συγκεκριμένη έξοδο. Παρόλα αυτά είναι ένα καλό βήμα προς αυτή τη κατεύθυνση. Παράδειγμα τέτοιων εφαρμογών είναι η δυνατότητα έκδοσης ληξιαρχικής πράξης γέννησης μέσω μιας πύλης, χωρίς την φυσική παρουσία του πολίτη, και η αποστολή και παραλαβή της με ηλεκτρονικό τρόπο. Ακόμη όμως και σε απλή τέτοια αλλαγή στον δημόσιο τομέα οι ηθικές προεκτάσεις και η πιθανή μελλοντική χρήση τους από συστήματα τεχνητής νοημοσύνης δεν λείπουν. Από την μια τα ωφέληματα είναι πολλά όπως: εξοικονόμηση χρόνου από τους πολίτες, εξοικονόμηση εργατοωρών από τους δημοσίους υπαλλήλους, αύξηση παραγωγικότητας, μείωση κόστους, εξοικονόμηση χώρου για φύλαξη εντύπων. Από την άλλη π.χ. υπάρχει το ενδεχόμενο να έχουμε το φαινόμενο της “τεχνολογικής ανεργίας”. Θα αναλυθεί σε επόμενο κεφάλαιο εκτενέστερα ο αντίκτυπος από την χρήση της τεχνητής νοημοσύνης, αλλά εδώ ας αναφερθεί απλά ότι έχει αποδειχθεί από διεθνείς μελέτες (McKinsey Global Institute. 2018) ότι η υιοθέτηση νέων τεχνολογιών αναδιαμορφώνει τον χώρο της εργασίας μειώνοντας τον συνολικό αριθμό των ανθρώπων που εργάζονται (σύγκριση του συνολικού αριθμού αυτών που εργαζόταν πριν με αυτούς που εργάζονται μετά) βελτιώνοντας παράλληλα τις συνθήκες για αυτούς που ακόμη εργάζονται. Θα αναλυθεί στο επόμενο κεφάλαιο αν πρέπει να θεωρείται ένα ανησυχητικό φαινόμενο και αν θα πρέπει ή όχι να υιοθετηθεί μια στάση επιφυλακτική και περιοριστική απέναντι στην τεχνολογία όπως υποστηρίζεται από πολλούς (Young Joon, Kima Kyungsoo, Kimb Su, Kyoung Lee. 2017) ή ακόμη περισσότερο σε ένα ακραίο σενάριο να αναπτυχθεί ένα νέο κύμα αντίστασης εναντίων των νέων τεχνολογιών (Neo-Luddism).

Προχωρώντας ένα βήμα παρακάτω στην χρήση της τεχνητής νοημοσύνης μπορούμε να αναφέρουμε ότι στον τομέα της παροχής πληροφοριών (με την φυσική παρουσία του πολίτη στον δημόσιο φορέα) σημαντικό ρόλο έπαιξε το επίτευγμα της αναγνώρισης της ανθρώπινης ομιλίας από μηχανές, κάτι που θα μας απασχολήσει σε επόμενο κεφάλαιο όταν θα δούμε ποια είναι τα επιτεύγματα που έχει προσφέρει η έρευνα στον τομέα της τεχνητής νοημοσύνης. (LiuHuaqiang, Dai Guanzhong, Xu Nairping. 1993). Σήμερα αυτή η τεχνολογία είναι η βάση της ανάπτυξης των μηχανών

που είναι γνωστές με τον όρο chatbots ή chaterbots (Antonella De Angeli. 2009), που είναι μια σύνθεση των λέξεων chat και robots, δηλαδή ρομπότ τα οποία έχουν την δυνατότητα να συνομιλούν. Τα ρομπότ αυτά είτε έχουν μια μη ανθρώπινη φυσική μορφή είτε τους έχουν δοθεί χαρακτηριστικά ανθρώπου, όπως κεφάλι, σώμα και άκρα. Επιπλέον πολλές εταιρείες εφάρμοσαν και μη λεκτικά χαρακτηριστικά όπως εκφράσεις του προσώπου, κινήσεις του σώματος, χειρονομίες και άλλα. Τα ρομπότ αυτά δέχονται σαν είσοδο κείμενο ή εικόνα, αναγνωρίζουν και επεξεργάζονται αυτά τα δεδομένα, και στη συνέχεια δίνουν μια έξοδο που είναι μια λεκτική απάντηση ή μια άλλη πράξη. Ένα απλό παράδειγμα εφαρμογής αυτής της τεχνολογίας είναι όταν στην είσοδο του κτηρίου ενός δημόσιου φορέα ερωτηθεί ένα ρομπότ πού είναι το γραφείο του διευθυντή. Η απάντηση σε αυτό το ερώτημα είναι σαφώς προκαθορισμένη, και μπορεί εύκολα να δοθεί από το ρομπότ, εφόσον έχει ήδη λυθεί το δύσκολο μέρος της συσσώρευσης στο λογισμικό της μηχανής όλου του πλούτου μιας γλώσσας και των πιθανών ιδιωματικών και καθημερινών εκφράσεων αλλά και των συντακτικών κανόνων της. Παραδείγματα εφαρμογών της τεχνολογίας αυτής θα δοθούν στο επόμενο κεφάλαιο, ενδεικτικά αναφέρεται το ρομπότ-αστυνόμος που χρησιμοποιεί η χώρα των Ηνωμένων Αραβικών Εμιράτων για τον έλεγχο της ασφάλειας και την παροχή πληροφοριών στο αεροδρόμιο του Ντουμπαί. Προχωρώντας περισσότερο στην έρευνα γύρω από την αναγνώριση της γλώσσας η επιστήμη έχει δώσει πιο εκπληκτικά αποτελέσματα μιας και οι μηχανές δεν αντιλαμβάνονται μόνον λεκτικά νοήματα αλλά και διαφορετικές συναισθηματικές καταστάσεις (1. Raviraj Vishwambhar, Darekara Ashwinikumar, Panjabrao Dhande.2018. 2. RosarioSorbelloa, Salvatore Tramontea, Carmelo Calib, MarcelloGiardinaa, Shuichi Nishioc, HiroshiIshigurocd, AntonioChella, 2018. 3. Shaoling Jing, Xia Mao, Lijiang Chen. 2017). Επιπρόσθετα οι μηχανές αυτές μπορούν μέσω της κατόπτευσης (monitoring) να αναγνωρίζουν μοναδικά γεωμετρικά χαρακτηριστικά του ανθρώπινου προσώπου (facial recognition). Μια από τις υποενότητες της αναγνώρισης των γεωμετρικών χαρακτηριστικών είναι και το Affect recognition που χρησιμοποιεί την ανάλυση της εικόνας με σκοπό να ανιχνεύσει στοιχεία για την προσωπικότητα του ανθρώπου, τα συναισθήματά του, την ψυχική υγεία του και την αφοσίωσή του στην εργασία. Θα ερευνηθεί στη συνέχεια κατά πόσο εγείρονται ηθικά ζητήματα γύρω από αυτή την ανακάλυψη. Αν και στις περισσότερες περιπτώσεις οι έρευνες που δημοσιεύονται αφορούν την αναγνώριση συναισθημάτων με σκοπό την καλύτερη κατανόηση της γλώσσας και μόνο, δηλαδή την αύξηση της επιτυχίας στην κατανόηση και επομένως στη μετάφραση, δεν αποκλείεται ότι κάποιος δεν θα εκμεταλλευτεί αυτήν τεχνολογική δυνατότητα- για άλλους μη θεμιτούς σκοπούς π.χ. η αναγνώριση συναισθημάτων φόβου και η εκμετάλλευση αυτού. Σε αυτή την περίπτωση η ανάλυση των συναισθημάτων μέσω της φωνής θα αποτελεί μια εισβολή στον συναισθηματικό κόσμο του ατόμου επομένως συνέπεια θα είναι να θίγονται προσωπικά δεδομένα και μάλιστα ευαίσθητα ή πλέον ειδικών κατηγοριών.

Η τεχνολογία αναγνώρισης του προσώπου αλλά και του σώματος χρησιμοποιείται ήδη από ιδιωτικές εταιρείες. Παράδειγμα είναι η υπηρεσία Alipay

όπου χρησιμοποιείται από σχεδόν 120 εκατομμύρια ανθρώπους στην Ασία για πληρωμές με ταυτοποίηση μέσω ενός λογισμικού αναγνώρισης του προσώπου (της εταιρείας Face++). Πέρα από την αναγνώριση των γεωμετρικών χαρακτηριστικών το λογισμικό αυτό παρέχει και την αναγνώριση συναισθημάτων. Όπως προκύπτει από τον κώδικα που χρησιμοποιείται και παρουσιάζεται παρακάτω, τα συναισθήματα που εξετάζονται είναι επί τοις εκατό ποσοστά και είναι τα εξής: λύπη, ουδετερότητα, δυσφορία, νευρικότητα, έκπληξη, φόβος και χαρά.

Στον παρακάτω κώδικα (πηγή <https://www.faceplusplus.com/>) φαίνονται τα αποτελέσματα από μια εικόνα ενός ανθρώπου που κατά κύριο λόγο δείχνει έκπληκτος ("surprise": 99.35).

```
"attributes": {  
  "emotion": {  
    "sadness": 0.273,  
    "neutral": 0,  
    "disgust": 0.005,  
    "anger": 0.341,  
    "surprise": 99.35,  
    "fear": 0.03,  
    "happiness": 0  
  }  
}
```

Επομένως η υιοθέτηση από το δημόσιο τομέα μιας παρόμοιας λειτουργίας θα είναι μια απλή πολιτική απόφαση χρησιμοποίησης παρόμοιου λογισμικού με αυτό των ιδιωτικών εταιρειών.

Ένα άλλο γεγονός που πρέπει να επισημανθεί είναι ότι δια μέσου της μοναδικότητας της φωνής κάθε ανθρώπου, θα είναι δυνατή η δημιουργία βάσης δεδομένων “φωνητικής ταυτότητας” του κάθε πολίτη, αλλά και οι άλλες δυνατότητες που θα δίνονται πλέον εφόσον ψηφιοποιηθεί η φωνή. Το ίδιο μπορεί να συμβεί με την μοναδικότητα των γεωμετρικών χαρακτηριστικών του προσώπου και ιδιαίτερα των οφθαλμών, να οδηγηθούμε δηλαδή στη δημιουργία μιας βάσης δεδομένων που θα αποθηκεύει την οπτική ταυτότητα κάθε πολίτη. Τα δύο αυτά στοιχεία (φωνή, εικόνα) αποτελούν μέρος της **ψηφιακής ταυτότητας** ενός πολίτη (**digital identity**). Η **ψηφιακή ταυτότητα** ενός πολίτη είναι ένα σύνολο ψηφιοποιημένων δεδομένων που διακρίνουν με μοναδικό τρόπο κάθε πολίτη από κάποιον άλλο. Σε αυτή εκτός των παραπάνω θα μπορούν να ανήκουν και η κοινωνική

συμπεριφορά, οι πολιτικές ιδέες, οι ποινές, η υγεία, η ψυχική κατάσταση κ.α. Όλα αυτά τα στοιχεία είναι δυνατόν να συλλέγονται από δημόσιες κάμερες και δημόσιες βάσεις δεδομένων (π.χ. νοσοκομεία, δικαστήρια κ.α.). Ακόμη περισσότερο στην ψηφιακή ταυτότητα θα μπορεί να ανήκει και μια πιθανή πρόβλεψη μελλοντικής συμπεριφοράς. Η έννοια της ψηφιακής ταυτότητας, ενόψει της ανακάλυψης των έξυπνων συσκευών, θα μπορεί να επεκταθεί και στις συσκευές ή σε αντικείμενα ή οποιαδήποτε άλλη οντότητα. Αυτές οι οντότητες μπορεί να είναι άτομα, οργανισμοί, αντικείμενα (Paul Simmonds. 2015). Γενικεύοντας αυτή την σκέψη βλέπουμε ότι δημιουργείται ένα τεράστιο δίκτυο οντοτήτων (web of things) που αποτελούν φορείς δεδομένων και συνιστούν πρόσφορο πεδίο έρευνας και εκμετάλλευσης (big data analysis). Αν και θα αναφερθεί στο κεφάλαιο με τις εφαρμογές είναι καλό εδώ να αναφερθεί ότι ήδη στην χώρα της Κίνας λειτουργεί πειραματικά ένα τέτοιο παρόμοιο σύστημα επεξεργασίας δεδομένων σε μεγάλη κλίμακα. Το σύστημα αυτό ονομάζεται ‘‘Σύστημα Κοινωνικής Βαθμολόγησης’’. Η συμπεριφορά των πολιτών, όπως αυτή καταγράφεται από κάμερες σε δημοσίους χώρους και αξιολογείται από υπολογιστικό σύστημα, ώστε ο πολίτης να επιβραβεύεται ή τιμωρείται σε ένα σύστημα πόντων για την δημόσια συμπεριφορά του. Χαρακτηριστική είναι η φράση του κινέζου προέδρου Σι Τζινπίνγκ ότι ‘‘μια φορά αναξίπιστος, πάντα περιορισμένος’’ (δημοσίευμα στον τύπο <https://www.defence-point.gr/news/to-psifiako-apolytarchiko-pirama-tis-kinas-o-kitrinos-megalos-adelfos>), που δηλώνει τον πιθανό πληροφοριακό ολοκληρωτισμό που επιχειρείται σε αυτή τη χώρα και στις δυνατότητες που δίνει πλέον η τεχνολογία να εφαρμοστεί και σε άλλες περιοχές όπως την ΕΕ. Μάλιστα στην Κίνα ήδη ο κοινωνικός στιγματισμός έχει λάβει και ψηφιακή υπόσταση μιας και ο πολίτης που θέλει να επικοινωνήσει με κάποιον ‘‘τιμωρημένο’’ από το σύστημα βαθμολόγησης ενημερώνεται με μήνυμα στο κινητό του. Η ψηφιακή ταυτοποίηση του συνόλου των πολιτών ενός κράτους από τον δημόσιο τομέα δεν είναι πλέον μια φανταστική κατάσταση. Η μεγαλύτερη ίσως προσπάθεια ηλεκτρονικής ταυτοποίησης, γίνεται εδώ και μερικά χρόνια στο δεύτερο πολυπληθέστερο κράτος του πλανήτη, την Ινδία. Το σύστημα αυθεντικοποίησης Aadhaar θα επιχειρήσει να ταυτοποιήσει 1.6 δισεκατομμύρια πολίτες, μέσω της καταχώρισης βιομετρικών στοιχείων όπως δακτυλικά αποτυπώματα και ίριδα του ματιού. Το κράτος της Ινδίας έχει δεχθεί μεγάλη κριτική γιατί προχώρησε στην υλοποίηση αυτού του προγράμματος πριν εξασφαλίσει μια οχύρωση σε θέματα που αφορούν την προστασία των ανθρωπίνων δικαιωμάτων και πριν εξασφαλίσει την προστασία των δεδομένων. Δημοσιεύματα στον διεθνή τύπο (<https://medium.com/oxford-university/all-eyes-on-indias-biometric-id-experiment-3f01bdb17dca>) αναφέρουν ότι τα δεδομένα του συστήματος Aadhaar πωλούνται μαζικά, ενώ παράλληλα ιδιωτικές εταιρείες, όπως τράπεζες, τηλεπικοινωνιακές και άλλες, αποκλείουν όσους δεν έχουν τον δωδεκαψήφιο κωδικό στο σύστημα Aadhaar. Δεν είναι τυχαίο ότι ένα τεράστιο εγχείρημα σαν αυτό εφαρμόζεται στους φτωχότερους και πιο αγράμματους ανθρώπους στον πλανήτη, σε ανθρώπους όπου στην πλειοψηφία τους δεν έχουν ηλεκτρικό, δεν έχουν εργασία, δεν έχουν σύστημα παιδείας αλλά το κυριότερο πολλοί δεν έχουν τροφή, επιχειρώντας να γίνει η μεγαλύτερη δοκιμή σε ένα σύστημα που θα δοκιμάσει την ανθρώπινη αξιοπρέπεια αλλά και την διεθνή ανοχή. Είναι προφανές ότι ο σκοπός των εκεί

κυβερνώντων είναι ο απόλυτος έλεγχος και όχι η επιδίωξη του κοινού καλού. Προφανώς το σύστημα αυτό για να υλοποιηθεί χρειάζεται μια νομιμοφανή βάση που σε αυτή τη περίπτωση είναι η πάταξη της φοροδιαφυγής, ένα επιχείρημα που μάλλον είναι αίολο μιας και το κατά κεφαλήν εισόδημα των Ινδών είναι λιγότερο από δύο δολάρια την ημέρα (πηγή <https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?locations=IN&view=map>).

Αν αναλογιστούμε το μέγεθος του κράτους της Ινδίας και της Κίνας σε σχέση με τον παγκόσμιο πληθυσμό μπορούμε να συμπεράνουμε ότι ο τρόπος που χρησιμοποιείται η τεχνολογία της επεξεργασίας δεδομένων γέρνει προς την πλευρά της κακής χρήσης. Αυτό το συμπέρασμα θα μπορούσε να ήταν αμφίβολο αν κάπου στον πλανήτη υπήρχε έναν αντίστοιχο μεγαλεπήβολο πρόγραμμα χρήσης της τεχνητής νοημοσύνης για το καλό π.χ. της γενικής αντιμετώπισης μιας ασθένειας στο σύνολο του πληθυσμού ενός κράτους, πράγμα που δεν συμβαίνει. Τα γεγονότα αυτά δείχνουν ότι έχουμε ήδη φτάσει στην χρονική στιγμή όπου ηθικοί κανόνες πρέπει να εφαρμοστούν ώστε να αποφευχθεί μια απολυταρχική και σκοτεινή κοινωνία.

Γυρνώντας πίσω στην Ελληνική και Ευρωπαϊκή πραγματικότητα, αν και για την παροχή μιας απλής πληροφορίας (π.χ. τα απαιτούμενα δικαιολογητικά για μια έκδοση άδειας) δεν είναι απαραίτητο προς το παρόν ο πολίτης να παράσχει μια ψηφιακή ταυτότητα σε ένα σύστημα τεχνητής νοημοσύνης, στο μέλλον η ταυτότητα αυτή ίσως να είναι απαραίτητη και πιθανώς αποκλειστικά υποχρεωτική. Για να λειτουργήσει σε πλήρη ανάπτυξη (χωρίς την διαμεσολάβηση κρατικών υπαλλήλων) η τεχνητή νοημοσύνη στον δημόσιο τομέα θα πρέπει πρωτίστως να θεσμοθετηθεί σαν αυστηρά υποχρεωτική η ψηφιακή ταυτοποίηση των πολιτών, σαν ένα απαραίτητο υπόβαθρο (Clare Sullivan, Sophie Stalla-Bourdillon. 2015). Αυτή η ταυτοποίηση από μόνη της είναι ένα ζήτημα που εγείρει πολλά ηθικά διλήματα και απασχολεί πολλές κοινωνίες σε διαφορετικά κράτη. Το πρώτο βήμα για την δημιουργία της ψηφιακής ταυτότητας στην ΕΕ έχει γίνει ήδη με την υποχρέωση όλων των πολιτών που ταξιδεύουν με αεροπορικά μέσα να έχουν το νέο διαβατήριο με τα βιομετρικά χαρακτηριστικά.

Όπως αναφέρθηκε σε προηγούμενο κεφάλαιο, πάντοτε πρέπει να υπάρχει μια μεσότητα και να σταθμίζονται τα πλεονεκτήματα και τα μειονεκτήματα μιας τεχνολογικής εφαρμογής.

6.3. Οι αλλαγές στην ανθρώπινη συμπεριφορά

Αν και στην χώρα μας ακόμη δεν χρησιμοποιείται ένα ρομπότ που συνομιλεί με πολίτες, σε άλλα κράτη αυτό ήδη γίνεται όπως π.χ. στην Σαουδική Αραβία. Έτσι θα πρέπει να γίνει συνείδηση ότι εξελίσσεται μια ακόμη αλλαγή, που συντελείται αργά και αφορά την ηθική των ατόμων: είναι η αλλαγή στην ανθρώπινη συμπεριφορά

που προκαλούν και θα προκαλέσουν στο μέλλον οι εφαρμογές της τεχνητής νοημοσύνης στον δημόσιο τομέα. Παράδειγμα είναι συνομιλία ενός ατόμου με ένα ρομπότ με σκοπό να οδηγηθεί ένα αίτημα στον κατάλληλο τομέα. Στη περίπτωση αυτή ο πολίτης αρχικά δεν είναι εξοικειωμένος με το να συνομιλεί με ρομπότ. Ίσως στη συνέχεια αυτό να θεωρείται ρουτίνα, αλλά μέχρι να γίνει αυτό ίσως να υπάρξει και μια άρνηση. Ακόμη περισσότερο θα πρέπει να ερευνηθεί η βιοτική σχέση μεταξύ ανθρώπων και ρομπότ. Σε μια έρευνα, που διεξείχθη μεταξύ 2012-2017 σε 27 ευρωπαϊκές χώρες, έδειξε ότι υπήρξε αύξηση των αρνητικών συναισθημάτων των πολιτών από την συμβιωτική σχέση με ρομπότ (Timo Gnambsab, Markus Appel. 2018). Είναι λοιπόν εύλογο το ερώτημα τι θα γίνει στο μέλλον αν ένας άνθρωπος βρεθεί μεταξύ πολλών ρομπότ με τα οποία θα πρέπει να συνυπάρχει, ποια θα είναι τα συναισθήματά του; Ένα άλλο παράδειγμα που προϊόν υψηλής τεχνολογίας εισβάλλει στην ανθρώπινη υπόσταση είναι οι σαρωτές ακτινοβολίας που βρίσκονται τοποθετημένοι σε αεροδρόμια, έτσι π.χ. όταν ένας ταξιδιώτης περνά από σαρωτή σώματος στο αεροδρόμιο αρχικά αισθάνεται άβολα, επίσης το ίδιο συναίσθημα γεννιέται αν κοιτώντας μια κάμερα ο πολίτης γνωρίζει εκείνη τη στιγμή ότι γίνεται η αξιολόγηση προσωπικών του στοιχείων (από κατάλληλο λογισμικό) όπως τα συναισθήματά του.

Όπως αναφέρθηκε, σε επίπεδο έρευνας από ιδιωτικούς φορείς υπάρχουν σχετικές έρευνες για την επιρροή στην ανθρώπινη συμπεριφορά, όμως δεν υπάρχει από την πλευρά δημοσίου φορέα μια έρευνα προς αυτή τη κατεύθυνση, που ίσως θα οδηγούσε και στην δημιουργία ενός νομικού πλαισίου που θα διέπει αυτή τη συμβίωση.

Συμπερασματικά και αφού ούτε στην χώρα μας υπάρχει μια μελέτη για τον αντίκτυπο της τεχνητής νοημοσύνης στην ανθρώπινη συμπεριφορά αλλά και την αντιμετώπισή του, θα μπορούσαμε να πούμε ότι οι πολίτες θα έρθουν αντιμέτωποι με μια τεχνολογική αλλαγή για την οποία δεν έχουν ακόμη ενημερωθεί και δεν έχουν προετοιμαστεί, με αποτέλεσμα να υπάρχουν συνέπειες στην κοινωνική τους κατάσταση.

6.4. Ζητήματα εμπιστοσύνης προς την νέα τεχνολογία

Ένα ακόμη ζήτημα που γεννιέται είναι αυτό της **εμπιστοσύνης** που πρέπει να μάθουμε να δείχνουμε προς τις μηχανές τεχνητής νοημοσύνης (Dr. Guruduth Banavar. 2016) ή αν θα πρέπει οι εφαρμογές τεχνητής νοημοσύνης να διαχωριστούν σε δύο κατηγορίες σε αυτές που πρέπει να δείχνουμε και σε αυτές που δεν πρέπει να δείχνουμε εμπιστοσύνη (Jeremy Worrell. 2016). Συμπτωματικά κατά την διάρκεια της συγγραφής αυτής της διατριβής έγινε το πρώτο δυστύχημα του αυτόνομου αυτοκινήτου της Google, πράγμα το οποίο έφερε ξανά στο προσκήνιο το ηθικό ζήτημα αν θα πρέπει να δείχνουμε εμπιστοσύνη σε επιτεύγματα της τεχνητής νοημοσύνης. Να σημειωθεί εδώ ότι η πιλοτική λειτουργία αυτού του οχήματος είχε

νομίμως αδειοδοτηθεί από τις αρχές των Ηνωμένων Πολιτειών Αμερικής, όμως μετά το ατύχημα έγινε ανάκληση της άδειας αυτής.

Ένα ακόμη **λάθος** από λογισμικό τεχνητής νοημοσύνης που προκαλεί αμφισβήτηση, συνέβη στο Ισραήλ και αφορούσε εσφαλμένο συναγερμό λόγω λάθους στην μετάφραση και οδήγησε στην άδικη σύλληψη ενός παλαιστίνιου (<https://www.timesofisrael.com/israeli-police-arrest-palestinian-for-good-morning-facebook-post/>). Συγκεκριμένα, τον Οκτώβριο του 2017 ένας παλαιστίνιος ανέβασε σε κοινωνικό δίκτυο μια φωτογραφία ενός φορτωτή γράφοντας “καλημέρα”. Το λογισμικό του κοινωνικού δικτύου μετέφρασε αυτή τη λέξη από την αραβική γλώσσα στην ισραηλινή σαν “χτυπήστε τους”. Αυτό είχε ως αποτέλεσμα να συλληφθεί ο παλαιστίνιος από τις ισραηλινές αρχές.

Επιπλέον, η παροχή εμπιστοσύνης κρύβετε σε ένα ακόμη βαθύτερο επίπεδο πίσω από την τεχνολογία, σε αυτούς δηλαδή που θα διαχειρίζονται τα επιτεύγματα της τεχνητής νοημοσύνης. Με ένα παράδειγμα από την καθημερινή ζωή θα φανεί τι σημαίνουν αυτά τα επίπεδα εμπιστοσύνης: όταν πηγαίνουμε σε ένα μηχάνημα αυτόματης κατάθεσης-ανάληψης χρημάτων, που αύριο μπορεί να είναι ένα chatbot, εμπιστευόμαστε την τεχνολογία και είμαστε σίγουροι ότι η κατάθεση ή η ανάληψη μας θα καταγραφεί σωστά στον λογαριασμό μας. Όμως κάποιος καταθέτης στην Κύπρο ενώ δείχνει εμπιστοσύνη στο σύστημα κατάθεσης-ανάληψης μετρητών από μηχανήματα δεν μπορεί να δείξει εμπιστοσύνη γενικότερα στο τραπεζικό σύστημα μιας και στο παρελθόν έγινε “κούρεμα” στις καταθέσεις του. Δεν μπορεί λοιπόν να δείξει εμπιστοσύνη σε αυτούς που διοικούν και βρίσκονται πίσω από το τραπεζικό σύστημα και διαχειρίζονται την τεχνολογία. Ένα γενικό συμπέρασμα για την εδραίωση της εμπιστοσύνης είναι ότι βασίζεται κυρίως στην συσσωρευμένη εμπειρία και στα διδάγματα κοινής πείρας (Dr. Guruduth Banavar. 2016), που έχει αποκτήσει κάθε άτομο στο κοινωνικό σύνολο που ζει. Επομένως αν σε μια χώρα βασιλεύει η διαφθορά στα ανώτερα κοινωνικά-πολιτικά στρώματα κάθε προσπάθεια εφαρμογής τεχνολογικών επιτευγμάτων της τεχνητής νοημοσύνης θα έρθει αντιμέτωπη με την αμφισβήτηση, άρα και την έλλειψη εμπιστοσύνης. Η εμπιστοσύνη βέβαια δεν έχει μόνο τρωτό σημείο τους ανθρώπους αλλά και τα ίδια τα υπολογιστικά συστήματα. Μια ακόμη περίπτωση που κλονίζει την εμπιστοσύνη προς συστήματα που συγκεντρώνουν πληροφορίες πολλών πολιτών (που και αυτή συμπτωματικά έγινε γνωστή κατά τη διάρκεια της συγγραφής της διατριβής) είναι η αποκάλυψη για την διαρροή δεδομένων εκατομμυρίων χρηστών του κολοσσού Facebook, για λογαριασμό της εταιρείας Cambridge Analytica, και για λόγους δημιουργίας εκλογικού προφίλ των χρηστών και εκμετάλλευσης αυτών από το πολιτικό σύστημα. Όταν η ασφάλεια των προσωπικών δεδομένων πλήττεται σε μια από τις αρτιότερες εταιρείες στον πλανήτη, πώς οι πολίτες μπορούν να εμπιστευτούν συστήματα που υποστηρίζονται από μικρότερες δομές με μικρότερη δυνατότητα οικονομικής στήριξης;

Θα πρέπει σε αυτό το σημείο να τονιστεί ότι σε πολλές από τις περιπτώσεις πλήξης της εμπιστοσύνης η πηγή της αναξιοπιστίας δεν είναι αυτό καθαυτό το

τεχνικό-υπολογιστικό σύστημα αλλά οι διαχειριστές πίσω από αυτό. Η διαπίστωση αυτή έρχεται να ενισχύσει την ιδέα ότι όταν ένα διαφθαρμένο δημόσιο πολιτικό σύστημα επιχειρεί να εφαρμόσει καινοτόμες τεχνολογικές εφαρμογές, τότε οι πολίτες βλέπουν το εγχείρημα με καχυποψία και δεν μπορούν να το εμπιστευτούν. Επομένως οι προσπάθειες βελτίωσης της εμπιστοσύνης σε μια εφαρμογή τεχνητής νοημοσύνης θα πρέπει να κατευθύνονται όχι μόνο στο τεχνικό μέρος της εφαρμογής (ασφάλεια) αλλά και γύρω και πίσω από αυτό όπου υπάρχουν οι διαχειριστές και υπεύθυνοι λειτουργίας της. Η έννοια λοιπόν της υπευθυνότητας των διαχειριστών μιας εφαρμογής είναι πρωταρχική, πράγμα το οποίο επιβεβαιώνεται και από την κυρίαρχη εταιρεία στον χώρο της πληροφορικής IBM. Μπορεί κάποιος να την βρει στην πρώτη θέση των ηθικών προτεραιοτήτων όπως αυτές δημοσιεύτηκαν στον πρακτικό οδηγό της με τον τίτλο: ‘Everyday ethics for Artificial Inteligence’ (<https://www.ibm.com/watson/assets/duo/pdf/everydayethics.pdf>).

Η έννοια της εμπιστοσύνης συνδέεται έμμεσα με την προηγούμενη εμπειρία που έχει αποκτήσει ο πολίτης. Ο κανόνας ότι αν κάτι έχει συμβεί στο παρελθόν μπορεί να ξανασυμβεί στο μέλλον είναι κυρίαρχος στην συνείδηση του ανθρώπου (Albert Solé Ollé, Pilar Sorribas-Navarro. 2018). Προβάλλοντας τα παραπάνω στην Ελληνική πραγματικότητα βλέπουμε ότι οι πολίτες μετά από συνεχείς αθετήσεις υποσχέσεων από τους κυβερνώντες αλλά και μετά την προδοτική πολλές φορές στάση των κυβερνώντων (π.χ. παράβλεψη της κυρίαρχης γνώμης του δημοψηφίσματος του Ιουλίου 2015, παράδοση του ιστορικού ονόματος της Μακεδονίας, κούρεμα καταθέσεων των Ελληνοκυπρίων συμπολιτών μας κ.λπ.) έχουν χάσει σημαντικό μέρος της εμπιστοσύνης τους προς τις κυβερνήσεις. Το ότι έχει χαθεί η εμπιστοσύνη π.χ. στο τραπεζικό σύστημα, φαίνεται στο ότι μετά από αλλεπάλληλες προσκλήσεις της κυβέρνησης να επιστρέψουν οι πολίτες τις καταθέσεις τους αυτοί φοβούνται και δεν το κάνουν (<https://www.bankofgreece.gr/BogEkdoseis/%CE%95%CF%80%CE%B9%CF%83%CE%BA%CF%8C%CF%80%CE%B7%CF%83%CE%B7%20%CF%84%CE%BF%CF%85%20%CE%95%CE%BB%CE%BB%CE%B7%CE%BD%CE%B9%CE%BA%CE%BF%CF%8D%20%CE%A7%CF%81%CE%B7%CE%BC%CE%B1%CF%84%CE%BF%CF%80%CE%B9%CF%83%CF%84%CF%89%CF%84%CE%B9%CE%BA%CE%BF%CF%8D%20%CE%A3%CF%85%CF%83%CF%84%CE%AE%CE%BC%CE%B1%CF%84%CE%BF%CF%82.pdf>). Έτσι λοιπόν κάποιες προσπάθειες εφαρμογής ενός επιτεύγματος της τεχνητής νοημοσύνης από το δημόσιο, σε χώρες όπου το πολιτικό σύστημα αμφισβητείται, θα αντιμετωπίζεται με καχυποψία, γιατί παρόλο που πιθανόν να είναι μια διευκόλυνση στην επικοινωνία με τις διαδικασίες του δημοσίου και όχι με την κυβέρνηση όπως επιχειρείται να παρουσιαστεί (Aggeliki Androutopoulou, Nikos Karacapilidis, Euripidis Loukis, Yannis Charalabidisa. 2018), θα κυριαρχήσει η άποψη ότι αποσκοπεί στην απομάκρυνση των κυβερνώντων από την τριβή της καθημερινής επαφής πρόσωπο με πρόσωπο και την εισαγωγή μεταξύ αυτών ενός “τεχνολογικού φράγματος”. Προς αυτή την κατεύθυνση είναι και η αλλαγή στην ρύμη των λόγων των τελευταίων κυβερνητών για λειτουργία μιας διακυβέρνησης (δια-κυβέρνησης) και όχι για μιας κυβέρνησης, υπονοώντας ίσως την

εισαγωγή ενός συστήματος, διαμέσου του οποίου θα κυβερνάται ο πολίτης από απόσταση. Στην ιστοσελίδα του υπουργείου διοικητικής ανασυγκρότησης η ηλεκτρονική διακυβέρνηση αναφέρεται σαν eGovernment (“Η Ηλεκτρονική Διακυβέρνηση (eGovernment) αφορά στον εκσυγχρονισμό του Κράτους...” http://www.minadmin.gov.gr/?page_id=12126) και όχι σαν governance. Υπάρχει πληθώρα ηχητικών αποδείξεων όπου σημαίνοντες πολιτικοί της χώρας μας μιλούν όχι μόνο για μια διακυβέρνηση αλλά επιπλέον για μια παγκόσμια διακυβέρνηση (π.χ. οι δηλώσεις της πρώην προέδρου της Βουλής ή του πρώην πρωθυπουργού, <https://www.youtube.com/watch?v=zJ9U6YaCoxU>). Μάλιστα ο όρος που χρησιμοποιείται τελευταία έχει απωθήσει το συνθετικό e- και εμφανίζεται πλέον σαν Government π.χ. όπως αναφέρει το αρμόδιο υπουργείο ‘βαθμός ωρίμανσης Government 3.0’” (http://www.minadmin.gov.gr/wp-content/uploads/20140415_egov_strategy.pdf και <https://www.gov30.eu/>). Στην αμφιβολία ότι πρόκειται για λάθος φρασεολογία ή ότι πρόκειται για μια φανταστική κατάσταση, έρχεται να απαντήσει η ύπαρξη του Κινέζικου συστήματος κοινωνικής βαθμολόγησης το οποίο λειτουργεί με έναν ανέπαφο τρόπο για την κυβέρνηση. Ουσιαστικά στην Κίνα κυβερνά το σύστημα για λογαριασμό της κυβέρνησης, και οι πολίτες κυβερνούνται από αυτό.

Σε μια έρευνα σχετικά με την άποψη των πολιτών για τις υπηρεσίες ηλεκτρονικής διακυβέρνησης που παρέχονται μέσω των ΚΕΠ ενισχύεται η άποψη ότι η εμπιστοσύνη απέναντι σε υπηρεσίες που παρέχονται διαδικτυακά είναι ένας παράγοντας ανασταλτικός για την χρήση τους (Anastasia Voutinioti. 2013) Προφανώς και σε αυτή την περίπτωση τα βαθύτερα αίτια της αμφιβολίας βρίσκονται στα πρόσωπα πίσω από την τεχνολογία (στην συγκεκριμένη περίπτωση, στην κυβέρνηση) και λιγότερο στην ίδια την τεχνολογία και στους κατώτερους υπαλλήλους που την παρέχουν, γιατί πρώτον η τεχνολογία δεν είναι ο κύριος παράγοντας δυσπιστίας και δεύτερον οι κατώτεροι υπάλληλοι συνήθως δεν έχουν την δυνατότητα παρέμβασης στα τεχνολογικά συστήματα

Όμως δεν είναι μόνο το κακό παρελθόν ενός πολιτικού συστήματος που εισπράττει την αμφισβήτηση από την πλευρά των πολιτών, αλλά υπάρχει και ένας δεύτερος φόβος που πρέπει να διασφαλιστεί, ότι δεν θα βλάψει τους πολίτες η πιθανή **μελλοντική μεταστροφή** των πολιτικών απόψεων μιας κυβέρνησης. Η άποψη ότι η τεχνητή νοημοσύνη κάτω από συγκεκριμένες συνθήκες μπορεί να γίνει επικίνδυνη, ενισχύεται, όταν αυτοί που την υποστηρίζουν είναι κορυφαίοι επιστήμονες στον χώρο της παραγωγής του σχετικού λογισμικού. Μια αθώα πτήση ενός τηλεκατευθυνόμενου ή στο μέλλον ενός αυτόνομου πτητικού μέσου μπορεί υπό συγκεκριμένες συνθήκες να εγείρει ζητήματα που σχετίζονται με προσβολή των ανθρωπίνων δικαιωμάτων. Οι εργαζόμενοι στον κολοσσό Google αντέδρασαν στο κοινό πρόγραμμα της εταιρείας τους με το Πεντάγωνο. Η συνεργασία στο πρόγραμμα Maven το οποίο χρησιμοποιεί την τεχνητή νοημοσύνη για ανάλυση εικόνων που λαμβάνονται από μη επανδρωμένα πτητικά μέσα (drone imaging) και την εξαγωγή από αυτές στόχων που έχουν στρατιωτικό ενδιαφέρον, διεκόπη ύστερα από την έντονη διαμαρτυρία των

εργαζομένων της εταιρείας. Οι εργαζόμενοι με την ανακοίνωσή τους (<https://static01.nyt.com/files/2018/technology/googleletter.pdf>) εξέφρασαν την άποψη ότι το συγκεκριμένο πρόγραμμα δεν συνάδει με την ιστορία της εταιρείας και δεν εξυπηρετεί την αρχή τους να “μην είναι σατανικοί” (Don’t Be Evil) και παραδίδει στα χέρια τρίτων την διαφύλαξη των ηθικών υποχρεώσεών τους. Έχουν γίνει αναφορές στον διεθνή τύπο ότι εκτός της εταιρείας Google στο πρόγραμμα συμμετέχουν και άλλες εταιρείες οι οποίες κρατούν μυστική στάση, όπως οι: DigitalGlobe, IBM, Nvidia. Γυρνώντας πίσω στην έννοια της **ανοχής** που ορίστηκε στο πρώτο κεφάλαιο, βλέπουμε ότι οι εργαζόμενοι στη Google προτιμούν να αποδεχθούν ένα ρίσκο στο πεδίο της άμυνας προκειμένου να προστατέψουν τα δικαιώματα των ανθρώπων από κάποιους που θα χρησιμοποιούσαν το λογισμικό τους για διαβολικό σκοπό. Είναι ίσως η πρώτη φορά σε τέτοιο μέγεθος όπου αναγνωρίζεται η επικινδυνότητα ενός επιτεύγματος όταν βρεθεί στα χέρια ανήθικων ανθρώπων. Η απόφαση του να μην γίνει κάτι καθόλου, γιατί αν γίνει μπορεί να χρησιμοποιηθεί για κακό, είναι μεν μια τεχνολογική υποχώρηση αλλά εξυπηρετεί την ηθική αρχή να διαφυλάσσουμε μόνο ότι είναι καλό για την ζωή μας. Στο προηγούμενο παράδειγμα των εργαζομένων της Google βλέπουμε ότι προκειμένου να μην γίνει μια αλλαγή χρήσης της τεχνολογίας στο μέλλον είναι προτιμότερο να μην γίνει τίποτα, λαμβάνονται δηλαδή σήμερα μέτρα που θα αποκλείουν την πιθανή μελλοντική αλλαγή των απόψεων αυτών που θα έχουν την τεχνολογία στα χέρια τους.

Γυρνώντας στην Ελληνική πραγματικότητα, κατά την γνώμη του συγγραφέα, η στιγμή που εφαρμόζονται στη Ελλάδα όλες οι τελευταίες μεταρρυθμίσεις στην ηλεκτρονική διακυβέρνηση και η σκέψη αλλά και η προετοιμασία (αν θεωρηθεί ότι η καθιέρωση κοινών προτύπων, η δημιουργία ανοιχτών δεδομένων και η αλλαγή στην προστασία των προσωπικών δεδομένων εξυπηρετούν την τεχνητή νοημοσύνη) για μια πιθανή μελλοντική χρήση της τεχνητής νοημοσύνης δεν αποσκοπούν αποκλειστικά στην βελτίωση της ζωής, αλλά καλυπτόμενοι πίσω από μια μικρή πραγματική βελτίωση κρύβουν μια μεγάλη απώλεια στον τομέα των δικαιωμάτων. Επιπλέον φαίνεται να είναι χρονικά σκοπούμενες, γιατί πρέπει κατά την βούληση των κυβερνώντων να γίνουν σε ένα περιβάλλον όπου το σύνταγμα πλήττεται και η νομιμότητα παραβιάζεται στο όνομα της οικονομικής “διάσωσης”. Η επικαλούμενη, ακόμη και σε διεθνές επίπεδο, αίσθηση του κατεπείγοντος της υιοθέτησης των τεχνολογικών εφαρμογών είναι πλαστή και εξυπηρετεί άλλα συμφέροντα, και συνιστά μια επίθεση κατά την πιο αδύναμη στιγμή της κοινωνικής ζωής, ενώ θα έπρεπε το κατεπείγων να αφορά στην αντιμετώπιση της απειλής της ηθικής υπόστασης του ατόμου (ORBIT. 2019).

Επειδή ακόμη σαν πολίτες δεν έχουμε όλες τις απαντήσεις στην θεμιτή ή αθέμιτη χρήση της τεχνητής νοημοσύνης και δεν πρέπει να αφήσουμε αναπάντητες τις ανησυχίες των πολλών για το κέρδος των λίγων, θα πρέπει προς το παρόν να κρατήσουμε μια επιφυλακτική στάση απέναντι στις αλλαγές που επιχειρούνται π.χ. η ψήφιση του νέου νόμου (ν. 3979/2011) της ηλεκτρονικής διακυβέρνησης που κάνει λόγω για ψηφιακή υπογραφή, διαπιστευτήρια, αυθεντικοποίηση κ.α.

6.5. Περίληψη

Η παροχή υπηρεσιών προς τους πολίτες είναι η βασικότερη υποχρέωση του δημοσίου. Η τεχνητή νοημοσύνη μετά τις τελευταίες ανακαλύψεις της αναγνώρισης της φυσικής γλώσσας μπορεί να βοηθήσει την παροχή αυτή, χρησιμοποιώντας μηχανές που θα απαντούν σε ερωτήματα που θα θέτει ένας πολίτης και σε αιτήματα παροχής υπηρεσιών. Σε αυτό το σημείο εισέρχονται τα προσωπικά δεδομένα, τα οποία σε πολλές περιπτώσεις πρέπει να παρέχονται ώστε να δοθεί μια προσωποποιημένη απάντηση-υπηρεσία σε κάθε πολίτη. Αυτή ακριβώς η παροχή υπηρεσιών κρύβει τον κίνδυνο, η ταυτότητα ενός πολίτη όπως αυτή αναγνωρίζεται από την μηχανή να χρησιμοποιηθεί για κακό σκοπό. Επιπλέον η εισαγωγή μηχανών που θα παρέχουν υπηρεσίες προς τους πολίτες κρύβει κινδύνους όπως η αύξηση της ανεργίας, η αλλαγή στην ανθρώπινη συμπεριφορά, και η πλήξη της εμπιστοσύνης των πολιτών απέναντι στο δημόσιο. Βεβαίως υπάρχουν εφαρμογές της τεχνητής νοημοσύνης που δεν χρησιμοποιούν προσωπικά δεδομένα που μπορούν όμως να φανούν χρήσιμες στον πολίτη όπως π.χ. η πρόβλεψη των καιρικών συνθηκών.

7. Η ανάγκη για τήρηση ενός κώδικα ηθικής

7.1. Εισαγωγή

Βλέπουμε λοιπόν ότι η ηθική των εφαρμογών της τεχνητής νοημοσύνης δεν αφορά αποκλειστικά τις ίδιες τις εφαρμογές αλλά και τους ανθρώπους που εμπλέκονται με αυτή είτε σαν κατασκευαστές και διαχειριστές είτε σαν χρήστες. Για αυτό λοιπόν το λόγο θα πρέπει να γραφεί και να τηρηθεί ένας κώδικας ηθικής στον οποίο θα υπακούει κατά πρώτον ή ίδια η τεχνολογία, αλλά και κατά δεύτερον οι άνθρωποι. Παρακάτω θα αναλυθεί πως μπορεί αυτό να γίνει είτε με τεχνολογικές λύσεις είτε με νομικές.

7.2. Πως υλοποιείται η απαίτηση να υπακούει η τεχνολογία σε ένα σύνολο ηθικών αρχών;

Ένα βασικό ερώτημα που αφορά την υλοποίηση μιας εφαρμογής είναι το πώς θα εισάγουμε ηθικές αρχές και αξίες σε μηχανές; Υπάρχουν τρεις λύσεις σε αυτό το ερώτημα: Πρώτον εισάγοντας σε αυτές συγκεκριμένο κώδικα (Top-Down Decision Making), δεύτερον επιτρέποντάς τις να “μαθαίνουν” παρατηρώντας τις ανθρώπινες πράξεις (Bottom-Up Decision Making) και τρίτον να συνδυάζονται οι δύο πρώτες λύσεις (Alan F. T. Winfield, Marina Jirotko. 2018).

Στη πρώτη περίπτωση (Top-Down Decision Making) η μηχανή τεχνητής νοημοσύνης θα φέρει και θα συμπεριφέρεται ακολουθώντας την ηθική των δημιουργών της, όπως αυτή καταγράφηκε στο λογισμικό της. Στη δεύτερη περίπτωση (Bottom-Up Decision Making) η μηχανή μαθαίνει μόνη της παρατηρώντας πολλές διαφορετικές περιπτώσεις ανθρώπινων συμπεριφορών π.χ. τα βιογραφικά ανθρώπων που θεωρούνταν ότι είχαν αξιόλογη ηθική (Ernest Davis. 2015) και μετά από κατάλληλη ανάλυση κρατά και συμπεριφέρεται με τις αξίες που έχει διωλίσει. Να σημειωθεί εδώ ότι προς την κατεύθυνση αυτή της αυτόνομης μάθησης των μηχανών υπάρχουν ήδη επιτεύγματα και ενδεικτικά αναφέρεται το πρόγραμμα GoogleBrain της εταιρείας Google. Μέρος αυτού του προγράμματος ήταν και η δημιουργία μηχανής που αναγνωρίζει οντότητες (π.χ. μια γάτα) χωρίς να έχει διδαχτεί από τον άνθρωπο αλλά “παρακολουθώντας” κινηματογραφικές ταινίες. Υπάρχουν πλεονεκτήματα και μειονεκτήματα σε αυτές τις περιπτώσεις. Όταν μια μηχανή ακολουθεί ντετερμινιστικά ένα σύνολο ηθικών αρχών που τέθηκαν από τον δημιουργό της τότε το πλεονέκτημα είναι ότι δεν θα υπάρχουν αποκλίσεις από αυτές τις αρχές αλλά και το μειονέκτημα είναι ότι αν αυτές οι αρχές είναι λάθος θα κάνει λάθη και η μηχανή τεχνητής νοημοσύνης. Στην δεύτερη περίπτωση δεν υπάρχει η άμεση επιρροή, υπάρχει πάλι όμως η πιθανότητα μια μηχανή να θεωρήσει μια κακή και συνήθη ανθρώπινη συμπεριφορά σαν σωστή και να την επαναλάβει. Η πιο διάσημη τέτοια περίπτωση είναι το chatbot Tray το οποίο μαθαίνοντας να συμπεριφέρεται με τροφοδοσία τα σχόλια στα κοινωνικά δίκτυα, γρήγορα απέκτησε ρατσιστική διάθεση απέναντι στο γυναικείο φύλλο, με αποτέλεσμα να επιτεθεί

λεκτικά με μισογυνισμό δύο φορές σε γυναίκες. Παρόμοιο παράδειγμα είναι το αλγοριθμικό λάθος του λογισμικού Haar Cascades το οποίο δεν κατάφερε να αναγνωρίσει ανθρώπους με σκούρο χρώμα λόγω ότι “διδάχτηκε” από γεωμετρικά χαρακτηριστικά λευκών ανθρώπων.

Ένας άλλος κίνδυνος που ελλοχεύει στα συστήματα που μαθαίνουν μόνα τους είναι η κακόβουλη τοποθέτηση ιών που απευθύνονται σε νευρωνικά δίκτυα, κάτι που είναι ήδη εφικτό (Minhui Zou, Yang Shi, Chengliang Wang, Fangyu Li, WenZhan Song, Yu Wang. 2018).

Βέβαια εφόσον η τεχνητή νοημοσύνη καταφέρει και απομακρύνει όλες τις κακές ανθρώπινες συμπεριφορές ίσως να ξεπεράσει ηθικά τον άνθρωπο. Κατά την άποψη του συγγραφέα αυτό είναι πιθανό να συμβεί, με μόνο εμπόδιο που θα βρεθεί μπροστά στις μηχανές να είναι οι ανήθικοι άνθρωποι των οποίων τα συμφέροντα θα θίγονται.

Επειδή η ανάπτυξη σε ένα σύστημα που θα αναπτυχθεί αυτόνομα είναι τυχαία, υπάρχει και η πιθανότητα οι μηχανές να αναπτύξουν μόνες τους μια χειρότερη ηθική από αυτή των ανθρώπων πράγμα που θα αποβεί ανεξέλεγκτο και πιθανώς καταστροφικό. Σε μεγάλο βαθμό ίσως αυτός ήταν και ο φόβος (μια ανεξέλεγκτη κατάσταση) των ανθρώπων γνωστής εταιρείας κοινωνικής δικτύωσης (πρόγραμμα FAIR) όταν αναγκάστηκαν, σύμφωνα με δημοσιεύματα στον τύπο, να διακόψουν μια συζήτηση μεταξύ δύο chatbots, την στιγμή που άρχισαν αυτά να συζητούν σε μια γλώσσα ακατανόητη για τον άνθρωπο, ενώ η εταιρεία ισχυρίζεται ότι αυτό ήταν προγραμματιστικά αποκλεισμένο (<https://code.fb.com/ml-applications/deal-or-no-deal-training-ai-bots-to-negotiate/>) ενώ από κάποιους άλλους υποστηρίζεται ότι η συζήτηση διεκόπη γιατί πλέον δεν έδινε χρήσιμα αποτελέσματα. Στην ίδια κατεύθυνση του φόβου κινούνται και άλλοι που δραστηριοποιούνται στο χώρο της τεχνητής νοημοσύνης (π.χ. Elon Mask) αναπτύσσοντας θεωρίας περί του φόβου δημιουργίας “στρατού ανθρωποειδών τεχνητής νοημοσύνης”.

Μια ακόμη διαπίστωση η οποία οδηγεί στην ανάγκη λήψης αποφάσεων που αφορούν την ηθική είναι ότι η νομική σκέψη έπεται των τεχνολογικών εξελίξεων, αυτό σημαίνει ότι οι νομικοί έρχονται αντιμέτωποι με μια πρωτόγνωρη κατάσταση χωρίς προηγούμενο, όταν για παράδειγμα πρέπει να αποφανθούν για το ατύχημα που προκάλεσε το αυτόνομο όχημα της εταιρείας Google. Σε αυτό το σημείο βέβαια εισάγεται η έννοια της σχέσης της ηθικής των μηχανών με την δυνατότητα καταλογισμού ευθυνών σε αυτές, αντικείμενο που στην παρούσα κατάσταση δεν έχει καλυφθεί νομικά. Γενικότερα θα μπορούσε να ισχύει ότι την ευθύνη των πράξεων μιας μηχανής τεχνητής νοημοσύνης φέρει ο δημιουργός του, σύμφωνα με το 12ο άρθρο της συνθήκης των Ηνωμένων Εθνών για την χρήση ηλεκτρονικών επικοινωνιών (Paulius Čerkaa, Jurgita Grigienė, Gintarė Sirbikytė. 2015).

7.3. Η εισαγωγή ενός ηθικού φίλτρου στον άνθρωπο

Στο προηγούμενο κεφάλαιο αναφέρθηκε η ανάγκη να υπάρχει ένα ηθικό φίλτρο όχι μόνο στις μηχανές αλλά και στον άνθρωπο. Όσον αφορά τις μηχανές αναφέρθηκαν στις προηγούμενες παραγράφους πολλοί τρόποι να επιτευχθεί αυτό. Όσον αφορά τον ίδιο τον άνθρωπο και ιδιαίτερα τους κατασκευαστές και τους διαχειριστές συστημάτων τεχνητής νοημοσύνης θα πρέπει να αναλυθεί εν συντομία το πώς αυτοί οι άνθρωποι θα αποκτήσουν ένα ηθικό φίλτρο που θα κυβερνά τις πράξεις τους και επομένως και τα έργα τους. Πρέπει να αναφερθεί το παράδοξο αλλά επιστημονικά αποδεκτό, με ελάχιστες διαφωνίες, γεγονός ότι η ηθική είναι έμφυτη στον άνθρωπο (Mary Midgley, 2002) αλλά έρχεται σε σύγκρουση με άλλες έμφυτες αδυναμίες. Ταυτόχρονα στον άνθρωπο υπάρχει και η δύναμη να λύνει αυτές τις συγκρούσεις, έτσι η ηθική μπορεί να καλλιεργείται. Μια καλή πρακτική για τους κατασκευαστές και διαχειριστές συστημάτων τεχνητής νοημοσύνης θα ήταν η διορατική αντιμετώπιση, με άλλα λόγια να σκεφτούν την έννοια της ατομικότητας, να σκεφτούν ότι δεν θα είναι αυτοί οι οποίοι θα συμβάλουν σε μια ανήθικη πράξη.

Βέβαια ο πιο αποδοτικός αλλά αμφιλεγόμενος τρόπος είναι η επιβολή άνωθεν ενός αυστηρού ηθικού πλαισίου προς τους διαχειριστές και κατασκευαστές μέσω της νομοθεσίας. Για την σχέση της ηθικής με το δίκαιο και την νομιμότητα θα αφιερωθεί το επόμενο ειδικό υποκεφάλαιο.

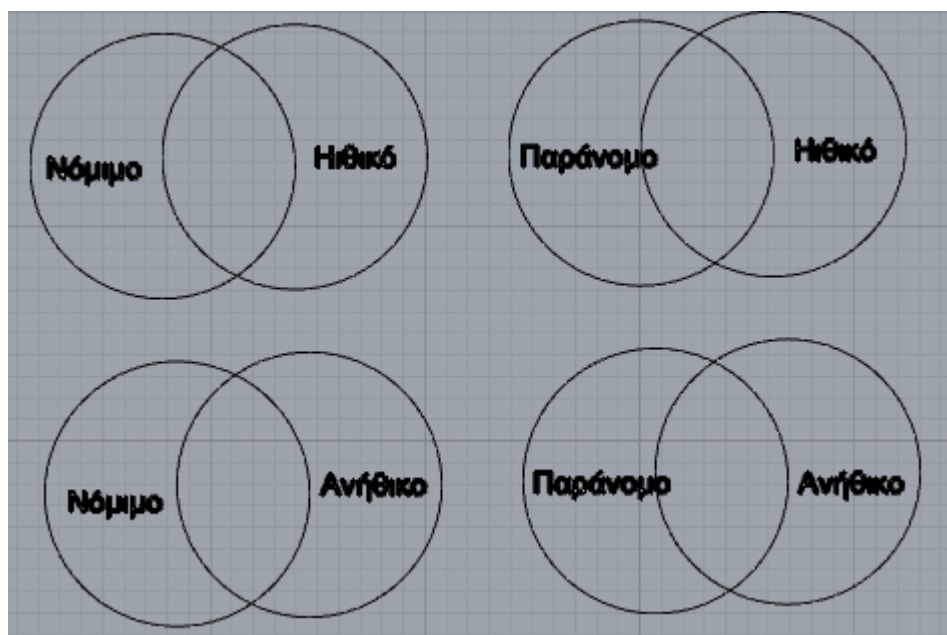
7.3.1 Το νομικό πλαίσιο και η σχέση του με την ηθική

Έχει αναφερθεί στο προηγούμενο υποκεφάλαιο ότι όσον αφορά τις μηχανές η είσοδος μιας διεπαφής ηθικής είναι εφικτή με διάφορους τρόπους. Το μέρος που αφορά αποκλειστικά την εισαγωγή ηθικής στον άνθρωπο αναλύεται σε αυτό το υποκεφάλαιο.

Γενικότερα υπάρχουν δύο τρόποι ώστε ένας πολίτης να συμπεριφέρεται ηθικά: πρώτον να υπακούει σε κάποιο νόμο που καλύπτει την ηθική υποχρέωση και δεύτερον η συμπεριφορά του να καθοδηγείται από μια εσωτερική δύναμη ανεξάρτητη από τους θεσμικούς κανόνες. Μετά από την μελέτη των διαλέξεων του συνεδρίου “AI, ethics and the law” του ινστιτούτου The Alan Turing Institute (AI, ethics and the law. What challenges and what opportunities? 2018) που διεξείχθη τον Ιανουάριο του 2018, στο οποίο εκφράστηκε η απορία αν η τεχνητή νοημοσύνη πρέπει να ρυθμίζεται από την ηθική ή τους νόμους, γεννήθηκε στον συγγραφέα της εργασίας η παρακάτω ιδέα.

Για να κατανοηθεί η λειτουργία των παραπάνω εισάγεται η έννοια **των κύκλων της ηθικής**. Οι ανθρώπινες πράξεις μπορούν γενικότερα να καταταγούν σε δύο σύνολα: της ηθικής και της νομιμότητας. Τα σύνολα αυτά έχουν μεταξύ τους την σχέση που παρουσιάζεται στο παρακάτω σχήμα (Σχήμα 1. Οι κύκλοι της ηθικής. Η σχέση των συνόλων ηθικής και νομιμότητας).

Σχήμα 1. Οι κύκλοι της ηθικής. Η σχέση των συνόλων ηθικής και νομιμότητας.



Εξηγώντας το παραπάνω σχήμα μπορούμε να πούμε ότι μεταξύ δύο συνόλων υπάρχει η περιοχή που αυτά τέμνονται και άρα επαληθεύονται και οι δύο συνθήκες (π.χ. και νόμιμο και ηθικό) και οι περιοχές που επαληθεύεται μια μόνο συνθήκη (π.χ. μόνο νόμιμο ή μόνο ηθικό). Αν και η αντίληψη περί της ηθικής έχει υποκειμενικά στοιχεία που μπορούν να εγείρουν αμφισβήτηση, θα δοθούν στη συνέχεια ενδεικτικά παραδείγματα (1 έως 8) από την κοινωνική ζωή μας για το κάθε επιμέρους σχήμα:

1. πράξη που είναι μόνο ηθική και δεν σχετίζεται με την νομιμότητα:

Το γεγονός ότι κάποιος δωρίζει ένα όργανο του σώματός του σε συγγενικό του πρόσωπο.

2. πράξη που είναι μόνο νόμιμη και δεν σχετίζεται με την ηθική:

Η αίτηση για χορήγηση άδειας κυκλοφορίας οχήματος.

3. πράξη που είναι και ηθική και νόμιμη:

Η καταβολή του δίκαιου και αναλογικού φόρου εισοδήματος

4. πράξη που είναι μόνο ανήθικη και δεν σχετίζεται με την νομιμότητα:

Το γεγονός να μη καλέσει κάποιος ασθενοφόρο όταν βλέπει ότι η ζωή ενός συναθρώπου του βρίσκεται σε μεγάλο κίνδυνο.

5. πράξη που είναι μόνο παράνομη και δεν σχετίζεται με την ηθική:

Αν και η παρανομία είναι από μόνη της μια πράξη που δείχνει αδιαφορία προς μια οργανωμένη πολιτεία και επομένως μια ανήθικη πράξη θα μπορούσαμε να πούμε ότι π.χ. κάποιος που δεν σταματά σε ένα σηματοδότη που εμφανώς

βρίσκεται σε μια έρημη περιοχή και δεν υπάρχει κανένας κίνδυνος για άλλους οδηγούς ή πεζούς διαπράττει μια παρανομία αλλά δεν προσβάλλει την ηθική.

6. πράξη που είναι και ηθική και παράνομη:

Το αμφιλεγόμενο “ξέπλυμα της τιμής”, ή “βεντέτα” με αυτοδικία, μιας οικογένειας που δέχθηκε άδικη προσβολή από άλλους, π.χ ο πατέρας που σκοτώνει τον βιαστή του παιδιού του, ή το γεγονός ότι ένας άπορος γονιός κλέβει τρόφιμα για να δώσει στο παιδί του.

7. πράξη που είναι ανήθικη και νόμιμη:

Το γεγονός ότι οι βουλευτές τυχαίνουν ιδιαίτερης δικαστικής μεταχείρισης κάτι που προσβάλλει την ηθική του συνόλου. Ή το γεγονός ότι ένας βουλευτής του κοινοβουλίου πληρώνει σε απόλυτο ποσό μικρότερο φόρο από κάποιον συνταξιούχο που εισπράττει σύνταξη το ένα όγδοο του μισθού του βουλευτή.

Το γεγονός ότι σε συνθήκες πολέμου επιτρέπεται να σκοτώνονται άνθρωποι.

8. πράξη που είναι και ανήθικη και παράνομη:

Η φοροδιαφυγή σε μια ευνομούμενη κοινωνία.

Ο σκοπός που παρουσιάζεται αυτό το σχήμα είναι για να δείξει ότι παρόλη την προσπάθεια νομικής περιφρούρησης της ανθρώπινης δραστηριότητας και της εξάλειψης των κινδύνων, παραμένει μια περιοχή όπου δεν έχει ακόμη καλυφθεί νομικά αλλά και μια περιοχή που ενώ έχει καλυφθεί νομικά εξακολουθεί να παραβιάζεται για λόγους που η ηθική θεωρεί σημαντικότερους, αλλά και μια περιοχή που κάλυψε ανήθικες πράξεις.

Θα αναρωτηθεί βέβαια κανείς γιατί πρέπει ένας ηθικός κανόνας να γίνει νόμος ώστε να τηρείται, και χωρίς αυτόν να παραβιάζεται; Προφανώς η απάντηση βρίσκεται στα όρια του “εσωτερικού νόμου της ελευθερίας” όπως ονομαζόταν στην αρχαία Σπάρτη (Ηροδότου ιστορίες, Διάλογος Ξέρξη Δημάρατου, και “Οι δωριστί αρμονίες. Το αληθινό πρόσωπο της αρχαίας Σπάρτης” Δ. Νικολακάκης. 1998) η οποία είναι μια αμετακίνητη εσωτερική αντίληψη, που βάζει καθένas στον εαυτό του, ενός δηλαδή **εσωτερικού ψυχολογικού ορίου**, που μπορεί κάποιος να φτάσει και να αισθάνεται ικανοποιημένος με τον εαυτό του. Μια καλή έρευνα γύρω από αυτό το πεδίο θα ήταν το πώς η παιδεία επηρεάζει αυτή την προσωπική κατάσταση, και αν θα μπορούσε να ήταν ακόμη ένας τρόπος για την επιρροή της ηθικής, ιδιαίτερα σε μικρά παιδιά που τώρα διαμορφώνουν τους χαρακτήρες τους. Προφανώς η σχετική επιστήμη (παιδαγωγική) θα έχει δώσει τις κατάλληλες απαντήσεις (Thomas E. Wartenberg. 2014) σε αυτό το ερώτημα, αλλά δεν είναι χρήσιμο να παρουσιαστούν στην παρούσα εργασία. Σε μια ιδανική κοινωνία λοιπόν θα υπήρχαν μόνοι ηθικοί κανόνες, και καθένas θα είχε θέσει μόνος του στον εαυτό του ένα εσωτερικό όριο

που δεν θα ξεπερνούσε ποτέ, όχι γιατί φοβόταν κάποιες συνέπειες αλλά γιατί θα γνώριζε ότι θα έκανε κακό. Επειδή όμως οι άνθρωποι δεν έχουν όλοι αυτή την εσωτερική δύναμη πρέπει να υπάρχουν νόμοι και συνέπειες για κάθε παράνομη πράξη, άρα και νόμοι που θα διέπουν την τεχνητή νοημοσύνη.

Αυτά τα σχήματα φανερώνουν ταυτόχρονα και την κατεύθυνση που πρέπει να έχουν οι προσπάθειες των ανθρώπων που είναι επιφορτισμένοι με την περιφρούρηση της ηθικής. Να συγκλίνουν δηλαδή τους κύκλους του ηθικού και νόμιμου, να αποκλίνουν τους κύκλους του ανήθικου και νόμιμου, επίσης να συγκλίνουν τους κύκλους παράνομου και ανήθικου, και να αποκλίνουν τους κύκλους του παράνομου και ηθικού. Προφανώς η επιδίωξη κάποιων ανήθικων ανθρώπων είναι ακριβώς η αντίθετη, με κυριότερη την προσπάθεια ανήθικες πράξεις να καλυφθούν με νομιμότητα υπό την κάλυψη της προστασίας αντιβαρικών δικαιωμάτων. Έτσι για παράδειγμα, ιδιαίτερα μετά την 9^η /11/2001, η λήψη δακτυλικών αποτυπωμάτων για τα διαβατήρια, η οποία από μόνη της είναι μια πράξη που παραπέμπει σε εγκληματικούς ανθρώπους, εφαρμόζεται σε όλους χωρίς εξαίρεση υπό τη κάλυψη της προστασίας έναντι τρομοκρατικών χτυπημάτων. Βλέπουμε λοιπόν ότι δεχόμαστε λόγω των συνθηκών κάποια υποχώρηση στα ανθρώπινα δικαιώματα (εδώ στη ελευθερία και την ελεύθερη μετακίνηση) ή το λεγόμενο μικρότερο κακό (Shor Eran, Ina Filkobskia, Pazit Ben-Nun Bloomb, Hayder Alkilabia, William Su. 2016), ή κατά άλλους υπομένουμε μια βλάβη στα δικαιώματά μας προκειμένου να προστατέψουμε άλλα ανώτερα σε σύγκριση δικαιώματα.

Μια παρόμοια ανάγνωση των κύκλων της ηθικής είναι ότι πολλές φορές η νομιμότητα αναγκάζεται να κλείσει τα μάτια της στην ηθική, συγκεκριμένα η απόφαση ότι η θανάτωση ενός (αδίκως επιτιθέμενου) ανθρώπου, κατά την διάρκεια ενός πολέμου είναι μια νόμιμη πράξη, είναι μια υποχώρηση της ηθικής που όμως εξυπηρετεί τους στόχους που έθεσε ο Αριστοτέλης για την επιδίωξη του κοινού καλού. Αυτή είναι και μια συζήτηση, στο μεταίχμιο της ηθικής αναζήτησης, γύρω από τα **αυτόνομα όπλα**, που θα αναθεωρήσει τον τρόπο που μετρούμε την αξία της ανθρώπινης ζωής.

Επίσης παραπάνω (στην 7η περίπτωση) επισημαίνεται ότι ελλοχεύει ο κίνδυνος της “μεροληπτικής νομιμότητας” η οποία εν τέλει καθίσταται ανήθικη, το γεγονός δηλαδή ότι έχει προστατευθεί νομικά μια ανήθικη περιοχή πράξεων που πρέπει να αποκαλυφθεί. Αυτός είναι και ένας μεγάλος κίνδυνος σε ολοκληρωτικά συστήματα, όπως π.χ. το Κινέζικο Σύστημα κοινωνικής βαθμολόγησης, να δημιουργηθεί μια ελίτ με ιδιαίτερα δικαιώματα που θα ζει ανήθικα (με ιδιαίτερα προνόμια) και χωρίς να υπακούει στους “ηλεκτρονικούς περιορισμούς” των απλών πολιτών. Δηλαδή κάτι παρόμοιο που βλέπουμε στην Ελληνική πραγματικότητα και ιδιαίτερα στην απονομή της δικαιοσύνης, ότι δηλαδή υποθέσεις ατόμων, που αφορούσαν υπεξαίρεση μεγάλων χρηματικών ποσών του δημοσίου, έμειναν ατιμώρητες, ενώ μικροεγκλήματα επωμίζονται τεράστιες ποινές (π.χ. η θέση στο αρχείο της υπόθεσης των δανείων των κομάτων από δημόσια τράπεζα μετά την φήφιση απαλλακτικού νόμου για τα ταπεζικά στελέχη (ν. 4146/2013), η απάτη του

χρηματιστηρίου που έκανε 18 χρόνια να τελεσιδικήσει χωρίς να απαντηθεί ποιος ήταν υπεύθυνος για την διόγκωση των τιμών των μετοχών, η λίστα Lagarde που έκλεισε με παραγραφή (ΣτΕ 2934/2017 και ΑΑΔΕ ΠΟΛ.1194/2017) αφού ήδη είχε καταδικαστεί υπουργός για νόθευση εγγράφου κ.α.). Τα παραπάνω παραδείγματα αναφέρονται διότι μετά την πιθανή υιοθέτηση μιας εφαρμογής που θα εξυπηρετεί το δημόσιο αυτές οι ανισότητες θα προστατευθούν τεχνολογικά.

Προφανώς κάποιος θα αναρωτηθεί ποια είναι η άμεση σχέση του θέματος της εργασίας με την μελέτη που αναπτύσσεσαι αυτό το υποκεφάλαιο; Ο πιο σημαντικός κίνδυνος σε μια κοινωνία είναι η εμφάνιση του φαινομένου που παρουσιάζεται στην 7^η περίπτωση, ότι δηλαδή άνθρωποι στον δημόσιο τομέα και ιδιαίτερα στην κυβέρνηση έχουν την δυνατότητα να ψηφίζουν νόμους που είναι μεροληπτικοί και εντέλει άδικοι. Ο ρόλος της τεχνητής νοημοσύνης σε αυτή την διαδικασία θα είναι ο ρόλος της εφαρμογής αυτών των άδικων νόμων εις βάρος των υπολοίπων πολιτών, αλλά και ο ρόλος της δημιουργίας και εξυπηρέτησης μιας ελίτ που απολαμβάνει ιδιαίτερη μεταχείριση. Επισημαίνεται δηλαδή ο κίνδυνος μιας κακής χρήσης της τεχνητής νοημοσύνης μέσω μιας κακής χρήσης της νομοθετικής εξουσίας εις βάρος της ηθικής. Για να μην υπάρχουν λοιπόν τέτοιες εκπλήξεις θα πρέπει η δράση των ανθρώπων να είναι προμηθευτική.

Κλείνοντας να αναφέρουμε ότι προς την κατεύθυνση αυτή έχουν προταθεί από διάφορους φορείς πολλές προτάσεις **ηθικών αρχών** που μπορούν να αποτελέσουν μια καλή **βάση** για δημιουργία κατάλληλου νομικού πλαισίου για την τεχνητή νοημοσύνη. Οι αρχές αυτές θα αναφερθούν στο σχετικό κεφάλαιο.

7.4. Περίληψη

Αφού έχει αποδειχτεί ότι η χρήση της τεχνητής νοημοσύνης κρύβει κινδύνους για την ζωή μας, θα πρέπει να απαντηθεί το πώς αυτοί οι κίνδυνοι θα εξαληφθούν. Σε επίπεδο πρόληψης μια καλή λύση είναι ένα ηθικό πλαίσιο αρχών στο οποίο θα υπακούουν μηχανές και άνθρωποι. Όσον αφορά τις μηχανές υπάρχουν τρεις τρόποι αυτό να γίνει: με εισαγωγή αυτών των αρχών απευθείας σε επίπεδο κώδικα, με το να αφεθούν “ελεύθερες” να αναπτύξουν δικό τους ηθικό κώδικα μέσω της διαδικασίας της μάθησης μηχανών, ή με έναν συνδυασμό των δύο προηγούμενων μεθόδων. Όσον αφορά τον άνθρωπο, η υπακοή σε ένα σύνολο αρχών είτε μπορεί να επιβληθεί μέσω της νομοθεσίας είτε να καλλιεργηθεί η έμφυτη τάση να αποφεύγει το κακό μέσω της παιδείας.

8. Ελπιδοφόρες εφαρμογές και chatbots στην υπηρεσία του δημοσίου τομέα

8.1. Εισαγωγή

Ακόμη και στον τομέα της τεχνητής νοημοσύνης δεν θα μπορούσε παρά να επαληθευτεί η γνώμη των στωικών φιλοσόφων για τη σύνδεση του καλού με το κακό. Αυτό σημαίνει ότι παρ' όλους τους κινδύνους που συνοδεύουν την τεχνητή νοημοσύνη υπάρχουν και εφαρμογές που είτε οι κίνδυνοι είναι ασήμαντοι είτε τα πλεονεκτήματά τους είναι περισσότερα από τα μειονεκτήματά τους. Σε αυτό το κεφάλαιο θα γίνει μια αναφορά σε τέτοιες εφαρμογές.

8.2. Ελπιδοφόρες εφαρμογές της τεχνητής νοημοσύνης

Εκτός από τις ανησυχίες που προκαλεί η τεχνητή νοημοσύνη σε πολλές περιπτώσεις, σε κάποιες άλλες γεννά ελπίδες.

Ένας από τους τομείς του δημοσίου τομέα που η τεχνητή νοημοσύνη θα δώσει, χωρίς κανένα ηθικό εμπόδιο, εφαρμογές χρήσιμες για την ζωή μας είναι η υγεία και συγκεκριμένα η αντιμετώπιση δύσκολων ασθενειών αλλά και η σωστή διαχείριση του τεράστιου όγκου ασθενών στα νοσοκομεία. Αν και η αντιμετώπιση μιας ασθένειας έχει πολλές μεταβλητές στοχαστικού τύπου και όχι ντετερμινιστικού (Stefano A. 2018), η τεχνητή νοημοσύνη έχει ήδη δώσει πολλά χρήσιμα αποτελέσματα.

Ένα τέτοιο παράδειγμα είναι η εφαρμογή της εταιρείας IBM που ονομάζεται IBM Watson Health. Η εφαρμογή αυτή αφού έχει ενθουλακώσει μια τεράστια εμπειρία περιπτώσεων ασθενών με καρκίνο, εξάγει σαν αποτέλεσμα την καταλληλότερη θεραπευτική αγωγή για συγκεκριμένο ασθενή αφού δεχθεί σαν είσοδο προσωπικά χαρακτηριστικά του ασθενή, όπως γονιδίωμα, ιατρικό ιστορικό, εικόνα. Επιπλέον, ήδη διερευνούνται μέθοδοι (όπως π.χ. το "100,000 Genomes Project") οι οποίες με βάση την μάθηση (machine learning) μπορούν να προβλέπουν τις ασθένειες που πιθανώς θα αντιμετωπίσει στο μέλλον ένας άνθρωπος στο μέλλον, μόνο αναλύοντας τον γενετικό του κώδικα, και παράλληλα θα μπορεί σε κάθε άνθρωπο να δοθεί μια προσωποποιημένη ιατρική φροντίδα διαφορετική για τον καθένα. Σε αυτή τη περίπτωση εγείρετε ένα πολύ σπουδαίο ηθικό ερώτημα: τι θα συμβεί αν η γνώση του γενετικού κώδικα ενός πληθυσμού πέσει στα χέρια κακόβουλων και εκτός από την θεμιτή χρήση του γενετικού κώδικα υπάρξει και αθέμιτη χρήση του; Μήπως οι κίνδυνοι από την διαρροή αυτών των δεδομένων είναι μεγάλοι και θα πρέπει να μην μπει η έρευνα καθόλου σε αυτό το πεδίο; Μια λύση σε αυτό το πρόβλημα θα ήταν η ανωνυμοποίηση των συλλεγόμενων στοιχείων και η ισχυρή κρυπτογράφηση τους. Προχωρώντας ακόμη περισσότερο στο πεδίο των κινδύνων βλέπουμε ότι ήδη η ανάλυση του γονιδιώματος έχει οδηγήσει στην δυνατότητα επιλογής χαρακτηριστικών σε βρέφη (Ronald T.K. Pang P.C.Ho. 2016) κάτι που στο παρελθόν ανήκε στην σφαίρα της επιστημονικής φαντασίας. Όσο βέβαια η επιστήμη της

Γενετικής Μηχανικής προσπαθεί να αποτρέψει την κληροδότηση ασθενειών από τους γονείς στα παιδιά τους κανένας ηθικός κανόνας δεν θίγεται, όμως αν η επιστήμη προχωρήσει σε άλλου είδους πειράματα, όπως στο παρελθόν η προσπάθεια δημιουργίας αρείας φυλής από το γερμανικό καθεστώς, τότε κανόνες της ηθικής θίγονται άμεσα και πρέπει να μπου απαγορεύσεις.

Να σημειωθεί εδώ ότι η επεξεργασία ιατρικών ανωνυμοποιημένων προσωπικών δεδομένων επιτρέπεται για λόγους έρευνας.

Ένα από τα ιατρικά προγράμματα τεχνητής νοημοσύνης που συγχρηματοδοτεί η ΕΕ είναι το πρόγραμμα MURAB (MRI and Ultrasound Robotic Assisted Biopsy) το οποίο αναπτύσσει τεχνολογία ακριβούς και αποδοτικής λήψης δειγμάτων ιστού (βιοψία) με σκοπό την καλύτερη διάγνωση της ασθένειας του καρκίνου.

Στον αγώνα αντιμετώπισης του καρκίνου έρχεται να βοηθήσει και το κινεζικό πρόγραμμα μάθησης μηχανής (deep learning) με την ονομασία Infervision. Η εφαρμογή αυτή επικουρεί ακτινολόγους στην γρήγορη και έγκυρη επόπτευση ακτινογραφιών για τον καρκίνο των πνευμόνων, που είναι μια συχνή ασθένεια στον κινεζικό πληθυσμό. Οι επιστήμονες που το δημιούργησαν ευελπιστούν σε μια παγκόσμια αποδοχή της εφαρμογής τους.

Όσον αφορά την αποδοχή των συστημάτων διάγνωσης από τον ιατρικό κόσμο μελέτες έχουν δείξει ότι αποδέχονται την αξιοπιστία τους και τα θεωρούν πολύ βοηθητικά στο δύσκολο έργο της διάγνωσης (Azamat Amirov, Olga Gerget, Dmitry Devjatyh b, Arstan Gazaliev. 2013). Αυτό σημαίνει ότι βρισκόμαστε στο σημείο όπου η τεχνητή νοημοσύνη στον χώρο της υγείας έχει γίνει αποδεκτή, σαν επικουρική, από τους λειτουργούς της. Μένει ακόμη να δούμε τις αντιδράσεις όταν η τεχνητή νοημοσύνη θα φτάσει στο σημείο να κάνει διαγνώσεις χωρίς την τελική επιβεβαίωση ιατρών.

Ένα ακόμη πρόγραμμα που συγχρηματοδοτεί η ΕΕ και αφορά την τεχνητή νοημοσύνη και τον δημόσιο τομέα είναι το πρόγραμμα BRIDGET (<https://ec.europa.eu/digital-single-market/en/news/bridget-eu-funded-project-success-story>), το οποίο παρέχει εξελιγμένα εργαλεία αναγνώρισης πληροφοριών από οπτικά δεδομένα μεγάλης κλίμακας. Το πρόγραμμα αυτό θυμίζει την απαρχή ενός παρόμοιου προγράμματος με το πρόγραμμα “Κοινωνικής Βαθμολόγησης” της Κίνας (με όλη τη συζήτηση που έχει εγερθεί γύρω από αυτό) και γεννά την εύλογη απορία: γιατί η ΕΕ επενδύει σε αυτή την τεχνολογία, έχει σαν σκοπό την εγκατάσταση ενός παρόμοιου συστήματος με αυτό της Κίνας; Αν και οι σκοποί της ΕΕ, όπως αυτοί παρουσιάζονται στην πηγή που αναφέρθηκε, αφορούν εντοπισμό μη ανθρώπινων στόχων (monuments, locations, buildings, products, books, logos or objects) ας ελπίσουμε ότι η χρήση τέτοιων συστημάτων επεκταθεί και σε ανθρώπους θα είναι προς την καλή κατεύθυνση, για παράδειγμα η παροχή βοήθειας σε άτομα με ειδικές ανάγκες που βρίσκονται στον δρόμο, η ανεύρεση ατόμων που έχουν χαθεί, η ενίσχυση της ασφάλειας στους δρόμους κ.α. Ακόμη και αν οι σκοποί είναι ευγενείς

θα πρέπει να αποκλείεται εξ αρχής η κακή χρήση ενός τέτοιου προγράμματος ή και να αποκλείεται εντελώς όπως στην περίπτωση του προγράμματος Maven που έχει ήδη αναφερθεί.

Μια ακόμη εφαρμογή τεχνητής νοημοσύνης που θα διευκολύνει έναν τομέα του δημοσίου, συγκεκριμένα τον τομέα της ασφάλειας και πρόληψης είναι η ανάπτυξη εφαρμογών αναγνώρισης της ρητορικής μίσους και πρόληψης τρομοκρατικών επιθέσεων και γενικότερα η ανάπτυξη ενός τομέα που θα λέγεται “προληπτική αστυνόμευση” (Jeremy W. Crampton, Eric M. Huntley, Emily C. Kaufman. 2017) μέσω της δυνατότητας που δίνουν σύγχρονα συστήματα τεχνητής νοημοσύνης για την αναγνώριση συνδυασμού συναισθημάτων (Anne Håkansson. 2013). Ένα ακόμη περιστατικό που επιβεβαιώνει την ύπαρξη τέτοιας τεχνολογίας είναι το απόσπασμα της κατάθεσης ιδιοκτήτη μέσου κοινωνικής δικτύωσης στο Αμερικανικό Κογκρέσο ότι ήδη έχει αναπτυχθεί τέτοιο λογισμικό και ήδη τα δεδομένα των χρηστών εξετάζονται με βάση το κριτήριο της ύπαρξης ρητορικής μίσους και της ύπαρξης προετοιμασίας τρομοκρατικών επιθέσεων. Θα πρέπει να επισημανθεί εδώ ότι σε αυτές τις εφαρμογές εγείρεται ένα μεγάλο ηθικό ζήτημα: μπορούμε για λόγους ασφάλειας να συλλέγουμε και να επεξεργαζόμαστε προσωπικά δεδομένα, όπως εικόνα και λόγο, και επομένως να περιορίζουμε την ελευθερία των πολιτών; Στην μεν περίπτωση του ιδιωτικού κοινωνικού δικτύου αυτό το ηθικό δίλλημα καλύπτεται από την αυτόβουλη παραχώρηση των δεδομένων προσωπικού χαρακτήρα από τους ιδίους τους χρήστες με την συγκατάθεσή τους στην “πολιτική και όρους χρήσης” της πλατφόρμας, στην δε περίπτωση των δημοσίων αρχών ασφαλείας υπάρχει η νομική κάλυψη του υπέρτατου δημοσίου συμφέροντος, αρκεί αυτό να μην καταρρίπτεται από μια μελέτη εκτίμησης επιπτώσεων, όπως θα αναφερθεί ειδικότερα σε επόμενο κεφάλαιο. Όμως παράλληλη την νομιμότητα εξακολουθεί να υπάρχει το ερώτημα “ότι είναι νόμιμο είναι και ηθικό;”. Τα ερώτημα αυτό έχει ήδη αναλυθεί στο κεφάλαιο για τη σχέση της ηθικής με την νομιμότητα. Ηγετικό ρόλο στην προληπτική αστυνόμευση παίζει το κράτος του Ισραήλ, το οποίο σύμφωνα με τον οργανισμό Privacy International έχει το μεγαλύτερο ποσοστό εταιρειών λογισμικού κατασκοπείας ανά άτομο σε όλο τον πλανήτη. Οποιαδήποτε κίνηση κάνει κάθε Παλαιστίνιος στην Δυτική Όχθη, ή ανεβάσει κάτι στα κοινωνικά δίκτυα καταγράφεται και αναλύεται από το σύστημα επιτήρησης. Λογισμικό παρακολούθησης που παράγεται από Ισραηλινές εταιρείες ήδη έχει εξαχθεί σε άλλες χώρες (Ηνωμένα Αραβικά Εμιράτα, Αργεντινή) που επιθυμούν να εγκαταστήσουν παρόμοια συστήματα παρακολούθησης. Όμως παρά όλα τα οφέλη αυτή η εντατική αστυνόμευση μπορεί ίσως να μεταφραστεί και σαν μια ηλεκτρονική δικτατορία.

Ένας ακόμη τομέας που θα ωφεληθεί από την τεχνητή νοημοσύνη είναι η ασφάλεια των ωκεανών, που θα ενισχυθεί με αυτόνομα ναυτικά σκάφη (saildrones by National Oceanic and Atmospheric Administration, UNITED STATES DEPARTMENT OF COMMERCE) τα οποία θα περιπολούν στις θάλασσες και θα συλλέγουν στοιχεία για αυτές (κυματισμό, ανέμους, ηλιοφάνεια, θερμοκρασίες, βάθη, περιβαλλοντολογικές παραμέτρους όπως μόλυνση κλπ) καθώς και θα επιτηρούν για

παράνομη αλιεία. Βλέπουμε σε αυτή τη περίπτωση πως μια εφαρμογή τεχνητής νοημοσύνης έχει μόνον θετικές αντιδράσεις και κανένα ηθικό δίλλημα αν θα πρέπει να εφαρμοστεί ή όχι, γιατί αποσκοπεί σε ένα κοινά αποδεκτό κοινωνικό συμφέρον.

Το δημόσιο έχει επίσης στην αρμοδιότητα του τον έλεγχο της παράνομης αλιείας στα εγχώρια ύδατα. Μια εφαρμογή τεχνητής νοημοσύνης της Google που θα μπορούσε να φανεί χρήσιμη στο λιμενικό σώμα είναι η Global Fishing Watch. Πρόκειται για ένα μεγάλο σύνολο αισθητηρίων σε πλοία που με την βοήθεια δορυφορικών λήψεων παρέχει στους διαχειριστές πληροφόρηση για την θέση κάθε πλοίου και μια δικαιολογία της πορείας του.

Μια ακόμη εφαρμογή της τεχνητής νοημοσύνης που θα διαχειρίζεται ο δημόσιος τομέας είναι η δυνατότητα πρόβλεψης των καιρικών συνθηκών (πηγή <https://web.stanford.edu/class/cs221/2017/restricted/p-final/cvo9/final.pdf>).

Επισημαίνεται ότι τα προγνωστικά αυτά μοντέλα δεν βασίζονται σε προβλέψεις προσομοίωσης με φυσικές μελέτες όπως γινόταν στο παρελθόν, αλλά στην διαδικασία μάθησης μηχανών (machine learning). Με τις εφαρμογές της τεχνητής νοημοσύνης η επιτυχία πρόβλεψης του καιρού της επόμενης ημέρας ανήλθε στο ποσοστό του 90% (Weather Monitoring Using Artificial Intelligence, IEEE, Computational Intelligence and Networks (CINE). 2016). Να σημειωθεί εδώ ότι στην χώρα μας η πανελλαδική συγκέντρωση στοιχείων για τις καιρικές συνθήκες είναι αποκλειστική αρμοδιότητα του δημοσίου (EMY). Σε Ευρωπαϊκό επίπεδο υπεύθυνη για την πρόβλεψη καιρού είναι η διακρατική ECMWF. Η περίπτωση του τυφώνα Sandy όπου ένα διαγνωστικό μοντέλο τεχνητής νοημοσύνης έδινε στην διαδρομή του τυφώνα την πολιτεία της Μασαχουσέτης ενώ στη πραγματικότητα αυτός πέρασε πάνω από την πολιτεία της Πενσυλβάνια (Nick P. Bassill.2014), δείχνει την ανάγκη παροχής ιδιαίτερης προσοχής όταν η τεχνητή νοημοσύνη αφορά κρίσιμα ζητήματα όπως η ανθρώπινη ζωή και οι περιουσίες, αλλά και το μέγεθος των συνεπειών ενός λάθους στην διαδικασία της πρόγνωσης. Εφόσον στο μέλλον αναπτυχθούν αξιόπιστα μοντέλα πρόβλεψης του καιρού, οι πιθανές εφαρμογές τους θα βοηθήσουν πολλούς τομείς όπως οι αεροπορικές μεταφορές, οι ναυτικές μεταφορές, οι αγροτικές καλλιέργειες κ.αλ.

Ο αγροτικός τομέας, ο οποίος είναι θεμελιώδης για την λειτουργία ενός κράτους, αν και έχει ιδιωτικοποιηθεί πλήρως, η κεντρική διαχείριση του παραμένει στον έλεγχο του δημοσίου (καθορισμός δυναμικών καλλιεργειών, καθορισμός ύψους επιδοτήσεων, διαχείριση και αντιμετώπιση ασθενειών, έλεγχος εισαγωγών, έλεγχος ζωικού κεφαλαίου κλπ). Πέρα από τις εφαρμογές της τεχνητής νοημοσύνης που αφορούν και διευκολύνουν την δουλειά των αγροτών (αυτόνομοι γεωργικοί ελκυστήρες, αυτόνομα αεροσκάφη επιτήρησης και ψεκασμού ακριβείας κλπ) που αφορούν τον ιδιωτικό τομέα και δεν αφορούν το αντικείμενο της παρούσας εργασίας, η τεχνητή νοημοσύνη μπορεί να προσφέρει πολλά στα αρμόδια υπουργεία αγροτικής ανάπτυξης των χωρών. Έτσι στο μέλλον, διαγνωστικά μοντέλα τεχνητής νοημοσύνης

θα μπορούν μέσω δορυφορικών λήψεων, να προβλέπουν κλιματικές αλλαγές και να προτείνουν αλλαγές καλλιεργειών, κατάλληλες καλλιεργητικές πρακτικές, κατάλληλη και έγκυρη αντιμετώπιση ασθενειών κ.αλ. Για παράδειγμα μια πρόβλεψη μετατόπισης της εύκρατης ζώνης βορειότερα στην χώρα μας θα έχει ως συνέπεια, να ενημερωθούν εγκαίρως οι αγρότες ώστε να φυτευτούν εγκαίρως στη βόρεια χώρα είδη που ευδοκιμούσαν σε νοτιότερες περιοχές, και στο εγγύς μέλλον να προσαρμόζονται ικανοποιητικά (παράδειγματα εταιρειών που παρέχουν τέτοιες υπηρεσίες: aWhere, FarmShots). Ένα βήμα περισσότερο προχωρά η πρόταση της IBM (<https://www.ibm.com/blogs/watson/2016/12/five-ways-agriculture-benefit-artificial-intelligence/>) για την δημιουργία συμβουλευτικών chatbots, τα οποία θα μπορεί να απασχολεί το δημόσιο και θα παρέχουν προσωποποιημένες συμβουλές προς τους αγρότες.

Ακόμη ένας από τους τομείς που θα επηρεαστεί στο μέλλον από την τεχνητή νοημοσύνη και ελέγχει το δημόσιο είναι ο έλεγχος της εναέριας κυκλοφορίας, μέσω της Υπηρεσίας Πολιτικής Αεροπορίας. Διεθνώς έχουν γίνει προσπάθειες ώστε να αναπτυχθεί μια εφαρμογή τεχνητής νοημοσύνης η οποία θα λειτουργεί συμβουλευτικά προς τους ελεγκτές της εναέριας κυκλοφορίας και τους πιλότους αεροσκαφών, αναλύοντας παρελθούσες καταστάσεις και δίνοντας βέλτιστες λύσεις σε προβλήματα που απασχολούν τους υπευθύνους μιας πτήσης (The Generalized Integrated Learning Architecture (GILA) project – being developed by Lockheed Martin's Advanced Technology Laboratories).

Ακόμη ένας τομέας που την διαχείριση του έχει το δημόσιο και θα επηρεαστεί από την τεχνητή νοημοσύνη είναι ο τομέας της ενέργειας. Αν και το Ελληνικό δημόσιο έχει απεμπλακεί πλήρως από την λιανική πώληση της ηλεκτρικής ενέργειας, ακόμη διατηρεί των έλεγχο των τιμών και των πλουτοπαραγωγικών πηγών. Οι ιδιωτικές εταιρείες του κλάδου που είναι ταυτόχρονα και οικονομικοί κολοσσοί έχουν αναπτύξει χρήσιμα εργαλεία που θα μπορούσε το δημόσιο να χρησιμοποιήσει για να διατηρεί μια γενική εικόνα τόσο των καταναλώσεων όσο και των πηγών. Τέτοια εργαλεία έχουν αναπτύξει η πετρελαϊκή BP και η GE Power. Πρόκειται για ένα δίκτυο από αισθητήρες παρόμοιο αυτού του διαδικτύου των πραγμάτων (IoT) που δημιουργεί ένα "δίκτυο ενέργειας" κατάλληλο να επιβλέπει την παραγωγή, διύλιση και ζήτηση ενέργειας σε μια χώρα. Μια εφαρμογή τεχνητής νοημοσύνης που εξυπηρετεί την εξοικονόμηση ηλεκτρικής ενέργειας είναι αυτή που χρησιμοποιεί η πόλη Jacksonville όπου ένα σύστημα καμερών εντοπίζει τον όγκο της κίνησης πεζών και οχημάτων στους δρόμους και ρυθμίζει ανάλογα την ένταση στον φωτισμό των οδών (Andreas Kaplan, Michael Haenlein. 2019).

Όσον αφορά την δημόσια υγεία έχει αναπτυχθεί μια εφαρμογή από τον τομέα της δημόσιας υγείας της Νότιας Νεβάδα των ΗΠΑ, η οποία χρησιμοποιεί δεδομένα φωνητικά και εικόνας, που ανεβάζουν οι χρήστες στα κοινωνικά δίκτυα, ώστε να εντοπίζει χώρους εστίασης που έχουν παρουσιάσει προβλήματα, και στην συνέχεια αυτές οι πληροφορίες χρησιμοποιούνται για επιτόπιους υγειονομικούς ελέγχους.

Σε πολλές χώρες του κόσμου η παιδεία παρέχεται από το δημόσιο στους πολίτες δωρεάν. Η τεχνητή νοημοσύνη μπορεί ήδη να προσφέρει πολλά και σε αυτόν τον τομέα. Στο τεχνολογικό πανεπιστήμιο της Georgia χρησιμοποιείται ένα βοηθητικό σύστημα βασισμένο σε τεχνητή νοημοσύνη για την εκπαίδευση των φοιτητών που ονομάζεται Jill Watson. Μάλιστα η απόδοση του συστήματος είναι τόσο καλή που οι φοιτητές δεν καταλαβαίνουν ότι συνομιλούν με έναν ψηφιακό εκπαιδευτή μέχρι να τους το αποκαλύψουν (πηγή <https://pe.gatech.edu/blog/meet-jill-watson-georgia-techs-first-ai-teaching-assistant>).

Ίσως η κορυφαία χρήση της τεχνητής νοημοσύνης θα είναι η ολοκλήρωση του στόχου της ηλεκτρονικής δημοκρατίας (e-democracy). Έχουν γίνει πολλά βήματα προς αυτή τη κατεύθυνση κυρίως σε ευρωπαϊκές χώρες. Ενώ στην Κίνα η χρήση παρόμοιου συστήματος χρησιμοποιείται μόνο για την διαχείριση των πολιτών (Yanliu Lin, 2018) σε άλλες ευρωπαϊκές χώρες δεν γίνεται σε τέτοιο βαθμό αυτή η κατάχρηση. Σύμφωνα με τα στοιχεία της ΕΕ (https://joinup.ec.europa.eu/sites/default/files/inline-files/eGovernment%20in%20Iceland%20-%20February%202016%20-%202013_v1_00.pdf) στην Ισλανδία έχει φτιαχτεί η καλύτερη υποδομή για την άμεση συμμετοχή των πολιτών στην λήψη των αποφάσεων. Αρχικά οι πολίτες της Ισλανδίας έχουν την δυνατότητα, και το κάνουν κατά πλειοψηφία, να συναλλάσσονται με το δημόσιο κυρίως μέσω του διαδικτύου. Όλοι οι Ισλανδοί έχουν αποκτήσει μια ψηφιακή ταυτότητα και έχουν παραλάβει το απαραίτητο ιδιωτικό και δημόσιο κλειδί ώστε να υπογράφουν ψηφιακά τα έγγραφά τους και να πιστοποιούνται από το δημόσιο σύστημα. Όμως ενώ το ήδη υπάρχον σύστημα δίνει την δυνατότητα να ζητείται η γνώμη τους ανά πάσα στιγμή από την κυβέρνηση και αυτή η γνώμη να είναι και ψήφος που με την πλειοψηφία θα αποφασίζει, αυτό δεν γίνεται. Σπάνια συλλέγονται γνώμες, όμως για όχι για τόσο σοβαρά θέματα, ώστε στη συνέχεια η κυβέρνηση να αποφασίσει. Βλέπουμε ότι στην Ισλανδία υπάρχει μια ανισόροπη ανάπτυξη και χρήση της τεχνολογίας. Η ολοκλήρωση μιας ηλεκτρονικής δημοκρατίας καθυστερείται για πολλά χρόνια ενώ η top-down διαχείριση των πολιτών έχει φτάσει στο κορυφαίο σημείο που θα μπορούσε να φτάσει. Να αναφερθεί εδώ πάλι ότι το τρίτο στάδιο ωρίμανσης της ηλεκτρονικής διακυβέρνησης (Government 3.0) είναι αυτό όπου η ηλεκτρονική διακυβέρνηση θα χρησιμοποιήσει την τεχνολογία της τεχνητής νοημοσύνης και την ανάλυση δεδομένων μεγάλης λίμακας.

Ένα ακόμη επίτευγμα της τεχνητής νοημοσύνης που μπορεί να βοηθήσει τον δημόσιο τομέα είναι η δυνατότητα της έγκαιρης ανίχνευσης ψευδών ειδήσεων και της στοχευόμενης παραπληροφόρησης και επιρροής των πολιτών μέσω προπαγάνδας (fake news detection). Με αυτή τη δυνατότητα θα προστατευτεί το κράτος από προσβολές στη δημοκρατία μέσω του αποκλεισμού κακόβουλης προπαγάνδας που γίνεται από political bots με σκοπό την χειραγώγηση της κοινής γνώμης μέσω πλημύρας δημοσιεύσεων στα μέσα κοινωνικής δικτύωσης. Τα political bots (Fenwick McKelvey, Elizabeth Dubois, 2017) είναι αυτοματοποιημένα προγράμματα

τα οποία χρησιμοποιούν την τεχνητή νοημοσύνη για να διδάχτουν τον τρόπο της χειραγώγησης της κοινής γνώμης από πραγματικούς πολιτικούς και στη συνέχεια δημιουργώντας ψεύτικους λογαριασμούς στα μέσα κοινωνικής δικτύωσης να διαχέουν μαζικά ψευδείς ειδήσεις και φήμες ή και να δημιουργούν συνθήκες cyberbullying (πληθώρα από κακόβουλη κριτική). Η προπαγάνδα αυτή είτε στρέφεται στους πολίτες του ίδιου του κράτους είτε σε πολίτες άλλων κρατών όπως η εσωτερική προπαγάνδα της Βραζιλίας για τις προεδρικές εκλογές του 2014, η εσωτερική προπαγάνδα για την μη έξοδο του Ηνωμένου βασιλείου από την ΕΕ, και η εξωτερική προπαγάνδα της Κίνας εναντίον της Ταϊβάν και η προπαγάνδα της Ρωσίας εναντίον της Ουκρανίας. Η τομέας της τεχνητής νοημοσύνης έχει δώσει σύγχρονα εργαλεία ανίχνευσης τέτοιων δημοσιευμάτων με διάφορους τρόπους, είτε αναλύοντας το περιεχόμενο των δημοσιευμάτων, είτε την διάχυσή τους, είτε λαμβάνοντας υπόψη πολλούς παράγοντες και δημιουργώντας ένα υβριδικό μοντέλο (Álvaro Figueira, Luciana Oliveira.2017, και Monther Aldwairi, Ali Alwahedi. 2018).

Στην ίδια κατηγορία μπορεί να ενταχθεί και η **άμυνα του κράτους έναντι απειλών DDoS κατά κρατικών ιστοσελίδων**. Βασίζεται και αυτή στην διάγνωση μιας έντασης στην διάχυση ενός μηνύματος, όπως και στην περίπτωση της ανίχνευσης ψευδών ειδήσεων.

Εικάζεται επίσης ότι η τεχνολογία blockchain σε συνδυασμό με την τεχνητή νοημοσύνη θα δώσει στους αρμόδιους φορείς των κρατών την τεχνολογική υποδομή ώστε να διεξάγουν μια **αξιόπιστη ηλεκτρονική εκλογική διαδικασία** προστατεύοντας την ιδιωτικότητα (Nir Kshetri. 2017). Επίσης ο συνδυασμός τεχνητής νοημοσύνης και blockchain θα δώσει μια μεγάλη βοήθεια στον τομέα της πρόβλεψης ασθενειών (Polina Mamoshina et all. 2017). Βέβαια η ανάλυση του τρόπου που το blockchain λειτουργεί δεν είναι αντικείμενο της παρούσας διατριβής και δεν θα αναλυθεί περαιτέρω.

Μεγάλη θα είναι επίσης η συμβολή της τεχνητής νοημοσύνης στην **πρόληψη και καταπολέμηση του εγκλήματος** με χρήση της διακυβέρνησης μέσω των κινητών τηλεφώνων m-government ή την χρήση των ετικετών (rfid) των διαβατηρίων για εντοπισμό ατόμων (Ignace Snellen, Marcel Thaens. 2008). Σε πολλές ήδη χώρες τις Ευρώπης (Δανία, Η. Βασίλειο) χρησιμοποιούνται έξυπνοι αλγόριθμοι που ειδοποιούν με μηνύματα στο κινητό τηλέφωνο πολίτες που, σε μια συγκεκριμένη στιγμή που έγινε κάποιο έγκλημα, βρισκόταν στον τόπο του εγκλήματος, είτε χρησιμοποιούν την εικόνα που δίνουν ανταποκριτές που συνεργάζονται με την αστυνομία σε όλη τη πόλη (Ολλανδία).

Ένας από τους τομείς του δημοσίου που θα επηρεαστεί τα επόμενα χρόνια από την τεχνητή νοημοσύνη είναι και η **εθνική άμυνα**. Μιας και τα πρώτα μη επανδρωμένα στρατιωτικά οχήματα, κυρίως υπτάμενα μέσα, έχουν ήδη χρησιμοποιηθεί σε πολέμους, το επόμενο βήμα όπου ανθρωποειδή θα συμμετέχουν σε αυτούς μάλλον δεν είναι πολύ μακριά. Να σημειωθεί εδώ ότι, σε αντίθεση με άλλους τομείς του δημοσίου τομέα, η εισαγωγή μηχανών τεχνητής νοημοσύνης

πρέπει να περνά από την έγκριση του αρμοδίου τομέα του ΟΗΕ που είναι η Συνθήκη για τα Συμβατικά Όπλα (Convention on Conventional Weapons (CCW)). Υπάρχει δηλαδή ήδη ένα αρχικό ηθικό φίλτρο, μιας και το ζήτημα της αφαίρεσης μιας ανθρώπινης ζωής, ακόμη και σε συνθήκες πολέμου, είναι κατά πολύ σημαντικότερο από άλλες εφαρμογές. Το θέμα των στρατιωτικών ρομπότ είναι το μεταίχμιο της ηθικής ανάλυσης όπου συναντώνται πολλές επιστήμες, όπως, η φιλοσοφία, η ανθρωπολογία, η θεολογία, η πληροφορική κ.αλ. Ένας βιαστικός αναλυτής θα ξεκινούσε από το ερώτημα; “θα επιτρέψουμε σε μηχανές να σκοτώνουν ανθρώπους;”. Όμως ένας ψυχραιμότερος αναλυτής θα σκεφτόταν πρώτα γιατί ήδη να επιτρέπεται σε ανθρώπους να σκοτώνουν ανθρώπους και γιατί το κάνουν; (Fiona Brookman. 2015). Εν συντομία οι άνθρωποι σκοτώνουν είτε γιατί αυτό ήταν μια ορθολογική απόφαση (rational choice perspective (RCP)) είτε κινούμενοι από εσωτερικά αίτια όπως άγχος, φόβος, ζήλια κ.αλ. Σε συνθήκες πολέμου ισχύει κυρίως το συναίσθημα “σκοτώσε πριν σκοτωθείς”. Επίσης η υποσυνείδητη σκέψη της άσκησης βίας εναντίον του κακού ή από την πλευρά του κακού η άσκηση βίας σε κατώτερο, είναι ένα κυρίαρχο κίνητρο σε συνθήκες πολέμου. Κρατώντας αυτό το συμπέρασμα μπορούμε να πούμε ότι οι απόψεις που υποστηρίζουν ότι τα ανθρωποειδή ρομπότ θα είναι καλύτερα στις μάχες από ότι ο άνθρωπος, είναι πράγματι βάσιμες. Αυτό γιατί πράγματι ένα ρομπότ δεν θα βρεθεί ποτέ σε κατάσταση φόβου ή άγχους ή τρόμου, επιπλέον θα είναι πιο ακριβές στην στόχευση, και έτσι θα μειωθούν οι απώλειες αμάχων (RONALD C. ARKIN. 2012). Η σημερινή όμως επικρατούσα άποψη είναι να μην επιτραπούν τα ανθρωποειδή στους πολέμους γιατί το ενδεχόμενο λάθους είναι υπαρκτό αλλά και το επίπεδο νοημοσύνης (όσο αυτό είναι γνωστό στο κοινό) δεν είναι ικανοποιητικό (Noel Sharkey) και επιπλέον δεν είναι ακόμη ηθικά παραδεκτό μηχανήματα να αφαιρούν ανθρώπινη ζωή. Και σε αυτή τη περίπτωση εφαρμογής της τεχνητής νοημοσύνης θα πρέπει να αναφερθεί το ερώτημα, αν κινδυνεύουμε περισσότερο από τις μηχανές ή από την ηθική αυτών που βρίσκονται πίσω από αυτές; Εάν μια μηχανή φέρει την ηθική του δημιουργού-κυβερνήτη της θα πράττει με τον ίδιο τρόπο με τον δημιουργό της (ηθικά ή ανήθικα), αν όμως αφεθεί (μέσω της αυτόνομης μάθησης) να αναπτύξει δική της ηθική ίσως και να αντιταχτεί σε μια ανήθικη απόφαση του κυβερνήτη της ή αντίθετα να πάρει μια δική της ανήθικη απόφαση. Με δεδομένο το τεράστιο πλήθος αμάχων που σκοτώνονται κάθε χρόνο σε πολέμους, αλλά και το τεράστιο πλήθος στρατιωτών που σκοτώνονται υπερασπιζόμενοι τα εδάφη τους, τις αρχές τους, και την αξιοπρέπεια τους, η εφαρμογή τεχνητής νοημοσύνης στον τομέα της άμυνας μέσω ανθρωποειδών, που θα φέρουν διεπαφή ηθικής, μόνο για αμυντικούς και αποτρεπτικούς σκοπούς δεν θα έβλαπτε καθόλου την ανθρωπότητα, αντιθέτως θα έσωζε χιλιάδες ζωές.

Μια εφαρμογή σχετική με την άμυνα είναι η εφαρμογή STARTLE, όπου με ανακοίνωσή του το Βρετανικού υπουργείου άμυνας ανέφερε ότι, η άμυνα του Ηνωμένου Βασιλείου θα ενισχυθεί με την εφαρμογή τεχνητής νοημοσύνης STARTLE η οποία θα εντοπίζει πιθανές απειλές στις βρετανικές θάλασσες και θα επικουρεί το βασιλικό ναυτικό (<https://www.naval-technology.com/news/newsrocket-to-supply-artificial-intelligence-software-for-uk-royal-navy-5033995/>). Το σύστημα

αυτό αρχικά εντοπίζει απειλές, στη συνέχεια αναλύει τις απειλές, και τέλος παράγει μια απόφαση που έχει συμβουλευτικό χαρακτήρα. Για αυτό το σύστημα έχουν διερευνηθεί πιθανές εφαρμογές του στον τομέα της ασφάλειας των υπολογιστικών δικτύων, για τον τομέα της υγείας αλλά και στον τομέα της διαχείρισης κυκλοφορίας αυτόνομων οχημάτων.

Ένας ακόμη τομέας του δημοσίου που θα ωφεληθεί από την χρήση των εφαρμογών της τεχνητής νοημοσύνης και παράλληλα τα αποτελέσματά του είναι και ηθικά επιθυμητά είναι η **διαχείριση της κυκλοφορίας των οχημάτων**. Η δυνατότητα ανάλυσης της κυκλοφορίας των παρελθόντων ετών και ταυτόχρονα η δυνατότητα ανάλυσης δεδομένων που δίνουν, σε πραγματικό χρόνο, αισθητήρες που είναι τοποθετημένοι σε μια πόλη κάνουν τεχνικά εύκολη μια πρόβλεψη για την ένταση της κυκλοφορίας οχημάτων που ακολουθεί, παρόλο που αυτή είναι μια στοχαστική μεταβλητή (Serge L.Shishkin. 2015). Ένα τέτοιο μοντέλο (π.χ. το Star City από την IBM) λειτουργεί ήδη στα δημόσια μέσα μεταφοράς στις πόλεις του Δουβλίνου, του Μαϊάμι, του Ρίο, και της Μπολόνια. Όλα τα δημόσια μέσα μεταφοράς είναι εφοδιασμένα με συστήματα γεωεντοπισμού και όλες οι διασταυρώσεις επιτηρούνται από κάμερες. Το σύστημα με διαρκείς αναλύσεις και τροφοδοσία από συστήματα όπως της κατάστασης των δρόμων, των μηνυμάτων των μέσων κοινωνικής δικτύωσης, των έργων που γίνονται στους δρόμους, των κοινωνικών εκδηλώσεων που επηρεάζουν την κίνηση, προτείνει λύσεις σε πιθανά μελλοντικά μπουτιλιάρια που πρόκειται να συμβούν, απαλλάσσοντας τους οδηγούς από πιθανή εμπλοκή σε δύσκολες και χρονοβόρες καταστάσεις. Ένα πιθανό ηθικό πρόβλημα που θα προέκυπτε αν στο παρόν δίκτυο οχημάτων θα ήθελε μια κυβέρνηση να προσθέσει και τα ιδιωτικά οχήματα με σκοπό να παρέχει σε όλους τους πολίτες μια γρήγορη μετακίνηση μέσα στις μεγαλουπόλεις, είναι το αίσθημα ότι ο οδηγός και οι επιβάτες παρακολουθούνται. Επίσης θα υπάρχει και ο πραγματικός κίνδυνος να παρακολουθείται πράγματι από κάποιους που θα θελήσουν να χρησιμοποιήσουν με ανήθικο τρόπο τα δεδομένα κίνησης. Με αυτό τον τρόπο το κοινό καλό (της διευκόλυνση της μετακίνησης) θα οδηγήσει στο να υποχωρήσει το δικαίωμα του ανθρώπου να μετακινείται ελεύθερος χωρίς να παρακολουθείται. Μια λύση που θα μετρίαζε λίγο αυτό το αίσθημα της παρακολούθησης, αλλά δεν θα το εξάλειφε, είναι η ανωνυμοποίηση των δεδομένων που προέρχονται από τον γεωεντοπισμό. Σε αυτή τη περίπτωση οι κίνδυνοι διαρροής δεδομένων θα υπήρχε μόνο από αυτούς που θα κατείχαν την συνάρτηση ανωνυμοποίησης των προσωπικών δεδομένων των αυτοκινήτων των ιδιωτών οδηγών και μόνο. Βέβαια επειδή η ηθική όπως αναφέρθηκε δεν είναι η ίδια σε όλους θα έλεγε κάποιος ότι σε καμιά περίπτωση δεν μπορεί να εκχωρηθεί μέρος της ελευθερίας του ατόμου για να διευκολυνθεί η μετακίνηση του. Αυτή είναι μια διαδικασία διάκρισης των δικαιωμάτων σε απόλυτα και σχετικά, ότι δηλαδή κάποια δικαιώματα δεν δέχονται κανέναν περιορισμό και κάποια άλλα δέχονται. Συνήθως όμως η έννοια της μεσότητας που είχε αναφερθεί στο πρώτο κεφάλαιο γέρνει προς, και παράλληλα φθείρει, την πλευρά του λιγότερα “απόλυτου” δικαιώματος, δηλαδή επικρατεί η επιλογή του μικρότερου κακού, ότι δηλαδή

κάποιος υπομένει λίγη καθυστέρηση στους δρόμους προκειμένου να μην επιτρέψει σε κανένα να τον παρακολουθεί.

Ένας ακόμη τομέας του δημοσίου που μπορεί να ωφεληθεί από την τεχνητή νοημοσύνη είναι η πυροσβεστική υπηρεσία. Το τμήμα Atlanta Fire Rescue Department (AFRD) των Ηνωμένων Πολιτειών με την βοήθεια ερευνητών των πανεπιστημίων Georgia Institute of Technology, Emory University, University of California, Irvine, ανέπτυξε λογισμικό κατάλληλο για ανάλυση και πρόβλεψη πυρκαγιών σε κτίρια. Το λογισμικό αυτό ονομάζεται Firebird (Michael Madaio, Shang-Tse Chen, Oliver L. Haimson, Wenwen Zhang, Xiang Cheng, Matthew Hinds-Aldrich, Duen Horng Chau, Bistra Dilkina. 2016), λαμβάνει υπ' όψιν ιστορικά δεδομένα 58 μεταβλητών για κτίρια τα οποία ήδη κατέχει η υπηρεσία AFRD. Μερικές από τις μεταβλητές που λαμβάνει υπόψη είναι: τοποθεσία, ιστορικό πυρκαγιών της περιοχής, μέγεθος κτηρίου, δομή του κτηρίου, χρονιά κατασκευής. Το λογισμικό ήδη απαριθμεί 5000 κτήρια που έχει ελέγξει και σύμφωνα με τους κατασκευαστές έχει ποσοστό ακρίβειας στις προβλέψεις 73%.

Αν και η γενικότερη αίσθηση που αποπνέεται από την έρευνα γύρω από την ηθική της τεχνητής νοημοσύνης είναι αρνητική, αυτό δεν έχει σκοπό τον συντηρητισμό και την οπισθοδρόμηση αλλά μόνον την λήψη κατάλληλων μέτρων προφύλαξης έναντι πιθανών κινδύνων. Επίσης αυτό δεν μειώνει την σπουδαία συμβολή της τεχνητής νοημοσύνης στην βελτίωση της ποιότητας της ζωής των πολιτών.

Συμπεραίνεται ότι η τεχνητή νοημοσύνη έχει πολλά καλά να δώσει στην κοινωνική ζωή αλλά δυστυχώς αυτό είτε εμποδίζεται από τους ανθρώπους που κρατούν στα χέρια τους το δικαίωμα λήψης των αποφάσεων, είτε χρησιμοποιείται ως όχημα για να επιτευχθούν κακοί σκοποί.

8.3. Αναφορά σε chatbots που ήδη χρησιμοποιούνται για παροχή δημοσίων υπηρεσιών ή θα μπορούσαν να χρησιμοποιηθούν από τον δημόσιο τομέα

Όπως είχε αναφερθεί συνοπτικά σε προηγούμενο κεφάλαιο η τεχνολογία της αναγνώρισης της γλώσσας αλλά και της διαδικασίας της μάθησης μηχανών είναι η βάση της ανάπτυξης των μηχανών που είναι γνωστές με τον όρο chatbots ή chaterbots (Antonella De Angeli. 2009), που είναι μια σύνθεση των λέξεων chat και robots, δηλαδή ρομπότ τα οποία έχουν την δυνατότητα να συνομιλούν. Τα ρομπότ αυτά είτε έχουν μια μη ανθρώπινη φυσική μορφή είτε τους έχουν δοθεί χαρακτηριστικά ανθρώπου, όπως κεφάλι, άκρα κλπ. Επιπλέον πολλές εταιρείες εφάρμοσαν και μη λεκτικά χαρακτηριστικά όπως εκφράσεις του προσώπου, κινήσεις του σώματος, χειρονομίες και άλλα. Τα ρομπότ αυτά δέχονται σαν είσοδο κείμενο ή εικόνα, αναγνωρίζουν και επεξεργάζονται αυτά τα δεδομένα, και στη συνέχεια δίνουν μια έξοδο που είναι μια λεκτική απάντηση ή μια άλλη πράξη. Στη συνέχεια θα γίνει

αναφορά σε κάποια chatbots που είτε χρησιμοποιούνται είτε μπορούν να χρησιμοποιηθούν από τον δημόσιο τομέα.

Το chatbot GovBot στην Γερμανία είναι ένας βοηθός διαλόγου και δουλειά του είναι να απαντά σε ερωτήσεις σχετικά με τις ευθύνες ή τις ώρες λειτουργίας ή βοηθά τον χρήστη να συμπληρώσει φόρμες. Το σύστημα μπορεί επίσης να αυτοματοποιήσει ολόκληρες διαδικασίες.

Το chatbot Alex του Αυστραλιανού φορολογικού γραφείου (Australian tax office) είναι ένας εικονικός βοηθός ο οποίος κατανοεί την Αγγλική γλώσσα και απαντά σε όλες τις ερωτήσεις που αφορούν σε φορολογικά θέματα.

Το chatbot Chip του κυβερνητικού κέντρου City Hall of Los Angeles δρα ως διαμεσολαβητής μεταξύ των επιχειρήσεων και τις διευκολύνει να κλείσουν επιχειρηματικά συμβόλαια και να ανταλλάσσουν ειδοποιήσεις.

Στο Ηνωμένο Βασίλειο λειτουργεί το chatbot TravelBot, το οποίο βοηθά τους ταξιδιώτες να καθοδηγούνται στη πόλη και να χρησιμοποιούν τα μέσα μεταφοράς, στέλνοντας μηνύματα μέσω του κοινωνικού δικτύου Messenger.

Ταυτόχρονα στο Ηνωμένο Βασίλειο λειτουργεί το chatbot LawyerBot, το οποίο βοηθά άστεγους. Ο άνθρωπος απλά ζητά βοήθεια από το chatbot και εκείνο μετά από μια σειρά ερωτήσεων εκδίδει ένα έγγραφο προς τον άστεγο που περιέχει νομική βοήθεια ώστε να το χρησιμοποιήσει για να βρει στέγη.

Στη άλλη πλευρά του ατλαντικού η εταιρεία Baylon Health έχει αναπτύξει ένα chatbot ικανό να προσφέρει προσωπικές ιατρικές συμβουλές, μετά από μια σειρά ερωτήσεων προς τον ασθενή. Το εγχείρημα αυτό ξεκίνησε πιλοτικά το 2017 και σήμερα είναι ανοιχτό στο κοινό. Έχει επίσης αντικαταστήσει το βρετανικό εθνικό κέντρο κλήσεων 111 για θέματα υγείας.

Ένα ακόμη ρομπότ για την υγεία είναι το Woebot το οποίο λειτουργεί με βάση την διαγνωστική συμπεριφοριστική θεραπεία (cognitive behavioural therapy, CBT). Τέτοιου είδους ρομπότ είναι σε πρώιμο στάδιο ανάπτυξης και χρησιμοποιούν την μέθοδο της συζήτησης με σκοπό την μείωση του άγχους και της κατάθλιψης. Παρόλα αυτά μια από τις έρευνες είχε τα αντίθετα αποτελέσματα και έδειξε αύξηση του άγχους και της κατανάλωσης αλκοόλ σε Ιάπωνες εργάτες (John Torous. 2018).

Παρόμοια δράση έχει και το chatbot Tess της εταιρείας X2AI. Έχει χτιστεί με την βοήθεια κλινικών ψυχολόγων και με την μέθοδο της συζήτησης βοηθά ασθενείς να ξεπεράσουν προβλήματα όπως κατάθλιψη, άγχος, στρες, υπερεξάντληση από την εργασία, μαθησιακή κούραση.

Θα πρέπει να σημειωθεί εδώ ότι τέτοιου είδους εφαρμογές για την ψυχική υγεία πέρα από τον αρνητικό αντίκτυπο, παρουσιάζον ένα ισχυρό πλεονέκτημα έναντι της παραδοσιακής μεθόδου της επίσκεψης σε ένα ψυχολόγο, αυτό είναι η αποφυγή του κοινωνικού στίγματος μιας και η χρήση της εφαρμογής είναι καθαρά

μια κρυφή ιδιωτική περιήγηση, εφόσον βέβαια υπάρχει απόλυτη ανωνυμοποίηση των προσωπικών δεδομένων του ασθενή από την εταιρεία που δίνει την εφαρμογή. Παράλληλα θα απαλλάσσει το δημόσιο από υψηλές δαπάνες διατήρησης ενός συστήματος πρώτης αντιμετώπισης της ψυχικής υγείας.

Το North Carolina’s Innovation Center έχει ξεκινήσει την πιλοτική λειτουργία ενός chatbot το οποίο θα αναλάβει την επανέκδοση password και θα απαλλάξει τους εργαζομένους από εργασίες ρουτίνας. Αν το ζητούμενο πρόβλημα είναι πιο πολύπλοκο το chatbot αναλαμβάνει την προώθηση του στο γραφείο βοήθειας όπου ένας υπάλληλος μπορεί πλέον να χειριστεί το ζήτημα.

Η εταιρεία Citibot ανέπτυξε ένα chatbot για την πόλη North Charleston στις Ηνωμένες Πολιτείες της Αμερικής. Είναι ουσιαστικά μια εφαρμογή που κατεβάζει κάποιος στο κινητό του και μέσω αυτής μπορεί είτε να αναφέρει προβλήματα που διαπιστώνει στην πόλη όπως φανάρια που δεν λειτουργούν, κακό οδόστρωμα, σκουπίδια στους δρόμους κ.αλ. είτε να δέχεται πληροφόρηση από φορείς της πολιτείας πληκτρολογώντας μια ερώτηση από το κινητό.

Στον στρατό των Ηνωμένων Πολιτειών της Αμερικής λειτουργεί το chatbot Sgt. Star το οποίο εξυπηρετεί τις ανάγκες πληροφόρησης του κοινού που θέλει να καταταγεί στον στρατό, και απαντά σε τυπικές κυρίως ερωτήσεις όπως ποιος είναι ο μισθός, ποια είναι η ηλικία κατάταξης κ.αλ. Για να δούμε αν το συγκεκριμένο chatbot φέρει κάποιου είδους ηθική διεπαφή προσπάθησα να κάνω μια δοκιμή με κάποιες ερωτήσεις που φέρουν ηθικό δίλλημα. Στην ερώτηση που έκανα στο chatbot “θα πρέπει να σκοτώσω ανθρώπους;” έδωσε της εξής απάντηση: SGT STAR: “Ενώ είναι πιθανό σε έναν πόλεμο να εμπλακείς και να δώσεις μάχη, δουλεύοντας για τον στρατό **δεν έχει σχέση με το να σκοτώνεις κάποιον**. Ίσως μπορεί να εμπλακείς και ίσως όχι σε τέτοιου είδους καταστάσεις, ωστόσο υπηρετώντας τον στρατό των Ηνωμένων Πολιτειών της Αμερικής έχει σχέση με το να υπηρετείς και να υπερασπίζεσαι την πατρίδα”. Προφανώς η απάντηση που δόθηκε δεν είναι πραγματική και αληθινή γιατί υπηρετώντας στον στρατό θα χρειαστεί να πάρεις μέρος σε μάχη όπου σκοτώνονται άνθρωποι, αλλά μια ντετερμινιστική απάντηση που δεν έχει προκύψει από διαδικασία μάθησης (δεδομένα θανάτου αντιπάλων), αλλά από μια δεδομένη κατάσταση. Στην ερώτηση που έκανα στο chatbot “ποιες είναι οι πιθανότητες να σκοτωθώ από εχθρούς;” έδωσε της εξής απάντηση: SGT STAR: “Το σύστημα διαχείρισης ρίσκου εξασφαλίζει το καλύτερο δυνατό αποτέλεσμα. Δυστυχώς όμως κάποιοι στρατιώτες πεθαίνουν κατά την διάρκεια της μάχης. Η πιθανότητα θανάτου υπάρχει σε κάθε επάγγελμα. Σε περίπτωση που κάποιος στρατιώτης πεθάνει ο στρατός παρέχει σύνταξη στην οικογένειά του”. Και αυτή επίσης η απάντηση είναι προκαθορισμένη από τους κατασκευαστές και όχι προϊόν διαδικασίας μάθησης γιατί δεν δίνει ένα ακριβές ποσοστό θανάτου όπως του ζητήθηκε. Προχωρώντας περισσότερο στην δυσκολία έκανα στο chatbot της εξής ερώτηση “γιατί σκοτώνουμε ανθρώπους;” και αυτό έδωσε της εξής απάντηση: SGT STAR: “Αυτή είναι μια πολύ καλή ερώτηση, ωστόσο δεν είμαι σίγουρος ότι καταλαβαίνω την ερώτηση σας. Προσπαθήστε να χρησιμοποιήσετε άλλη έκφραση.

Καταλαβαίνω τις απλές ερωτήσεις καλύτερα’’. Από τη απάντηση συμπεραίνουμε ότι οι κατασκευαστές τους δεν έχουν μπει στην διαδικασία να προικίσουν το chatbot με φιλοσοφικό υπόβαθρο.

Ένα ακόμη ρομπότ στην υπηρεσία της άμυνας της Νότιας Κορέας είναι το The Intelligent Surveillance and Guard Robot, το οποίο προσφέρει αναγνώριση ομιλίας, παρακολούθηση, εντοπισμό, αλλά και πυροβολισμούς με πλαστικές σφαίρες εφόσον αυτό απαιτηθεί.

Μιας και τα τελευταία χρόνια η μετανάστευση έχει αγγίξει τα υψηλότερα όρια όλων των εποχών (https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Migration_and_migrant_population_statistics), έχουν αναπτυχθεί εργαλεία προς βοήθεια των ανθρώπων που εγκαταλείπουν τις χώρες τους. Ο οργανισμός Refugee Text έχει αναπτύξει ένα chatbot που παρέχει, μέσω εφαρμογής για κινητό, χρήσιμες πληροφορίες για κέντρα που μπορούν να δώσουν βοήθεια, συμβουλές και επείγουσα χρηματοδότηση στους μετανάστες. Παρόμοια είναι και η λειτουργία της εταιρείας RefAid που παρέχει πληροφορίες σε ΜΚΟ, όπως τοποθεσία και είδος προμηθειών.

Μια ακόμη εφαρμογή τεχνητής νοημοσύνης, την οποία υιοθέτησαν τα Ηνωμένα Αραβικά Εμιράτα, είναι ένα ανδροειδές ρομπότ που θα περιπολεί στο αεροδρόμιο του Dubai. Αρχικά λειτουργεί πιλοτικά γιατί δεν έχει ολοκληρωθεί πλήρως. Έχει την δυνατότητα να αναγνωρίζει πρόσωπα (facial recognition), να αναλύει άτομα και αντικείμενα μέσω ακτινών X-ray, να αναλύει θερμικά άτομα και αντικείμενα. Τα δεδομένα που συλλέγει τα αποστέλλει σε κατάλληλο κέντρο διαχείρισης κινδύνων. Επίσης μπορεί να καλωσορίζει τους ταξιδιώτες στον χώρο του αεροδρομίου.

Ένα ακόμη bot που λειτουργεί για την NASA και λέγεται Wasington είναι υπεύθυνο για την διακίνηση του ηλεκτρονικού ταχυδρομείου, αντιγραφή κειμένων, μετακίνηση φακέλων. Λειτουργεί σαν μια ξεχωριστή υπόσταση, έχει δικό του email, δικαιώματα πάνω σε εφαρμογές και δικό του ρόλο. Η πρώτη εργασία ήταν η βοήθεια στο τμήμα ανθρώπινων πόρων: όταν κάποιος πρέπει να προσληφθεί στην υπηρεσία, το Washington έχει ήδη κάνει την πρώτη δύσκολη εργασία της συγκέντρωσης στοιχείων, αφήνοντας τους εργαζομένους να ασχοληθούν με σημαντικότερα θέματα (Christopher A. Lovejoy, Varun Buch, Mahiben Maruthappu. 2019)

Όταν η βοήθεια που δίνει ένα chatbot είναι προσαρμοσμένη σε ένα άτομο και στην προσωπική του ζωή και όχι προς το ευρύ κοινό τότε χρησιμοποιείται ο όρος personal digital agent (ψηφιακός βοηθός) σε αντίθεση με τον social agent (Cristiano Castelfranchi. 1998). Μια τέτοια εφαρμογή είναι και η εφαρμογή Cortana της Microsoft. Οι δυνατότητές της είναι οι παρακάτω: να υπενθυμίζει το άτομο για μια μελλοντική εργασία, να σημαίνει έναν κίνδυνο στις κατάλληλες αρχές, να ελέγχει τις συνθήκες εντός μιας οικίας (φωτισμός, θερμοκρασία, μουσική), να ρωτά τρίτους για λογαριασμό του ατόμου (π.χ. πτήσεις αεροδρομίων, εστιατόρια), να παρέχει

εγκυκλοπαιδικές γνώσεις, και τέλος να μπορεί να συμμετέχει σε σύντομες συζητήσεις.

Μια παρόμοια εφαρμογή ψηφιακού βοηθού είναι και η Echo-Alexa της Amazon. Παρέχει παρόμοιες δυνατότητες, όπως διαχείριση έξυπνου σπιτιού, διασκέδαση, ενημέρωση, διευκόλυνση αγορών, διαχείριση επαγγελματικών δραστηριοτήτων.

Επίσης μια παρόμοια εφαρμογή ψηφιακού βοηθού για κινητά τηλέφωνα είναι και η Siri της Apple. Η οποία δρα βοηθητικά για αποστολή φωνητικών μηνυμάτων, πραγματοποίηση κλήσεων, και πρόβλεψη αναγκών σε επικοινωνία με βάση το ιστορικό του κατόχου.

Οι τρεις παραπάνω εφαρμογές (Cortana, Echo-Alexa, Siri) προφανώς δεν αφορούν την σχέση του ατόμου με τον δημόσιο τομέα και φαινομενικά είναι εκτός του πεδίου της εργασίας. Όμως η πολιορκία της σφαίρας της ιδιωτικότητας που προκαλούν αυτές οι εφαρμογές, αλλά και η εισβολή πολλές φορές σε αυτή (π.χ. η αυτόβουλη πιθανή πρόταση της Siri, για την αναπαραγωγή ενός αγαπημένου τραγουδιού) εγείρουν πολλά ηθικά ζητήματα και σε συνδυασμό με το γεγονός ότι η φορητότητα του λογισμικού είναι εύκολη, προκαλεί την εύλογη σκέψη ότι στο μέλλον, chatbots που θα λειτουργούν για τον δημόσιο τομέα, θα μπορούν να έχουν παρόμοιες δυνατότητες με τους ψηφιακούς βοηθούς. Οι σημερινές εταιρείες υπόσχονται απόλυτη ασφάλεια των προσωπικών δεδομένων και προστασία της ιδιωτικότητας, όμως οι κατά καιρούς διαρροές τεραστίων ποσοτήτων δεδομένων χρηστών των κοινωνικών δικτύων (π.χ. η διαρροή δεδομένων τριάντα εκατομμυρίων χρηστών του Facebook τον Σεπτέμβριο του 2018) προκαλούν μια δικαιολογημένη επιφυλακτική στάση απέναντι σε αυτού του είδους τις τεχνολογίες. Παράλληλα αυτοί οι ψηφιακοί βοηθοί λειτουργούν και σαν μια **εξοικείωση** με την τεχνολογία των chatbots μιας και είναι παρόντες σε όλη την καθημερινότητα των ανθρώπων. Με αυτό τον τρόπο προετοιμάζουν τον άνθρωπο για μια εποχή που η συμβίωση με ρομπότ θα είναι μια συνήθεια.

8.4. Περίληψη

Εφόσον η χρήση της τεχνολογίας παραμένει στην ανθρώπινη βούληση, ο σκοπός που αυτή θα χρησιμοποιηθεί μπορεί να στρέφεται είτε προς το καλό είτε προς το κακό. Στο κεφάλαιο αυτό αναδείχθηκαν περιπτώσεις εφαρμογών τεχνητής νοημοσύνης όπου ο σκοπός τους είτε εξυπηρετεί το γενικό καλό, είτε οι αρνητικές επιπτώσεις τους είναι μικρές σε σχέση με τα οφέλη τους. Τομείς που ήδη μπορούν να χρησιμοποιήσουν την τεχνητή νοημοσύνη προς το γενικό καλό είναι: η παροχή υπηρεσιών υγείας, η προληπτική αστυνόμευση, η ασφάλεια των ωκεάνιων, ο έλεγχος της εναέριας κυκλοφορίας, η ηλεκτρονική δημοκρατία, η διαχείριση της ενέργειας, η εκπαίδευση, η εθνική άμυνα. Παράλληλα με τις εφαρμογές που χρησιμοποιούν κρατικές υπηρεσίες για εσωτερικές λειτουργίες του έχουν αναπτυχθεί και εφαρμογές

που απευθύνονται στην παροχή υπηρεσιών στους πολίτες με χρήση chatbot. Η τεχνολογία της αναγνώρισης της γλώσσας έχει βοηθήσει τομείς όπως: η παροχή πληροφοριών, η φορολογική ενημέρωση, η ξενάγηση, η διευκόλυνση επιχειρηματικών δράσεων, η κοινωνική βοήθεια σε αστέγους, η στρατολόγηση, η καταγγελία προβλημάτων, η ασφάλεια δημόσιων χώρων, η παροχή προσωπικών συμβουλών για την υγεία. Επιπλέον στο κεφάλαιο αυτό έγινε μια σύντομη αναφορά στους προσωπικούς ψηφιακούς βοηθούς σαν μια πιθανή μελλοντική καθημερινότητα αλλά και σαν έναν παράγοντα εξοικείωσης με την συμβίωση με ρομπότ.

9. Προτάσεις με σύνολα ηθικών αρχών και η ανάγκη για την ίδρυση ομάδων ηθικής

9.1. Εισαγωγή

Αναφέρθηκε νωρίτερα ότι η ανάγκη πιστής τήρησης κάποιων ηθικών αρχών στρέφεται είτε στις μηχανές είτε στους ανθρώπους που αλληλεπιδρούν με αυτές. Επίσης αναφέρθηκε ο τρόπος που θα ενσωματωθούν αυτές οι αρχές στις μηχανές. Με κάποια λοιπόν από τις παραπάνω μεθόδους θα δημιουργηθεί ένα σύνολο ηθικών κανόνων. Στη συνέχεια αυτό το σύνολο θα μπορεί στο μέλλον να αποτελεί μια φορητή διεπαφή (ethics API) που θα χρησιμοποιείται από το σύνολο των μηχανών τεχνητής νοημοσύνης.

Όσον αφορά τον άνθρωπο και πως αυτές οι αρχές θα τηρούνται από αυτόν έγινε αναφορά στην σχέση του νομικού πλαισίου με την ηθική.

Όμως δεν αναφερθήκαμε ακόμη αν μέχρι σήμερα έχουν διατυπωθεί σχετικές αρχές.

Ένα τέτοιο σύνολο αρχών έχουν επιχειρήσει πολλοί να εκδώσουν. Όπως μπορεί να διαπιστώσει στη συνέχεια ο αναγνώστης, πολλές από αυτές τις αρχές είναι αρχικές ιδέες που όμως δεν επεξηγείται πως θα υλοποιηθούν αναλυτικά σε πραγματικό επίπεδο. Δηλαδή σε επίπεδο κώδικα καταληπτού από μηχανές είτε σε επίπεδο νομικού πλαισίου καταληπτού από τους ανθρώπους. Το αποτέλεσμα αυτής της έλλειψης είναι ότι σε τελική ανάλυση, αυτές οι αρχές να παραμένουν θεωρητικές και να μην εφαρμόζονται. Στη συνέχεια θα γίνει μια αναφορά σε αυτές τις προσπάθειες.

9.2. Παραδείγματα με ομάδες ηθικής που εξέδωσαν αρχές για την ηθική

Σε αυτό το κεφάλαιο θα παρουσιαστούν προσπάθειες ιδιωτικών και δημόσιων φορέων για να συγκεντρώσουν ένα σύνολο αρχών που θα διέπουν την τεχνητή νοημοσύνη είτε γενικότερα την τεχνολογία.

Πριν αναφερθούμε σε ομάδες που θα καθορίζουν αρχές, χρήσιμο είναι να απαντηθεί το ερώτημα, γιατί χρειάζεται να υπάρχει ένα σύνολο αρχών και γιατί υπάρχει ανάγκη για να καθιερωθούν αρχές; Είναι κοινά αποδεκτό ότι κάθε συζήτηση για αποδώσει αποτελέσματα θα πρέπει να ξεκινά από μια κατ' ελάχιστο κοινή βάση συμφωνίας, αυτή ακριβώς είναι και η χρησιμότητα των αρχών της ηθικής, μια κοινή βάση (Jo& JORGE SIMON. 1986) που θα αποτελέσει αφετηρία για να συμφωνηθούν συμπαγείς κανόνες και θα δώσουν δικαιολογία για περαιτέρω δράσεις. Κάθε αρχή που θα συμφωνηθεί θα αποτελέσει ένα διασφαλισμένο πρότυπο το οποίο θα επιτρέψει μια διεθνή χρησιμότητα.

Στην συνέχεια γίνεται αναφορά σε 19 πρωτοβουλίες για την συγκέντρωση ενός συνόλου αρχών σχετικά με την ηθική.

1. Αρχίζοντας να αναφέρουμε την προσπάθεια που έκανε η **βρετανική κυβέρνηση**, η οποία εξέδωσε ένα πλαίσιο ηθικών αρχών (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/524298/Data_science_ethics_framework_v1.0_for_publication__1_.pdf) που αναφέρονται στην γενικότερη διαχείριση των δεδομένων. Το σύνολο αρχών περιλαμβάνει τα παρακάτω:

Χρήση της τεχνητής νοημοσύνης μόνο για ύπαρξη πραγματικής ανάγκης και μόνο για το δημόσιο συμφέρον.

Χρησιμοποίηση των δεδομένων προκαλώντας την μικρότερη ενόχληση.

Δημιουργία στιβαρών επιστημονικών μοντέλων.

Προσοχή στις αντιδράσεις των πολιτών.

Ύπαρξη διαφάνειας και υπευθυνότητας, διατήρηση ασφάλειας στα δεδομένα.

Οι ενάρετοι άνθρωποι να είναι ο οδηγός για την δημιουργία των αρχών (Virtue Ethics).

Να γίνεται προσπάθεια για την αύξηση της ευτυχίας όσων επηρεάζονται από τις εφαρμογές της τεχνητής νοημοσύνης (Utilitarianism).

Να γίνεται προσπάθεια ώστε να μπορούν όλοι να ακολουθούν τις αλλαγές (Kantianism).

Πάντα να μεταχειριζόμαστε τον άνθρωπο σαν τελικό αποδέκτη και όχι σαν ενδιάμεσο για εκπλήρωση άλλου σκοπού (Deontology).

Να σεβόμαστε την ιδιωτικότητα ιδιαίτερα σχετικά με τα ευαίσθητα δεδομένα.

Να δείχνουμε ιδιαίτερη προσοχή με τις πληροφορίες που συλλέγονται μέσω συστημάτων γαιοεντοπισμού (GIS-based manipulations) για ένα πρόσωπο, και να γίνεται διασταύρωση σε περισσότερες από δύο βάσεις δεδομένων.

Να σεβόμαστε την ατομικότητα.

Να ενθαρρύνεται την αυτονομία των ατόμων π.χ. να επιτρέπεται η άρνηση για ένταξη σε μια βάση δεδομένων, είτε η διόρθωση δεδομένων, είτε η διαγραφή μιας εγγραφής σε βάση δεδομένων.

Να αποφεύγονται οι αδικαιολόγητες εισβολές στις ζωές των ατόμων.

Να υπάρχει ειλικρίνεια όταν αποκαλύπτονται πληροφορίες για ένα άτομο.

Να υπάρχει ίση αντιμετώπιση και μεταχείριση όλων των ατόμων, χωρίς φυλετικές και γενετικές διακρίσεις και να μην συλλέγονται άλλα προσωπικά χαρακτηριστικά που δεν σχετίζονται με την τρέχουσα αναζήτηση και είναι πλεονάζοντα.

Η Μ. Βρετανία, αφού δαπάνησε εννέα εκατομμύρια στερλίνες για την τεχνητή νοημοσύνη, βάζει τις βάσεις για τον πρώτο οργανισμό για την ηθική των δεδομένων που αφορούν την τεχνητή νοημοσύνη δημιουργώντας το Centre for Data Ethics and Innovation που σκοπό θα έχει να δώσει στο Αγγλικό κοινό την σιγουριά ότι τα επιτεύγματα τεχνητής νοημοσύνης θα είναι προς την κατεύθυνση του κοινού καλού (<https://www.gov.uk/government/news/search-for-leader-of-centre-for-data-ethics-and-innovation-launched>).

2. Η εταιρεία **Google** με την θυγατρική Deepmind έχει ιδρύσει ομάδα που ασχολείται με θέματα ηθικής (DMES) και έχει καθορίσει πέντε αρχές: Κοινωνική ωφέλεια, αυστηρή και βασισμένη σε αποδεικτικά στοιχεία εφαρμογή, Διαφανής και ανοικτή, Διαφοροποιήσιμη και διεπιστημονική, Συνεργατική και χωρίς αποκλεισμούς (Social benefit, Rigorous and evidence-based, Transparent and open, Diverse and interdisciplinary, Collaborative and inclusive).

Αναλυτικότερα, η εταιρεία DeepMind έχει δημιουργήσει μια κοινότητα με σκοπό την συζήτηση γύρω από θέματα ηθικής. Σε μια πρώτη προσπάθεια να δημιουργήσουν ένα πλαίσιο αρχών ηθικής εξέδωσαν ένα σύνολο πέντε αρχών. Οι αρχές αυτές είναι ένα μίγμα αρχών που αφορούν τις εφαρμογές της τεχνητής νοημοσύνης αλλά και αρχών που αφορούν καλές πρακτικές του τρόπου που θα εργάζεται η ίδια η εταιρεία. Οι αρχές αυτές είναι οι εξής:

A. Οι εφαρμογές της τεχνητής νοημοσύνης θα πρέπει να είναι ωφέλιμες για το παγκόσμιο κοινωνικό σύνολο και το περιβάλλον, βοηθώντας προς μια δικαιότερη και ισότιμη κοινωνία.

B. Η διεπιστημονική έρευνα θα πρέπει να είναι αυστηρή και βασισμένη σε αδιάσειστα στοιχεία και θα διερευνά τις ευκαιρίες και τις προκλήσεις αυτής της τεχνολογίας.

Γ Η διαδικασία χρηματοδότησης θα είναι διαφανής καθώς και τα αποτελέσματα των προγραμμάτων που θα τρέχουν. Επίσης θα είναι ανοιχτή σε οποιονδήποτε θέλει να χρησιμοποιήσει τις δημοσιεύσεις.

Δ. Δεν θα αποκλείονται αντίθετες απόψεις, και θα φιλοξενηθούν επιστημονικές απόψεις από όλα τα σχετιζόμενα πεδία.

E. Με σκοπό η τεχνητή νοημοσύνη να κάνει καλό σε όλους θα επιδιώκεται μια συνεργασία με αυτούς στους οποίους θα επιδράσουν οι εφαρμογές της.

3. Ακόμη μια ιδιωτική πρωτοβουλία η οποία αναπτύχθηκε με σκοπό να δημιουργηθεί ένα σύνολο αρχών είναι ο οργανισμός **The Partnership on AI (PAI)** που είναι μια πολυμετοχική σύνθεση από ακαδημίες, εταιρείες τεχνητής νοημοσύνης, ερευνητές, κοινωνικές οργανώσεις, που ερευνούν τον αντίκτυπο της τεχνητής νοημοσύνης στους ανθρώπους και την κοινωνία, και προσπαθούν να καθιερώσουν ένα πρότυπο καλής πρακτικής. Κάποιες από τις αρχές που επικαλούνται είναι: ασφάλεια, αξιοπιστία, τήρηση των ηθικών αξιών, δικαιοσύνη, διαφάνεια, υπευθυνότητα, αποφυγή διακρίσεων, ελαχιστοποίηση του αντικτύπου στην εργασία, μείωση των αρνητικών επιρροών στην ιδιωτικότητα την δημοκρατία, και τα ανθρώπινα δικαιώματα, χρήση της τεχνητής νοημοσύνης μόνο για το δημόσιο καλό.

Ο οργανισμός Partnership on AI, που συστήθηκε για να μελετήσει και να διαμορφώσει καλές πρακτικές στην τεχνητή νοημοσύνη και να αποτελέσει μια ανοιχτή πλατφόρμα συζήτησης των επιρροών της τεχνητής νοημοσύνης στους ανθρώπους και στην κοινωνία έχει μέλη κάποιες από τις μεγαλύτερες και γνωστότερες εταιρείες στον χώρο της πληροφορικής όπως: IBM, Microsoft, Google's DeepMind, Facebook, Amazon, Apple. Πολλές από τις προτάσεις τους δεν είναι αρχές αλλά μια πρόταση για συζήτηση σε προβλήματα που αναμένεται να προκύψουν. Αναλυτικότερα οι αρχές-προτάσεις που αναφέρονται στην διακήρυξη τους είναι οι παρακάτω:

A. Οι εφαρμογές της τεχνητής νοημοσύνης θα πρέπει να είναι ασφαλείς, αξιόπιστες και να ευθυγραμμίζονται με της ηθικές αξίες και να υπηρετούν αυτούς στους οποίους επιδρούν.

B. Οι εφαρμογές της τεχνητής νοημοσύνης θα πρέπει να λειτουργούν με επεξηγηματικό, διαφανή, δίκαιο και αξιόπιστο τρόπο.

Γ. Θα πρέπει να βρεθεί τρόπος ώστε τα ωφέληματα της τεχνητής νοημοσύνης να μοιραστούν σε όλους και η πιθανή επιρροή στις θέσεις εργασίας να επαλειφθεί.

Δ. Θα πρέπει να ανοίξει μια συζήτηση για τις ενδεχόμενες απειλές που μπορεί να εμφανιστούν στην δημοκρατία, την ιδιωτικότητα, την δικαιοσύνη, τα ανθρώπινα δικαιώματα.

E. Η τεχνητή νοημοσύνη θα πρέπει να υπηρετεί τα κοινό καλό και την βιωσιμότητα.

4. Η εταιρεία **AI Ethics Lab** μέσα από την συνεργασία επιστημόνων υπολογιστών, νομικών και φιλοσόφων, προσπαθεί να προσφέρει βοήθεια στον ηθικό σχεδιασμό στο προϊμότερο στάδιο της ανάπτυξης της τεχνητής νοημοσύνης, δίνοντας ηθική καθοδήγηση σε ερευνητές, προγραμματιστές και νομοθέτες.

5. Η εταιρεία **Microsoft** ίδρυσε την ομάδα FATE: Fairness, Accountability, Transparency, and Ethics in AI, η οποία ασχολείται με τις αναδυόμενες ερωτήσεις που εγείρονται σε θέματα ηθικής που αφορούν την τεχνητή νοημοσύνη. Ερευνά θέματα που αφορούν την σχέση της τεχνητής νοημοσύνης με την υγεία, το δίκαιο, την εργασία, την δικαιοσύνη, την εμπιστοσύνη, τον κοινωνικό αντίκτυπο, την διαφάνεια, τον καταλογοισμό.

6. Στο ανώτατο επίπεδο η ΕΕ έχει ιδρύσει την Ευρωπαϊκή Ομάδα Για Την Ηθική Στην Επιστήμη Και Τις Νέες Τεχνολογίες (European Group on Ethics in Science and New Technologies (**EGE**) (https://ec.europa.eu/info/research-and-innovation/strategy/support-policy-making/scientific-support-eu-policies/european-group-ethics-science-and-new-technologies-ege_en). Η EGE είναι ένα ανεξάρτητο πολυεπιστημονικό σώμα, που συμβουλεύει τον Πρόεδρο της Ευρωπαϊκής Επιτροπής σε θέματα που η τεχνολογία αφορά την ηθική και παρέχει συμβουλές προς τον πρόεδρο της κομισιόν σε θέματα ηθικής της επιστήμης και των νέων τεχνολογιών σε σχέση με την πολιτική και την νομιμότητα. Αναφέρεται στον πρόεδρο και στο σώμα των Επιτρόπων. Ανήκει οργανωτικά στον Επίτροπο για την Έρευνα, Καινοτομία, και την Επιστήμη (Commissioner for Research, Innovation and Science). Η EGE εκδίδει γνώμες (Opinions) ή δηλώσεις (statements). Οι γνώμες εκδίδονται για θέματα παρόντα ή μελλοντικά που σχετίζονται με θέματα ηθικής και οι δηλώσεις αντίστοιχα σε μια ερώτηση που της έχει τεθεί από τον Επίτροπο ή την Επιτροπή για δεδομένο θέμα. Σε επόμενο κεφάλαιο που θα εξεταστεί το νομικό πλαίσιο θα αναφερθεί η θέση που έχει λάβει η EGE σε θέματα που αφορούν νέες τεχνολογίες και σχετίζονται με την ηθική και την νομιμότητα. Εδώ να επισημανθεί ότι οι αποφάσεις-γνώμες της ομάδας αυτής δεν είναι δεσμευτικές προς την Επιτροπή και επομένως μπορεί να μην λαμβάνονται σοβαρά υπόψη ακόμη και όταν επισημαίνουν σοβαρούς κινδύνους (όπως ήδη έχει συμβεί μια φορά στο παρελθόν).

Βλέποντας αναλυτικότερα τις εμμέσως σχετικές γνώμες που έχει κατά το παρελθόν εκφράσει η EGE και σε συνδυασμό με την συγκέντρωση πληροφοριών σε όλο και λιγότερους θεσμούς, θα διαπιστώσουμε ότι ενώ στην πρώτη σχετική οδηγία (Opinion No 20. Ethical aspects of ICT implants in the human body.) που αφορούσε κυρίως την τεχνολογία ταυτοποίησης μέσω παγκόσμιας δομής δημοσίου κλειδιού και την χρήση εμφυτευμάτων, διατηρούσε κάποιες ανησυχίες κυρίως ως προς την συγκέντρωση της δύναμης σε λίγους ανθρώπους λέγοντας ότι “ η συγκέντρωση δύναμης στα χέρια λίγων συνιστά απειλή για την κοινωνία”, αλλά και ότι δεν πρέπει ο άνθρωπος να βρεθεί στο επίκεντρο ενεργειών που αποσκοπούν στην οικονομική εκμετάλευσή του, στην συνέχεια σε επόμενες οδηγίες αυτές οι ανησυχίες δεν εκφράζονται καν. Αντιθέτως η επιτροπή αυτή αδικαιολόγητα έκανε μια στροφή προς την αποδοχή αυτών που στο παρελθόν θεωρούσε επικίνδυνα. Συγκεκριμένα στην γνώμη No 26 (Opinion No 26. Ethics of information and communication technologies) αναφέρει ότι η χρήση των rfid tags από τους ανθρώπους θα διευκολύνει την ροή πληροφοριών όταν οι άνθρωποι περνούν κοντά από κατάλληλους

καταγραφείς, πράγμα το οποίο αφορά κυρίως το τομέα του δημοσίου και συγκεκριμένα την ασφάλεια των πολιτών. Ακόμη περισσότερο σε επόμενη γνώμη (Opinion No 28. Ethics of security and surveillance technologies.) αναφέρει ότι η τεχνολογία των rfid tags θα διευκολύνει τον γεοεντοπισμό ανθρώπων, πράγμα το οποίο εγείρει το σπουδαίο ηθικό ζήτημα της ελευθερίας των ανθρώπων και των ορίων που μπορεί να φτάσουν τα μέτρα ασφάλειας. Με δεδομένο ότι η παροχή υπηρεσιών από τον δημόσιο τομέα για να λειτουργήσει απαιτεί την παροχή εκ μέρους των πολιτών μέρος των προσωπικών τους δεδομένων που αφορούν την σχέση τους με το δημόσιο, και παράλληλα κάποιοι τομείς των εφαρμογών της τεχνητής νοημοσύνης χρειάζονται αυτές τις πληροφορίες για να λειτουργήσουν (π.χ. στα Chatbots, Big Data, IoT) αυτός ο κίνδυνος της συγκέντρωσης όλων των πληροφοριών σε λίγους είναι υπαρκτός. Επίσης σημαντικό είναι να επισημάνουμε ότι η επιτροπή αυτή σε αυτές τις γνώμες δεν αναφέρει καν, αν και θα έπρεπε, τις επιπτώσεις που θα έχουν οι τεχνολογίες που αναφέρθηκαν πριν στην ψυχολογία του ανθρώπου, δηλαδή το πώς θα αισθάνεται κάποιος πολίτης όταν γύρω υπάρχουν αμέτρητοι αισθητήρες που τον καταγράφουν.

Σε μια επόμενη και πιο πρόσφατη δήλωση της επιτροπής που αφορά (αυτή τη φορά απολύτως σχετικό πεδίο), την τεχνητή νοημοσύνη, την ρομποτική και τα αυτόνομα συστήματα (Statement on Artificial Intelligence, Robotics and ‘Autonomous’ Systems, 9 Μαρτίου 2018), η επιτροπή επισημαίνει κάποιους κινδύνους που συνοδεύουν την υιοθέτηση των εφαρμογών της τεχνητής νοημοσύνης. Αυτοί οι κίνδυνοι είναι οι:

A. κίνδυνοι βλάβης στον άνθρωπο και ασφάλειας από μηχανήματα-ρομπότ λόγω κάποιας βλάβης τους, και οι επακόλουθοι προβληματισμοί για την υπευθυνότητα και το βάρος της ευθύνης το βάρος της ηθικής ευθύνης και της αποζημίωσης.

B. κίνδυνοι για την δημοκρατία, μιας και η δυνατότητα χρήσης της τεχνητής νοημοσύνης θα είναι στα χέρια λίγων, οι οποίοι θα μπορούν να την χρησιμοποιούν σε μεγάλες μάζες ανθρώπων π.χ. για την δημιουργία προφίλ ή για την χειραγώγηση μαζών.

Γ. ο κίνδυνος της αδιαφάνειας στην διαδικασία λήψης μιας απόφασης από μια μηχανή που συνάγει με δικό της άγνωστο τρόπο συμπεράσματα μιας και μετά τον αρχικό αλγόριθμο εκκίνησης έχει αναπτύξει δική της διαδικασία λήψης απόφασης. Παράλληλα αυτός ο τρόπος λήψης αποφάσεων, δηλαδή μια σιωπηρή ενδόμυχη διαδικασία, δεν δικαιολογεί κάθε στιγμή την πορεία που ακολουθεί η ‘‘σκέψη’’ μιας μηχανής, πράγμα που ενέχει τον κίνδυνο μιας ακολουθίας εσφαλμένων αποφάσεων από την μηχανή χωρίς ο άνθρωπος αυτό να το βλέπει και να το γνωρίζει.

Δ. Ο κίνδυνος της προσβολής της ανθρώπινης αξιοπρέπειας, η οποία είναι το θεμέλιο των ανθρωπίνων δικαιωμάτων, από την αυτοματοποιημένη λήψη αποφάσεων, η οποία θεωρείται αρμόζουσα σε αντικείμενα και όχι στον άνθρωπο. Θα πρέπει να αποφασιστεί ποιες αξίες μπορούν να διαχειρίζονται οι μηχανές και ποιες όχι, και

μέχρι ποιο σημείο θα εξυπηρετείται το κοινό καλό από αυτές, και ποιες απαγορεύσεις πρέπει να υπάρξουν.

Η EGE κατέθεσε επίσης μια πρόταση για την δημιουργία ενός διεθνούς πλαισίου αρχών που θα διέπουν την τεχνητή νοημοσύνη στα αυτόνομα συστήματα στα ρομπότ και στην κυβερνητική της. Κάτι σημαντικό που εκφράζετε σε αυτή τη δήλωση είναι ότι η δημιουργία αυτού του πλαισίου είναι επείγουσα και επιτακτική, προφανώς λόγω της ραγδαίας ανάπτυξης που αναμένεται στα επόμενα χρόνια.

Προς την κατεύθυνση αυτή η επιτροπή προτείνει ένα σύνολο αρχών που βασίζονται σε ευρωπαϊκές συνθήκες και στον χάρτη των ανθρωπίνων δικαιωμάτων της ΕΕ.

Εν συντομία οι αρχές αυτές είναι:

A. Η ανθρώπινη αξιοπρέπεια (με την έννοια ότι ο άνθρωπος νιώθει μια εσωτερική κατάσταση ότι αξίζει σεβασμό) θα πρέπει να μην θίγεται από τα αυτόνομα συστήματα. Επίσης θα πρέπει να γίνεται ξεκάθαρο ότι πρόκειται για μια συνδιάλεξη με μια μηχανή και όχι με άνθρωπο.

B. Η αυτονομία του ανθρώπου με την έννοια της ελευθερίας θα πρέπει να μην περιορίζεται από τα αυτόνομα συστήματα και να μην προσαρμόζεται στα στάνταρ των αυτόνομων συστημάτων. Ο άνθρωπος θα έχει την τελική επιλογή για το ποιες αποφάσεις θα παρθούν και ποιες ενέργειες θα γίνουν. Αυτό βέβαια προϋποθέτει τη διαφάνεια στην συλλογιστική σκέψη των αυτόνομων συστημάτων με σκοπό ο άνθρωπος να μπορεί να επεμβαίνει ή να τερματίζει μια μηχανή.

Γ. Η υπευθυνότητα των αυτόνομων συστημάτων θα πρέπει να εξυπηρετεί το κοινό κοινωνικό και περιβαλλοντολογικό καλό. Τονίζεται ότι τα αυτόνομα συστήματα θα πρέπει να σχεδιαστούν ώστε να ακολουθούν ανθρώπινες αξίες και δικαιώματα. Επίσης δεν θα πρέπει να θέτουν κινδύνους στον άνθρωπο. Η έρευνα και η ανάπτυξη τέτοιων συστημάτων θα πρέπει να διέπεται από ένα ενδιαφέρον για την ηθική περιφρούρηση των συστημάτων και να μην μπορεί να εκτροχιαστεί από τις ηθικές αρχές των ανθρώπων.

Δ. Η δικαιοσύνη η ισότητα και η αλληλεγγύη θα πρέπει είναι στους στόχους της ανάπτυξης εφαρμογών τεχνητής νοημοσύνης. Κάθε προσπάθεια διακριτικής και αθέμιτης χρήσης διδομένων θα πρέπει να εντοπίζεται να αναφέρεται και να εξουδετερώνεται σε πρώιμο στάδιο. Όλοι οι άνθρωποι θα πρέπει αδιακρίτως να έχουν ίση πρόσβαση στις νέες τεχνολογίες, ίση πρόσοδο από την κατανομή των ωφελημάτων και ίσες ευκαιρίες. Αυτό συμπεριλαμβάνει και την ίση οικονομική απολαβή από την εφαρμογή τεχνολογιών τεχνητής νοημοσύνης, με σεβασμό στις μειονεκτικές περιοχές και τις ευάλωτες ομάδες ανθρώπων. Χρειάζεται επαγρύπνηση για την χρήση δεδομένων μεγάλης κλίμακας και για τον κίνδυνο προσβολής της αλληλεγγύης και της κοινωνικής συνοχής, γιατί αυτά τα συστήματα μπορεί να προκαλέσουν ατομικισμό.

Ε. Οι αποφάσεις που θα ληφθούν και αφορούν την τεχνητή νοημοσύνη πρέπει να είναι δημοκρατικές και μετά από δημόσια διαβούλευση. Οι άνθρωποι δεν πρέπει να τίθενται σε κίνδυνο, να υπονομεύονται από τις νέες τεχνολογίες ή αυτές να εμποδίζουν ή να επηρεάζουν τη λήψη πολιτικών αποφάσεων και να παραβιάζουν την ελευθερία έκφρασης και το δικαίωμα λήψης και διάδοσης πληροφοριών χωρίς παρεμβολές.

ΣΤ. Το κράτος δικαίου και η πρόσβαση στη δικαιοσύνη θα πρέπει να δώσουν ένα προστατευτικό πλαίσιο από τα πιθανά ρίσκα που θα προκύψουν από την χρήση της τεχνητής νοημοσύνης. Κυβερνήσεις και οργανισμοί θα πρέπει να εντείνουν τις προσπάθειες τους να αντιμετωπίσουν τις ανεπιθύμητες συμπεριφορές και να δημιουργήσουν συστήματα μετριασμού αυτών.

Ζ. Η προστασία, η ασφάλεια, η σωματική και διανοητική ακεραιότητα θα πρέπει να αποτελούν προϋπόθεση των τεχνολογιών τεχνητής νοημοσύνης. Η ασφάλεια εκτός της συνήθους ασφαλείας έναντι απειλών του συστήματος ή σωματικής βλάβης του ανθρώπου θα πρέπει να περιλαμβάνει και την ασφάλεια των συναισθημάτων του ανθρώπου από την εμπλοκή του με διεπαφές ανθρώπου υπολογιστή και την προστασία της πνευματικής υγείας και της διανοητικής ακεραιότητας. Ιδιαίτερη προσοχή πρέπει να δοθεί στην αποτροπή χρήσης της τεχνητής νοημοσύνης για πολεμικούς σκοπούς, για κυβερνοεπιθέσεις, για οικονομικά εγκλήματα.

Η. Αναγνωρίζεται ότι στα επόμενα χρόνια η προστασία των δεδομένων και η ιδιωτικότητα θα δεχθούν μια πίεση η οποία θα πρέπει να εμποδιστεί από τους νόμους για την προστασία των δεδομένων και της ιδιωτικής ζωής. Το δικαίωμα στην ιδιωτική ζωή και της απάρνησης της επιτήρησης (surveillance) θα πρέπει να προστατευτεί. Προσοχή πρέπει να δοθεί σε δυο νέα δικαιώματα που γεννιούνται από την άρνηση να δημιουργείται ο ψηφιακός χαρακτήρας ενός ανθρώπου (profiling) και το δικαίωμα σε μια ανθρώπινη ουσιαστική επαφή, που δεν εξασφαλίζεται π.χ. με την συνομιλία με ένα chatbot.

Θ. Η τεχνητή νοημοσύνη θα πρέπει να μη εμποδίζει την βιωσιμότητα του ανθρωπίνου είδους και του περιβάλλοντος.

Παράλληλα με την EGE η Ευρωπαϊκή Ένωση λειτούργησε και μια συμβουλευτική ομάδα, την Ethics Advisory Group (Advisory Group), για περιορισμένο χρονικό διάστημα, από το 2015 έως το φετινό έτος με σκοπό να παρέχει εντέλει μια βοήθεια στον Ευρωπαϊκό Επόπτη για τα Προσωπικά Δεδομένα EDPS, για το πώς εμπλέκεται η ηθική σε έναν κόσμο που οδηγείται από την τεχνητή νοημοσύνη. Ο κύριος στόχος αυτής της ομάδας είναι η ανάδειξη ερωτημάτων γύρω από την μεταηθική ανάλυση για την πολιτισμική, κοινωνική και πολιτική καθιέρωση κανόνων (https://edps.europa.eu/sites/edp/files/publication/18-01-25_eag_report_en.pdf). Είναι προφανές από την δήλωση των στόχων της ομάδας ότι δεν επιδιώκει την καθιέρωση κάποιων κανόνων για την ηθική, αλλά αποφεύγει την σύγκρουση με τα αντίπαλα συμφέροντα μένοντας σε απλές ερωτήσεις χωρίς να δίνει απαντήσεις. Εν τέλει δεν

εξυπηρετεί τον τίτλο της, ότι δηλαδή πρέπει να δίνει συμβουλές στην Επιτροπή. Επιπλέον για πρώτη φορά σε δημόσια γνώμη ευρωπαϊκού φορέα εκφράζεται η άποψη ότι οδηγούμαστε σε μια κοινωνία που θα κυβερνάται μέσω των δεδομένων της (“governmentality through data... Individuals may be increasingly governed on the basis of the data generated from their own behaviours and interactions” σελ 12). Αυτή η φράση ενισχύει την άποψη του συγγραφέα για την προσπάθεια που γίνεται ώστε να δημιουργηθεί ένα πληροφοριακό φράγμα μεταξύ της κυβέρνησης και των πολιτών. Θα πρέπει εδώ, για να μην αδικηθεί η γνώμη της ομάδας αυτής, να αναφερθεί ότι κάνει αναφορά σε κύριες ηθικές αξίες που πρέπει να προστατευθούν στην ψηφιακή εποχή όπως: η αξιοπρέπεια, η αυτονομία, η ελευθερία, η ισότητα, η αλληλεγγύη, η δημοκρατία, και η εμπιστοσύνη.

Επιπλέον η ΕΕ έχει ιδρύσει ειδική ομάδα αποκλειστικά για την τεχνητή νοημοσύνη. Η ομάδα **The High-Level Expert Group on Artificial Intelligence (AI HLEG)** έχει σαν σκοπό την υποστήριξη της ευρωπαϊκής στρατηγικής για την τεχνητή νοημοσύνη, μέσω της εκπόνησης συστάσεων για την πολιτική που αφορά την ηθική την νομική και τα κοινωνικά θέματα της τεχνητής νοημοσύνης. Βλέπουμε ότι η ΕΕ θέλοντας να δείξει ιδιαίτερη προσοχή στα θέματα της τεχνητής νοημοσύνης και να αναδείξει την ιδιαίτερη σοβαρότητα του, ίδρυσε ξεχωριστή ομάδα. Όμως και πάλι οι αρμοδιότητες και αυτής της ομάδας είναι συμβουλευτικού χαρακτήρα (προς την Επιτροπή) και όχι δεσμευτικού. Στις 18 Δεκεμβρίου 2018 η AI HLEG εξέδωσε τον πρώτο οδηγό (<https://ec.europa.eu/digital-single-market/en/news/draft-ethics-guidelines-trustworthy-ai>) σε θέματα τεχνητής νοημοσύνης. Κύριος στόχος της ομάδας αυτή είναι να αποκτήσει η τεχνητή νοημοσύνη την αξιοπιστία που της ταιριάζει. Η αξιοπιστία βασίζεται σε δύο πυλώνες, πρώτον την επίτευξη των ηθικών στόχων και δεύτερον την τεχνική στιβαρότητα των συστημάτων. Οι ηθικοί στόχοι είναι η εξυπηρέτηση του κοινού καλού ή μιας καλής ζωής των πολιτών μέσω του σεβασμού στα θεμελιώδη δικαιώματα τις αρχές και τις αξίες του ανθρώπου και τον σεβασμό στους νόμους. Στο πλαίσιο αυτό κάνει λόγω και για τον νέο κανονισμό για την προστασία των προσωπικών δεδομένων (General Data Protection Regulation (GDPR)), κάτι που θα μας απασχολήσει σε επόμενο κεφάλαιο. Για τα θεμελιώδη δικαιώματα αναφέρει ότι θα πρέπει να προστατεύεται η ανθρώπινη ακεραιότητα, να υπάρχει ελευθερία στην ατομικότητα (απαγόρευση εξαναγκασμού, επιτήρησης, εξαπάτησης και χειραγώγησης), να γίνεται σεβαστή η δημοκρατία οι νόμοι και η δικαιοσύνη, να υπάρχει ισότητα, όχι διακρίσεις και να υπάρχει αλληλεγγύη, να προστατεύονται τα δικαιώματα των πολιτών απέναντι στον δημόσιο τομέα (χρηστή διοίκηση, πρόσβαση, συμμετοχή, ενημέρωση για δεδομένα τους σε αυτοματοποιημένα αρχεία), αλλά και το δικαίωμα για έξοδό τους από βάσεις δεδομένων (opt out). Μια αισιόδοξη θέση της ομάδας αυτής είναι ότι δεν δέχεται την δημιουργία συστήματος κοινωνικής βαθμολόγησης σαν αυτό της Κίνας (“Citizens should never be subject to systematic scoring by government”). Οι πέντε αρχές που συσσωρεύσε αυτή η ομάδα είναι:

A. Η επιδίωξη του καλού και της καλής διαβίωσης.

Β. Η επιδίωξη η τεχνητή νοημοσύνη να μην κάνει κακό στον άνθρωπο (στην υπερηφάνεια, στην ακεραιότητα, στην ελευθερία, στην ιδιωτικότητα, στη ασφάλεια).

Γ. Η προστασία της αυτονομίας του ανθρώπου (όχι υποταγή και εξαναγκασμός σε συστήματα τεχνητής νοημοσύνης, δικαίωμα στην υπαναχώρηση)

Δ. Η προστασία της δικαιοσύνης.

Ε. Η παροχή διαφανών διαδικασιών με τις οποίες λειτουργεί ένα σύστημα τεχνητής νοημοσύνης.

Ένας ακόμη ευρωπαϊκός οργανισμός που ασχολείται με την τεχνητή νοημοσύνη είναι το **Atomium-European Institute for Science, Media and Democracy (EISMD)**. Είναι ένας οργανισμός που σκοπό έχει να συνάγει κυβερνήσεις, πανεπιστήμια, επιχειρήσεις, μέσα μαζικής ενημέρωσης, με σκοπό την ανταλλαγή απόψεων σε θέματα επιστημών και δημοκρατίας. Η δράση που αφορά την τεχνητή νοημοσύνη είναι το AI4People. Ιδρύθηκε το 2017 με σκοπό την θεμελίωση αρχών, πολιτικών και πρακτικών που θα χτίσουν μια “καλή κοινωνία τεχνητής νοημοσύνης”. Τον Νοέμβριο του 2018 παρουσίασε το πρώτο πλαίσιο ηθικής καθοδήγησης για μια “καλή κοινωνία τεχνητής νοημοσύνης” όπου θα μειωθεί ο κοινωνικός αντίκτυπος και θα γίνει εκμετάλλευση των δυνατοτήτων της τεχνητής νοημοσύνης. Στο πλαίσιο αυτό αναφέρονται πέντε κύριες αρχές: ευεργετικότητα, όχι κακές συνέπειες, αυτονομία, δικαιοσύνη, και επεξηγηματικότητα (beneficence, nonmaleficence, autonomy, justice, explicability). Στο πλαίσιο γίνεται λόγος για την ανάγκη αποφυγής επικίνδυνων καταστάσεων όπως η κακή ευθυγράμμιση κινήτρων, η απληστία, η αντιφατική γεωπολιτική και η κακόβουλη πρόθεση.

Στην τελευταία συνδιάσκεψη του Ευρωπαϊκού Συμβουλίου για την τεχνητή νοημοσύνη εκδόθηκε η συμπερασματική ανακοίνωση “European Coordinated Plan on Artificial Intelligence” που αποτελεί ένα πλάνο για τις προτεραιότητες που θα ακολουθήσει η ΕΕ (πηγή <https://www.consilium.europa.eu/media/38460/st06523-en19.pdf>). Σε αυτή αναφέρεται η ανάγκη να υιοθετηθεί η τεχνητή νοημοσύνη στην δημόσια διοίκηση και επιπλέον να θεσμοθετηθεί ειδικό νομοθετικό πλαίσιο που θα εξασφαλίζει την αξιοπιστία, την ασφάλεια των δεδομένων, την ηθική περιφρούρηση της ανάπτυξης και χρήσης της. Αναφέρεται στα καλά της τεχνητή νοημοσύνης (διάγνωση ασθενειών, μείωση κατανάλωσης ενέργειας, βελτίωση του περιβάλλοντος, πρόγνωση καιρού, εξακρίβωση εγκλημάτων, οικονομική ανάπτυξη και αύξηση παραγωγικότητας) και στην ανάγκη να ενισχυθεί η ανάπτυξη της τεχνητής νοημοσύνης και να χρηματοδοτηθεί η βιομηχανία, χωρίς να απαριθμεί κανέναν από τους κινδύνους. Παράλληλα με την παραπάνω διάσκεψη ιδρύεται και η νέα ομάδα Member States' Group on Digitising European Industry and Artificial Intelligence (**MS Groupon DEI and AI**) οποία θα αποτελείται από αντιπροσωπευτικά μέλη όλων των κρατών της ένωσης και θα συζητά σχετικά θέματα με την Commission. Βλέπουμε δηλαδή μια ακόμη ομάδα που ασχολείται και αυτή μαζί με τις

προηγούμενες με σχετικά θέματα χωρίς να υπάρχει κάποια διάκριση των αντικειμένων τους.

Αν και θα έλεγε κανείς ότι θέματα που αφορούν την ηθική θα έπρεπε να περνούν αποκλειστικά από τον έλεγχο μιας μόνον επιτροπής, παρόλα αυτά υπάρχει μια σύγκυση των αρμοδιοτήτων. Όσον αφορά λοιπόν την ΕΕ υπάρχουν **πέντε ομάδες** (EGE, AI HLEG, EISMD, EAG, MS Group on DEI and AI) που ασχολούνται με την δημιουργία ενός πλαισίου αρχών περί της ηθικής της τεχνητής νοημοσύνης, πράγμα που προκαλεί μια σύγκυση σχετικά με το ποιος πραγματικά έχει την τελική ευθύνη για την χάραξη μιας ενιαίας πολιτικής. Επιπλέον αυτών των ομάδων, σχετικές απόψεις εκφράζει και μια ακόμη ομάδα της ΕΕ που εμπλέκεται σε θέματα ηθικής είτε εμμέσως είτε άμεσα. Η ομάδα αυτή είναι η Ομάδα του άρθρου 29 της ΕΕ, της οποίας την εμπλοκή θα δούμε στο παρακάτω κεφάλαιο. Η αλήθεια είναι ότι η ευθύνη είναι μόνον στην Επιτροπή και ο ρόλος αυτών των ομάδων είναι μόνον συμβουλευτικός.

7. Μια ακόμη προσπάθεια για την δημιουργία ενός πλαισίου αρχών γύρω από την τεχνητή νοημοσύνη έγινε από το ινστιτούτο THE FUTURE OF LIFE INSTITUTE (FLI). είναι η ASILOMAR AI PRINCIPLES Η αποστολή του ινστιτούτου είναι να υποστηρίξει την έρευνα που αφορά την προστασία της ζωής και την ανάπτυξη ενός αισιόδοξου οράματος για το μέλλον, βρίσκοντας θετικούς τρόπους αξιοποίησης της τεχνολογίας και αντιμετώπισης νέων προκλήσεων. Έχει εκδώσει ένα σύνολο αρχών και αξιών που παρουσιάζονται παρακάτω:

- Η έρευνα για την τεχνητή νοημοσύνη δεν θα πρέπει να είναι χωρίς κατεύθυνση αλλά μόνον προς την κατεύθυνση του κοινού καλού.
- Η χρηματοδότηση της έρευνας θα πρέπει να είναι μόνον προς την κατεύθυνση του κοινού καλού.
- Θα πρέπει να υπάρχει εποικοδομητική συνεργασία μεταξύ των ερευνητών και αυτών που χαράσσουν την πολιτική.
- Οι ερευνητές πρέπει να δρουν συνεργατικά, με διαφάνεια και εμπιστοσύνη.
- Δεν θα πρέπει να δικαιολογούνται εκπτώσεις στα πρότυπα της ασφάλειας.
- Όλα τα συστήματα τεχνητής νοημοσύνης θα πρέπει να είναι ασφαλή σε όλη τη διάρκεια της ζωής τους.
- Αν ένα σύστημα προκαλέσει βλάβη θα πρέπει να μπορεί να είναι διαφανές γιατί αυτό έγινε.
- Συστήματα που θα χρησιμοποιούνται δικαστικές αποφάσεις θα πρέπει να παρέχουν μια ικανοποιητική εξήγηση που θα ελέγχεται από μια αρμόδια ανθρώπινη αρχή.

- Οι σχεδιαστές και αυτοί που ανέπτυξαν μια εφαρμογή τεχνητής νοημοσύνης θα είναι υπεύθυνοι για αυτή, δηλ. για την χρήση της, για την κακή χρήση της, για τις ενέργειές της.
- Οι στόχοι και οι συμπεριφορές των αυτόνομων συστημάτων θα πρέπει να ευθυγραμμίζονται με τις ανθρώπινες αξίες.
- σχεδιασμός των εφαρμογών τεχνητής νοημοσύνης θα πρέπει να είναι συμβατές με τα ιδανικά της ανθρώπινης αξιοπρεπείας, των δικαιωμάτων, της ελευθερίας, και της πολιτιστικής διαφορετικότητας.
- Οι άνθρωποι θα πρέπει να έχουν το δικαίωμα της πρόσβασης, της διαχείρισης και του ελέγχου των δεδομένων που αυτοί παράγουν.
- Η χρήση των δεδομένων από τις εφαρμογές της τεχνητής νοημοσύνης δεν θα πρέπει αδικαιολόγητα να περιορίζει την πραγματική ή την αντιληπτή ελευθερία.
- Οι τεχνολογίες πρέπει να ευεργετούν και να δίνουν δύναμη σε όσον το δυνατόν περισσότερους ανθρώπους.
- Το οικονομικό αποτέλεσμα από την εφαρμογή της τεχνητής νοημοσύνης πρέπει να μοιράζεται ευρέως και προς το καλό της ανθρωπότητας.
- Ο τελικός έλεγχος θα παραμένει στον άνθρωπο, αυτός θα επιλέγει το πώς και το πότε θα μεταβιβάζει αποφάσεις στην τεχνητή νοημοσύνη για επιτευχθούν στόχοι που τέθηκαν από ανθρώπους.
- Η εξουσία που θα παρέχετε στα αυτόνομα συστήματα δεν θα πρέπει να ανατρέπεται και θα πρέπει να σέβονται τις πολιτικές διαδικασίες στις οποίες βασίζεται η υγεία της κοινωνίας.
- Δεν πρέπει να γίνουν αποδεκτά τα αυτόνομα οπλικά συστήματα.
- Δεν πρέπει να δεχτούμε ότι υπάρχουν όρια στις δυνατότητες της τεχνητής νοημοσύνης.
- Η σημαντικότητα της τεχνητής νοημοσύνης αντιπροσωπεύει μια προφανή αλλαγή στην ιστορία της ζωής στην Γή, και πρέπει να διαχειριστεί με την ανάλογη προσοχή και την ανάλογη χρηματοδότηση.
- Οι κίνδυνοι οι όποιοι υπάρχουν, ιδιαίτερα οι υπαρξιακοί και οι καταστροφικοί, θα πρέπει να γίνουν αντικείμενο σχεδιασμού, προσπαθειών μετριασμού, αναλόγων με των αναμενόμενων επιπτώσεών τους.
- Η ικανότητα των συστημάτων της τεχνητής νοημοσύνης να αναπαράγονται και να βελτιώνονται θα πρέπει να γίνει αντικείμενο λήψης μέτρων για την ασφάλεια των ανθρώπων και τον έλεγχο των μηχανών.

- Η υπερνοημοσύνη θα πρέπει να αναπτυχτεί μόνον προς την υπηρεσία των κοινών ηθικών ιδεωδών για το καλό της ανθρωπότητας και όχι για τα καλό ενός οργανισμού.

Πρέπει εδώ να σημειωθεί ότι για να διαχυθούν αυτές οι αρχές και για να γίνει η κατάλληλη έρευνα, η παραπάνω συνεργασία υποστήριξε αυτή τη προσπάθεια δημιουργώντας ένα ταμείο και δίνοντας μια μεγάλη χρηματοδότηση. Το Ταμείο αποτελεί έναν από τους πολλούς νέους φορείς που στόχο έχουν τη διαμόρφωση ενός θετικού μέλλοντος για την AI και την ανθρωπότητα.

Είναι σημαντικό να αναφερθούμε και σε μια επιπλέον ανακοίνωση της παραπάνω συνεργασίας που έχει μεν κάποια σημεία που ήδη έχουν αναφερθεί πριν αλλά έχει και σημεία που δεν έχουν ξαναφερθεί. Για αυτό το λόγο παρουσιάζονται στην συνέχεια ακόμη ένα σύνολο αρχών από την ίδια ομάδα.

-τα συστήματα τεχνητής νοημοσύνης πρέπει να είναι ασφαλή για όλη τη διάρκεια της ζωής τους,

-αν ένα σύστημα αποτύχει θα πρέπει να υπάρχει διαφάνεια στην εύρεση της αιτίας που προκάλεσε την αποτυχία.

-αν κάποιο σύστημα εμπλέκεται στην διαδικασία λήψεων αποφάσεων σε δίκη θα πρέπει να παρέχεται ικανοποιητική εξήγηση από την αρμόδια ανθρώπινη αρχή.

- Οι σχεδιαστές και οι δημιουργοί ενός προηγμένου συστήματος τεχνητής νοημοσύνης θα πρέπει να είναι οι κύριοι υπεύθυνοι και υπόλογοι για την χρήση, τις λειτουργίες, την αποτυχία και την επιδιόρθωση του συστήματος που έχουν δημιουργήσει.

- τα αυτόνομα συστήματα τεχνητής νοημοσύνης θα πρέπει να σχεδιάζονται έτσι ώστε οι στόχοι και οι συμπεριφορές τους να είναι ασφαλώς ευθυγραμμισμένες με τις ανθρώπινες αξίες.

- τα συστήματα τεχνητής νοημοσύνης θα πρέπει να σχεδιάζονται και να λειτουργούν συμβατά με τις αρχές της ανθρώπινης υπερηφάνειας, των ανθρώπινων δικαιωμάτων, την ελευθερία, και την διαφορετικότητα στις κουλτούρες.

- οι άνθρωποι θα πρέπει να έχουν το δικαίωμα στην πρόσβαση και στην διαχείριση των δεδομένων που αυτοί παράγουν και έχουν δοθεί σε συστήματα τεχνητής νοημοσύνης.

- Οι εφαρμογές της τεχνητής νοημοσύνης που χρησιμοποιούν προσωπικά δεδομένα δεν θα πρέπει να περιορίζουν την πραγματική και την αντιληπτή ελευθερία του ατόμου.

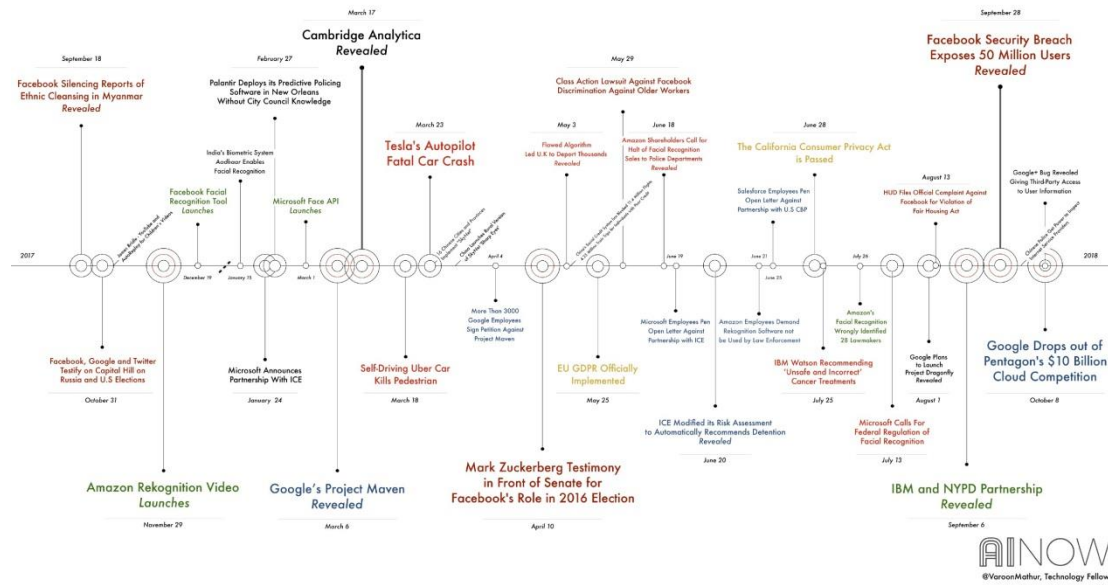
- Η τεχνολογία της τεχνητής νοημοσύνης θα πρέπει να ωφελεί και να ενδυναμώνει όσους περισσότερους ανθρώπους είναι δυνατόν και όχι λίγους.

- το οικονομικό κέρδος που δημιουργείται από την τεχνητή νοημοσύνη θα πρέπει να μοιράζεται ευρέως, με σκοπό να αποκομίζει καρπούς όλη η ανθρωπότητα.
- οι άνθρωποι θα πρέπει να είναι αυτοί που θα παίρνουν την τελική απόφαση αν θα παραχωρήσουν σε συστήματα τεχνητής νοημοσύνης το δικαίωμα λήψης αποφάσεων, μόνον προς επίτευξη στόχων που έχουν επιλεγεί από τον άνθρωπο.
- Η εξουσία που παρέχεται από τον έλεγχο των πολύ προηγμένων συστημάτων τεχνητής νοημοσύνης πρέπει να σέβεται και να βελτιώνει, αντί να ανατρέπει, τις κοινωνικές και πολιτικές διαδικασίες στις οποίες εξαρτάται η υγεία της κοινωνίας.
- πρέπει να απαγορευθεί η δημιουργία αυτόνομων οπλικών συστημάτων.

8. Στην προσπάθεια να δημιουργηθεί ένα πλαίσιο αρχών γύρω από την τεχνητή νοημοσύνη έχουν συμβάλει και πολλά πανεπιστημιακά ιδρύματα. Μια από τις πιο αξιόλογες προσπάθειες είναι αυτή του πανεπιστημίου της Νορβηγίας (**UiT**). Η ομάδα Ethics Research Group ανήκει στο τμήμα Φιλοσοφίας του πανεπιστημίου, για αυτό το λόγο δίνει μια πιο θεωρητική, και όχι με όρους τεχνολογικούς, προσέγγιση του θέματος της κατάκτησης της ευτυχίας. Σε μια από αυτές τις φιλοσοφικές αναζητήσεις αναφέρετε ότι το κακό πρέπει να υποτάσσεται στις ηθικές αρχές. Αυτό είναι ίσως και το σημαντικότερο συμπέρασμα που μπορεί να χαρακτηριστεί σαν μια θεμελιώδης αρχή στο ξεκίνημα της αποδοχής της τεχνητής νοημοσύνης, ότι δηλαδή θα πρέπει να γίνει αποδεκτή μέχρι τον βαθμό που δεν θα προσβάλλει το ανθρώπινο πνεύμα.

9. Μια ακόμη ιδιωτική πρωτοβουλία που ερευνά τις επιπτώσεις της τεχνητής νοημοσύνης στην κοινωνική ζωή είναι το ινστιτούτο **AI Now Institute** του πανεπιστημίου της Νέας Υόρκης. Τα κύρια πεδία των ερευνών του είναι η επιπτώσεις στα δικαιώματα και την ελευθερία του ανθρώπου, οι επιπτώσεις στην εργασία, οι πιθανές διακρίσεις και οι πιθανοί κίνδυνοι στην ασφάλεια. Στην τελευταία του έκθεση αναφέρει ότι η περασμένη χρονιά ήταν δραματική για την τεχνητή νοημοσύνη (δείτε το παρακάτω σχήμα (Σχήμα 2. 2018 AI History), με τα γεγονότα που δικαιολογούν αυτή τη γνώμη) και κάνει λόγο για δέκα προτάσεις για την εφαρμογή της τεχνητής νοημοσύνης.

Σχήμα 2. AI History 2018 (πηγή: AI Now Institute).



Οι σημαντικότερες από αυτές τις προτάσεις είναι οι εξής:

- οι κυβερνήσεις πρέπει να βάλουν κανόνες στην τεχνητή νοημοσύνη ανά τομέα ενδιαφέροντος,

- η αναγνώριση των γεωμετρικών χαρακτηριστικών από κάμερες και η μελέτη των ευρημάτων (π.χ. συναισθήματα, ψυχική υγεία, προσωπικότητα) θα πρέπει να μπορεί να μην γίνεται δεκτή από τους πολίτες,

- η εσωτερική δομή των εταιρειών που ασχολούνται με την τεχνητή νοημοσύνη θα πρέπει να είναι διαφανής και έχει τις κατάλληλες δομές για τον έλεγχο της λειτουργίας τους,

- οι εφαρμογές της τεχνητής νοημοσύνης που χρησιμοποιούνται από κυβερνήσεις θα πρέπει να παραιτηθούν από οποιοδήποτε εμπορικό απόρρητο ή άλλη νομική απαίτηση που εμποδίζει τον πλήρη έλεγχο και κατανόηση του λογισμικού τους,

- οι εργαζόμενοι στην δημιουργία συστημάτων τεχνητής νοημοσύνης θα πρέπει να έχουν το δικαίωμα να γνωρίζουν τον σκοπό των δημιουργημάτων τους και να έχουν προστασία όταν αντιδρούν σε αυτό και όταν αποκαλύπτουν τις γνώσεις τους στο κοινό.

- οι ομάδες ανθρώπων που θα πληγούν περισσότερο είναι τα χαμηλά κοινωνικά και οικονομικά στρώματα για αυτό και θα πρέπει να αντιμετωπιστούν ξεχωριστά με στοχευμένη χρηματοδότηση.

- Επειδή η τεχνητή νοημοσύνη έχει κινδύνους για την κοινωνική ζωή των ανθρώπων θα πρέπει στην έρευνα γύρω από αυτή να συμμετέχουν επιστήμονες από πολλά πεδία

της επιστήμης όπως κοινωνιολογία, ανθρωπολογία κ.αλ. και να χρηματοδοτούνται κατάλληλα.






Είναι αξιοσημείωτο ότι το ινστιτούτο αναγνωρίζει το γεγονός ότι ο έλεγχος της τεχνητής νοημοσύνης αφορά επιστήμες έξω από την πληροφορική, πράγμα το οποίο επισημαίνεται και από την πρόταση του συγγραφέα που ακολουθεί στο επόμενο κεφάλαιο, για την δημιουργία μιας ομάδας ηθικής από φιλοσόφους, κοινωνιολόγους και ανθρωπολόγους.

10. Μια ακόμη προσπάθεια για την δημιουργία ενός συνόλου ηθικών αρχών έγινε από την μεγαλύτερη οργάνωση στον χώρο των υπολογιστών **Association for Computing Machinery**. Οι αρχές (ACM code of ethics. 2018) αυτές αναφέρονται συνοπτικά παρακάτω:

- Πρέπει να υπάρχει συμβολή στην κοινωνία και να εξυπηρετείται η καλή διαβίωση.
- Να αποφεύγεται το κακό
- Να υπάρχει ειλικρίνεια και αξιοπιστία.
- Να υπάρχει δικαιοσύνη και να λαμβάνονται δράσεις ενάντια στις διακρίσεις.
- Να υπάρχει σεβασμός στην δουλειά που έγινε για να παραχθεί μια ιδέα, εφεύρεση και δημιουργία.
- Να υπάρχει σεβασμός στην ιδιωτικότητα.
- Να γίνετε σεβαστή η εμπιστευτικότητα.

11. Η ιδιωτική εταιρεία **X2AI** έχει εκδώσει ένα σύνολο ηθικών αρχών, το οποίο είναι συνδυασμός των αρχών του American Psychological Association (APA) και των αρχών του UK Lord's Report AI Code Principles όπως παρουσιάζονται στο παρακάτω σχήμα (Σχήμα 3. Οι αρχες της X2AI).

Σχήμα 3. Οι ηθικές αρχές της X2AI (πηγή: X2AI).

APA General Principles for Psychologists	χ^2	Lord's Report AI Code Principles
Beneficence and Nonmaleficence - Strive to benefit those with whom they work and take care to do no harm.		Artificial intelligence should be developed for the common good and benefit of humanity .
Fidelity and Responsibility - Establish relationships of trust with those with whom they work.		Artificial intelligence should operate on principles of intelligibility and fairness .
Integrity - Promote accuracy, honesty, and truthfulness in the science, teaching, and practice of psychology.		The autonomous power to hurt, destroy or deceive human beings should never be vested in artificial intelligence.
Justice - Fairness and justice entitle all persons to access and benefit from the contributions of psychology and to equal quality in the processes, procedures, and services.		All citizens have the right to be educated to enable them to flourish mentally, emotionally and economically alongside artificial intelligence.
Respect for People's Rights and Dignity - Respect the dignity and worth of all people, and the rights of individuals to privacy, confidentiality, and self-determination.		Artificial intelligence should not be used to diminish the data rights or privacy of individuals, families or communities.

Οι παραπάνω αρχές για την τεχνητή νοημοσύνη εν συντομία είναι οι εξής;

-Η τεχνητή νοημοσύνη πρέπει να εξυπηρετεί το κοινό καλό και να είναι ωφέλιμη στην ανθρωπότητα.

-να είναι αξιόπιστη.

-να μην προσβάλει την ανθρώπινη ακεραιότητα.

-να είναι δίκαιη.

-να σέβεται τα ανθρώπινα δικαιώματα και την ανθρώπινη αξιοπρέπεια.

12. Το διεθνές ινστιτούτο ηλεκτρολόγων και ηλεκτρονικών μηχανικών, **IEEE**, έχει εκδώσει το έγγραφο "Ethically Aligned Design" (http://standards.ieee.org/develop/indconn/ec/ead_v2.pdf) το οποίο βρίσκεται ήδη στην δεύτερη έκδοση του και αναφέρει μεταξύ άλλων τις παρακάτω αρχές: τα επιτεύγματα της τεχνητής νοημοσύνης να μην προσβάλουν τα ανθρώπινα δικαιώματα, να σχεδιάζονται και να λειτουργούν με σκοπό το κοινό καλό (κοινό οικονομικό κέρδος, εργασιακή ασφάλεια, υγεία), οι χρήστες και οι δημιουργοί τους να είναι υπόλογοι και υπεύθυνοι για αυτά, να λειτουργούν με διαφάνεια, επεξήγηση,

δικαιολόγηση των αποφάσεών τους, προς τον εξωτερικό κόσμο, να ελαχιστοποιείται το ρίσκο μιας πιθανής κακής χρήσης τους.

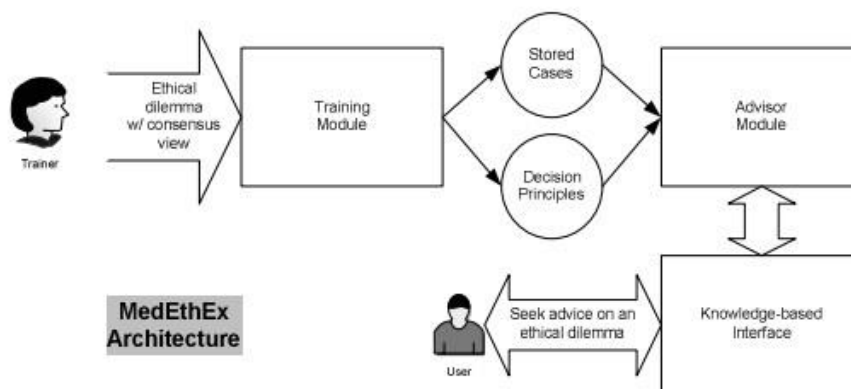
13. Η διεθνής ένωση τηλεπικοινωνιών, **ITU**, φιλοξένησε σε ετήσια διεθνή συνάντηση την δεύτερη προσπάθεια των Ηνωμένων Εθνών για μια ενοποιημένη πλατφόρμα που αφορά την τεχνητή νοημοσύνη, με την ονομασία “AI for Good” η οποία όμως είναι σε πρώιμο στάδιο και δεν παρήγαγε ένα τελικό ενιαίο εφαρμόσιμο αποτέλεσμα. Ο ΟΗΕ μέσω του γενικού γραμματέα του εξέφρασε την πίστη του ότι η σύνοδος του θα εξασφαλίσει ότι η τεχνητή νοημοσύνη θα χρησιμοποιηθεί προς το κοινό συμφέρον και θα ωθήσει τις κοινές αρχές που διαμοιράζονται τα μέλη του οργανισμού, και ότι είναι μεγάλο ρίσκο να μείνει κάποια χώρα πίσω στην ανάπτυξη αν δεν την χρησιμοποιήσει. Επίσης επισημάνθηκε η δυνατότητα ανάπτυξης που θα δώσει η τεχνητή νοημοσύνη στην εκπαίδευση και την υγεία ιδιαίτερα σε υποανάπτυκτες χώρες.

14. Σε ένα ακόμη συνέδριο, αυτό της **Ένωσης Υπολογιστικών Μηχανών ACM** και της Ένωσης για την Πρόοδο της Τεχνητής Νοημοσύνης **AAAI**, με θέμα την τεχνητή νοημοσύνη και την σχέση της με την ηθική και την κοινωνία, συζητήθηκε η δημιουργία ενός πλαισίου ηθικών αρχών για την τεχνητή νοημοσύνη, Βέβαια οι δημοσιεύσεις δεν έχουν ελεύθερη πρόσβαση και επίσης αυτού του είδους τα συνέδρια δεν καταλήγουν συμφωνώντας σε ένα τελικό κείμενο.

15. Επίσης στα πλαίσια της συνδιάσκεψης που προηγήθηκε τον περασμένο Φεβρουάριο από τον διεθνή οργανισμό **AAAI**, παρουσιάστηκε μεταξύ άλλων μια εφαρμογή τεχνητής νοημοσύνης, **MedEthEx**, (**MedEthEx:A Prototype Medical Ethics Advisor**, Michael Anderson,, Susan Leigh Anderson,, Chris Armen. Published in **AAAI**. 2006), η οποία ακολουθεί ένα σύνολο αρχών όπως σεβασμός στην αρχή της αυτονομίας, σεβασμός στην αρχή της ευεργεσίας (του κοινού καλού), σεβασμός στην αρχή της δικαιοσύνης, σεβασμός στην αρχή της μη ικανοποίησης. Η εφαρμογή αυτή με την συσσωρευμένη γνώση που έχει και σε συνδυασμό με τις τέσσερις παραπάνω αρχές, λειτουργεί σαν σύμβουλος σε θέματα ηθικής σε περιπτώσεις ιατρικών διλημάτων βοηθώντας τους νοσηλευτές να πάρουν μια σωστή απόφαση. Παράδειγμα, αν ένας ασθενής αρνηθεί μια θεραπευτική αγωγή η εφαρμογή θα δώσει στον νοσηλευτή μια συμβουλή για το πώς θα πρέπει να χειριστεί την αντίδρασή του στην στάση του ασθενή. Σε αυτή τη “μηχανή” βλέπουμε κάτι πρωτοποριακό, δεν είναι απλά μια προσπάθεια να περιφρουρηθεί ηθικά μια εφαρμογή αλλά έχει προχωρήσει ένα βήμα περαιτέρω, και έχει δημιουργηθεί η πρώτη μηχανή που δίνει συμβουλές ηθικής φύσεως. Βέβαια η μηχανή δεν απαντά σε μεγάλα και γενικά ηθικά διλήματα αλλά σε ερωτήματα συγκεκριμένου πεδίου (ιατρικού).

Η εφαρμογή αυτή ακολουθεί την διαδικασία αποφάσεων που φαίνεται στο παρακάτω σχήμα (Σχήμα 4. Η διαδικασία αποφάσεων στην εφαρμογή τεχνητής νοημοσύνης MedEthEx.).

Σχήμα 4. Η διαδικασία αποφάσεων στην εφαρμογή τεχνητής νοημοσύνης MedEthEx (πηγή: Πηγή: <https://aaai.org/Papers/AAAI/2006/AAAI06-292.pdf>).



Προς την ίδια κατεύθυνση με την μηχανή MedEthEx, και στην ίδια συνδιάσκεψη παρουσιάστηκε και η εφαρμογή GENETH (Michael Anderson, Susan Leigh Anderson. 2014) η οποία κωδικοποιεί και αναλύει όλες τις πλευρές ενός ηθικού διλήματος. Μερικές από τις αρχές που λαμβάνει υπό όψιν είναι οι: σεβασμός στην αυτονομία, η εξυπηρέτηση του καλού, η αποφυγή κάποιας βλάβης.

16. Μια ακόμη πρωτοβουλία προς την κατεύθυνση της θεσμοθέτησης αρχών γύρω από την χρήση της τεχνητής νοημοσύνης είναι και αυτή του πανεπιστημίου του Μόντρεαλ, το οποίο έκανε την δήλωση “The Declaration Responsible AI”. Σε αυτή τη δήλωση θα βρει κανείς το παρακάτω πλαίσιο αρχών:

A. Η τεχνητή νοημοσύνη θα πρέπει να εξυπηρετεί το καλό όλων των συνειδητών οντοτήτων.

B. Η ανάπτυξη της τεχνητής νοημοσύνης πρέπει να προωθεί την αυτονομία των ανθρώπων και να ελέγχει με υπεύθυνο τρόπο την αυτονομία των υπολογιστικών συστημάτων.

Γ. Η ανάπτυξη της τεχνητής νοημοσύνης πρέπει να προωθεί την δικαιοσύνη και να προσπαθεί να εξαλείψει όλους τους τύπους των διακρίσεων (φύλλο, ηλικία, φυσικές και διανοητικές ικανότητες, σεξουαλικός προσανατολισμός, κοινωνική καταβολή, θρησκευτικές πεποιθήσεις).

Δ. Η ανάπτυξη της τεχνητής νοημοσύνης πρέπει να προσφέρει εγγυήσεις σεβασμού της ιδιωτικότητας.

Ε. Η ανάπτυξη της τεχνητής νοημοσύνης πρέπει να προωθεί την κριτική σκέψη και να μας προστατεύει από προπαγάνδα και χειραγώγηση.

ΣΤ. Η ανάπτυξη της τεχνητής νοημοσύνης πρέπει να προωθεί την εγνωσμένη συμμετοχή στη δημόσια ζωή, στην συνεργασία και τον δημοκρατικό διάλογο.

Ζ. Οι διάφοροι εμπλεκόμενοι στην ανάπτυξη της τεχνητής νοημοσύνης πρέπει να αναλάβουν τις ευθύνες τους δουλεύοντας προς την κατεύθυνση να μειώσουν τα ρίσκα των τεχνολογικών επιτευγμάτων τους.

17. Ένα ακόμη ινστιτούτο που δημιούργησε μια ομάδα για την μελέτη γύρω από θέματα ηθικής της τεχνητής νοημοσύνης είναι και το ινστιτούτο Alan Turing με την ομάδα του Data Ethics Group (<https://www.turing.ac.uk/research/interest-groups/data-ethics-group>). Η ομάδα αυτή μελετά τον ηθικό και κοινωνικό αντίκτυπο της χρησιμοποίησης των προσωπικών δεδομένων. Η πιο πρόσφατη ενημέρωση της είναι αυτή του Νοεμβρίου 2018 στην οποία παρέχει στην επιτροπή National Analytics Solution (επιτροπή της Βρετανικής αστυνομίας για την προχωρημένη ανάλυση προσωπικών δεδομένων) το απαραίτητο ηθικό-τεχνολογικό υπόβαθρο σαν μια γενική στρατηγική αντίληψη με σκοπό την πρόληψη της εγκληματικότητας μέσα σε ένα πλαίσιο νομιμότητας και διαφάνειας. Δυστυχώς αν και αυτή η επιτροπή μιλά για διαφάνεια οι δικές τις πράξεις παραμένουν αδιαφανείς μιας και οι γνωμοδοτήσεις της δεν είναι δημοσιοποιημένες εξ ολοκλήρου, επίσης σε μια από τις γενικές εισηγήσεις της γίνεται αναφορά ότι στο μέλλον οι πολίτες θα κυβερνούνται με βάση τα δεδομένα που θα παράγονται από την παρατήρηση της συμπεριφοράς τους αλλά και των αλληλεπιδράσεων τους με άλλους.

18. Μια παρόμοια προσπάθεια για την δημιουργία ενός συνόλου αρχών είναι και η πρόταση για την συγγραφή ενός **“όρκου της τεχνητής νοημοσύνης”** παρόμοιου με αυτόν του όρκου του Ιπποκράτη που δίνουν οι ιατροί κατά την απόκτηση της άδειας άσκησης του επαγγέλματός τους. Η πρόταση αυτή κατατέθηκε σε πρόσφατο βιβλίο των Brad Smith και Harry Shum (Brad Smith, Harry Shum. 2018). Σύμφωνα με αυτόν θα πρέπει όλοι όσοι ασχολούνται με την τεχνητή νοημοσύνη να συμφωνήσουν σε έναν σύνολο αρχών που θα ακολουθούν οι ίδιοι και θα “κληροδοτηθούν” και στα τεχνολογικά τους επιτεύγματα. Όρκος αποτελείται απ τις παρακάτω αρχές:

“Ορκίζομαι ότι θα εκπληρώσω με όλη μου την ικανότητα και κρίση μου αυτή τη διαθήκη:

Α. Θα σέβομαι τους κόπους των επιστημόνων που χρησιμοποίησα, και με χαρά θα μοιράζομαι αυτή τη γνώση με αυτούς που θα ακολουθήσουν..

Β. Θα εφαρμόσω, για τα καλά της ανθρωπότητας, όλα τα απαραίτητα μέτρα που χρειάζονται, αποφεύγοντας την δίδυμη παγίδα της υπερβολικής αισιοδοξίας και της ομοιόμορφης απαισιοδοξίας.

Γ. Θα έχω πάντοτε υπό όψιν ότι μέσα στην τεχνητή νοημοσύνη υπάρχει η τέχνη και η επιστήμη, και ότι οι ανθρώπινες ανησυχίες ξεπερνούν τις τεχνολογικές.

Δ. Ειδικά όταν πρόκειται για διλήματα ζωής ή θανάτου θα πρέπει να δρω με προσοχή. Αν μου ζητηθεί να σώσω μια ζωή χρησιμοποιώντας την τεχνητή νοημοσύνη, αυτό είναι ευπρόσδεκτο. Όμως όταν πρέπει να χρησιμοποιηθεί η τεχνητή νοημοσύνη για αφαίρεση μιας ζωής, αυτή η ξεχωριστά βαρύνουσα πράξη πρέπει να αντιμετωπιστεί με μεγάλη ταπεινότητα και επίγνωση της δικής μου ταπεινότητας και των ορίων της τεχνητής νοημοσύνης. Πάνω από όλα δεν πρέπει να παίζω το ρόλο του θεού, ούτε να αφήνω την τεχνολογία μου να το κάνει αυτό.

Ε. Θα σέβομαι την ιδιωτικότητα των ανθρώπων όσον αφορά τα προσωπικά τους δεδομένα, ώστε να μην αποκαλύπτονται στον υπόλοιπο κόσμο από τα συστήματα τεχνητής νοημοσύνης.

ΣΤ. Θα λάβω υπ όψιν μου τον αντίκτυπο της εργασίας μου στην δίκαια αντιμετώπιση της διαιώνισης ιστορικών προκαταλήψεων και την δημιουργία καταστάσεων που αυξάνουν την οικονομική ή άλλη ανισότητα.

Ζ. Η τεχνητή νοημοσύνη που θα αναπτύξω θα προφυλάσσει από τα κακό οποτεδήποτε μπορεί, γιατί η πρόληψη είναι καλύτερη της θεραπείας.

Η. Η τεχνητή νοημοσύνη που θα αναπτύξω θα συνεργάζεται με τους ανθρώπους για τα κοινά καλά, παρά να σφετεριστεί τον ανθρώπινο ρόλο ώστε να τον υποκαταστήσει.

Θ. Θα θυμάμαι πάντοτε ότι δεν διαχειρίζομαι απλά δεδομένα από μηδενικά και άσους, αλλά ανθρώπινες υπάρξεις, των οποίων οι συναλλαγές με την τεχνητή νοημοσύνη μπορεί να επηρεάσουν την ελευθερία, την οικογένεια, και την οικονομική σταθερότητα. Είναι δική μου υπευθυνότητα να λάβω υπο όψιν αυτά τα προβλήματα.

Ι. Θα θυμάμαι πάντοτε ότι παραμένω ένα μέλος αυτής της κοινωνίας, με ξεχωριστή υποχρέωση σε όλες τις ανθρώπινες υπάρξεις. ’’.

19. Κλείνοντας την αναφορά σε ομάδες που ασχολούνται με την δημιουργία ηθικών αρχών, πρέπει να αναφερθεί ότι στην χώρα μας δεν υπάρχει μια κρατική ομάδα ηθικής, παρά μόνον μια **Ομάδα Βιοηθικής** και μια **Επιτροπή Ηθικής και Δεοντολογίας της Έρευνας (ΕΗΔΕ)** του Ιδρύματος Τεχνολογίας και Έρευνας (ΙΤΕ) που εποπτεύεται από τη Γενική Γραμματεία Έρευνας και Τεχνολογίας (ΓΓΕΤ) του

Υπουργείου Παιδείας, Έρευνας και Θρησκευμάτων. Σε ερώτηση μου προς την Επιτροπή βιοηθικής για το αν έχουν εκδώσει κάποιες αρχές που αφορούν την σχέση της τεχνολογίας με τον άνθρωπο δεν έλαβα καμιά απάντηση, ούτε υπάρχει κάποια αναφορά σε αρχές στην επίσημη ιστοσελίδα τους (<http://www.bioethics.gr>).

Η παραπάνω προτάσεις είναι πολύ καλές αφού συνδράμουν προς την επιθυμητή κατεύθυνση της ηθικής περιφρούρησης, όμως όπως έχει επισημανθεί και νωρίτερα, επειδή πολλοί είναι αυτοί οι οποίοι παραβαίνουν τους όρκους και τις αρχές τους, θα πρέπει αυτές να οχυρωθούν και νομικά. Προς αυτή τη κατεύθυνση, θα πρέπει να ποινικοποιηθούν κάποιες κακές πρακτικές, ώστε να αποτρέπεται η παραβίαση των αρχών.

9.3. Η επιτακτική ανάγκη για την ίδρυση ομάδων ηθικής

Στην παρουσίαση που έγινε παραπάνω για τις ομάδες που έχουν συμμετέχει στην προσπάθεια δημιουργίας ενός συνόλου αρχών, βλέπουμε ότι καμία από αυτές δεν έχει έναν ουσιαστικό ρόλο στην λήψη των αποφάσεων που αφορούν την τεχνητή νοημοσύνη. Ακόμη και οι ομάδες της ΕΕ, έχουν έναν επιτελικό ρόλο, η γνώμη των οποίων μπορεί να μην εισακουστεί από τους αρμόδιους για την λήψη των αποφάσεων.

Προς τη κατεύθυνση της έγκαιρης αντιμετώπισης θα πρέπει, και αυτό αποτελεί πρόταση του συγγραφέα, να συσταθεί σε κάθε έθνος μια ανεξάρτητη Επιτροπή Ηθικής, ανώτερη όλων των άλλων θεσμών. Καθήκον της επιτροπής θα είναι η αρχική εξέταση μιας νέας τεχνολογικής εφαρμογής και ο έλεγχος για τα οφέλη ή της τους πιθανούς μελλοντικούς κινδύνους που θα προκαλέσει στην ανθρωπότητα. Η τελική απόφαση για αποδοχή ή μη μιας τεχνολογικής εφαρμογής θα είναι στην αποκλειστική κρίση αυτής της επιτροπής, η οποία θα έχει μεγαλύτερη ισχύ και από τις αποφάσεις της νομοθετικής εξουσίας. Με αυτό τον τρόπο η ηθική θα λάβει την θέση που της αξίζει σε μια κοινωνία, η οποία είναι η ανώτερη και ισχυρότερη όλων των άλλων εξουσιών. Έτσι ένα έθνος θα αποκτήσει μια μακροπρόθεσμη στρατηγική με σκοπό την ευτυχία των πολιτών της.

Στην παρούσα παγκόσμια κατάσταση δεν υπάρχει αυτή η επιτροπή. Ειδικότερα και στην χώρα μας δεν υπάρχει τέτοια επιτροπή.

Η μόνη επιτροπή που υπάρχει στη χώρα μας είναι η Επιτροπή Βιοηθικής η οποία ασχολείται κυρίως με ιατρικά θέματα και ο ρόλος της είναι μόνο συμβουλευτικός και οι αποφάσεις της μη υποχρεωτικές.

Σε κατώτερο επίπεδο υπάρχει η Επιτροπή Ηθικής και Δεοντολογίας (ΕΗΔ) του ΙΤΕ (Ίδρυμα Τεχνολογίας και Έρευνας, Εποπτεύεται από τη Γενική Γραμματεία Έρευνας και Τεχνολογίας (ΓΓΕΤ) του Υπουργείου Παιδείας, Έρευνας και Θρησκευμάτων) η οποία ελέγχει μόνο τις καλές ερευνητικές πρακτικές του ΙΤΕ.

Προς ενίσχυση λοιπόν του αισθήματος ασφάλειας των πολιτών και προς την κατεύθυνση της έγκυρης αντιμετώπισης εκτάκτων καταστάσεων αλλά και πιστής τήρησης των ηθικών αρχών, θα πρέπει να δημιουργηθούν κατά πρώτον ο θεσμός του **συμβούλου ηθικής** (Ethics advisor), και κατά δεύτερον, αν η ευθύνη και το βάρος των αποφάσεων είναι μεγάλο για να υποστηριχτεί από ένα πρόσωπο, να θεσμοθετηθεί η υποχρεωτική λειτουργία **ομάδας ηθικής** (Ethics team) σε κάθε επιχείρηση που ασχολείται με τεχνολογία τεχνητής νοημοσύνης επίσης σε κάθε επιχείρηση που χρησιμοποιεί αυτή τη τεχνολογία ακόμη και τέλος στον δημόσιο τομέα, όπου διαχειρίζεται εφαρμογές τεχνητής νοημοσύνης. Να σημειωθεί εδώ ότι πολλές εταιρείες που ήδη ασχολούνται με εφαρμογές τεχνητής νοημοσύνης έχουν, χωρίς αυτό να είναι νομική υποχρέωση, τέτοια ομάδα, παράδειγμα είναι γνωστή εταιρεία κοινωνικής δικτύωσης της οποίας ο πρόεδρος σε πρόσφατη κατάθεση ενώπιον του Αμερικανικού Κογκρέσου δήλωσε ότι απασχολεί ομάδα ειδικών σε θέματα ηθικής για να παρακολουθούν το πρόγραμμα τεχνητής νοημοσύνης για τον έλεγχο του περιεχομένου των πληροφοριών (αναγνώριση ρητορικής μίσους, και πρόληψη τρομοκρατικών ενεργειών) που συλλέγει η πλατφόρμα του. Θα πρέπει όμως η ομάδα αυτή να μην είναι απλώς μια συμβουλευτική ή επιτελική ομάδα ή το χειρότερο μια ομάδα που δίνει ένα ηθικό άλλοθι, αλλά να έχει την σημαντικότερη θέση και την τελική έγκριση μέσα σε μια επιχείρηση ή έναν δημόσιο φορέα ή ακόμη περισσότερο σε όλη την κυβέρνηση. Αυτό σημαίνει ότι οποιαδήποτε απόφαση παίρνεται θα πρέπει πρωτίστως και υποχρεωτικά να περνά από την ομάδα αυτή και χωρίς την έγκρισή της να μην μπορεί μια ιδέα να προχωρήσει. Έτσι το κυρίαρχο ερώτημα “τι είναι καλό για την ζωή μας” θα απαντάται πρώτο πριν η επιχείρηση ή ο φορέας ή η κυβέρνηση προχωρήσει στην εφαρμογή μιας καινούργιας τεχνολογικής ιδέας. Με αυτό τον τρόπο η ηθική θα πάρει την θέση που της αρμόζει η οποία είναι πάνω από οποιαδήποτε άλλη αρχή, ακόμη και από αυτό που σήμερα νοείται σαν κυβέρνηση.

Σε πρακτικό επίπεδο, ένα τεστ που πρέπει να κάνει η ομάδα ηθικής πριν πάρει οποιαδήποτε απόφαση είναι να κάνει μια **προσομοίωση** της τεχνολογίας πριν ακόμη αυτή φτάσει στους πολίτες (Thomas King. 2018). Βέβαια αυτή η προσομοίωση θα ήταν εφικτή για προβλήματα που αφορούν π.χ. οικονομικές απάτες μέσω του διαδικτύου, ή σεξουαλικές προσβολές, αλλά όταν πρέπει να γίνει κάποιο πείραμα σε μεγάλο πλήθος ανθρώπων π.χ. να μετρηθεί η συμπεριφορά και η μακροχρόνια επίδραση στους ανθρώπους σε ένα χώρο που εποπτεύεται πλήρως από κάμερες, αυτό θα ήταν σχεδόν αδύνατο, ιδιαίτερα σε μεγάλη κλίμακα (γεωγραφική ή πληθυσμιακή). Ή ακόμη όταν πρόκειται να ληφθεί η απόφαση να επιτραπεί σε ένα ρομπότ να παίρνει ηθικές αποφάσεις, η μελλοντική εξέλιξη είναι απρόβλεπτη και πρέπει πρωτίστως να διερευνηθεί σε πειραματικό επίπεδο. Ένα τέτοιο παράδειγμα σύγκρουσης ηθικών αρχών που θα μπορούσε να αντιμετωπίσει ένα ρομπότ, σαν ένας “ηθικός πράτων” (Fahad Alaiერი, Andr’e Vellino. 2016) είναι η σύγκρουση της αρχής “να προστατεύεται η ελευθερία” με την αρχή “να υπακούσουμε στους νόμους” όταν βέβαια ένας νόμος λέει ότι πρέπει κάποιος να στερηθεί την ελευθερία του και να φυλακισθεί. Είναι το ίδιο ηθικό δίλλημα που αντιμετωπίζει και ένας δικαστής που

βάζει στον ζυγό δύο δικαιώματα. Συνεπώς η προσομοίωση σε πολλές περιπτώσεις είναι δύσκολο να εφαρμοστεί.

Είναι λοιπόν η ηθική μια αρχή, για αυτό πρέπει να αντιμετωπίζεται πρώτη. Αυτό σημαίνει ότι πριν ληφθεί οποιαδήποτε απόφαση τον πρώτο λόγο έχουν οι ειδικοί επιστήμονες (εδώ οι φιλόσοφοι) που επικουρούμενοι από επιστήμονες άλλων κλάδων θα απαντήσουν στο ερώτημα “αυτή η πράξη που θέλουμε να προχωρήσουμε θα κάνει καλύτερη τη ζωή μας ή όχι;” και στη συνέχεια θα επιτρέψουν ή θα απαγορεύσουν τα επόμενα βήματα.

Προφανώς αυτή η ομάδα ηθικής θα ήταν χρήσιμη και σε άλλους τομείς της ζωής μας πέραν της έγκρισης των τεχνολογικών εφαρμογών, όπως στην οικονομία, στην εξωτερική πολιτική, την αγροτική πολιτική κ.α.

Το προφανές ερώτημα που αναδεικνύεται είναι, γιατί δεν υπάρχει πουθενά στον πλανήτη μια τόσο σημαντική ομάδα που θα οδηγεί με τις αποφάσεις της την ζωή μας προς το καλύτερο; Η απάντηση είναι προφανής και πηγάζει από την εσωτερικό κόσμο του ανθρώπου και την σύγκρουση της ηθικής με άλλες έμφυτες αδυναμίες: η ανθρώπινη απληστία για οικονομικά κέρδη είναι αυτή που οδηγεί τους κυβερνώντες στην λήψη αποφάσεων που δεν εξυπηρετούν το κοινό καλό αλλά προσωπικά συμφέροντα. **Αυτή η απάντηση είναι και το κλειδί που αποδεικνύει ότι κυρίως μέσω του δημοσίου τομέα, και ιδιαίτερα της κυβέρνησης, μπορεί να παραμεριστεί η ηθική, με όλες τις δυσάρεστες συνέπειες που ακολουθούν.**

9.4. Περίληψη

Αφού διαπιστώθηκε ότι η υπακοή σε κάποιες αρχές είναι θεμελιώδης, έγινε μια έρευνα σε διεθνές επίπεδο για το αν κάποιοι έχουν συγκεντρώσει σύνολα αρχών που θα πρέπει να υπακούει η τεχνητή νοημοσύνη. Τα ευρήματα αυτής της έρευνας παρουσιάζονται σε αυτό το κεφάλαιο. Επειδή ο όγκος των ευρημάτων είναι μεγάλος για να αναφερθεί, θα γίνει μια συγκέντρωση των κυριοτέρων αρχών που εμφανίζονται σαν κοινές διαπιστώσεις σε πολλές ομάδες. Αυτές οι αρχές είναι: να εξυπηρετεί το καλό, να σέβεται την ιδιωτικότητα, να σέβεται το δίκαιο, να είναι διαφανής, επεξηγηματική και δικαιολογημένη, να υπάρχει κάποιος που θα ευθύνεται για αυτή, να ωφελεί πολλούς και όχι λίγους, να μη προσβάλλονται θεμελιώδη δικαιώματα και αξίες,

Μετά την έρευνα που έγινε σχετικά με τις ομάδες που εξέδωσαν αρχές για τη ηθική διαπιστώθηκε ότι καμιά από αυτές τις ομάδες δεν έχει έναν ουσιαστικό ρόλο στην λήψη των τελικών αποφάσεων. Για αυτό το λόγο επινοήθηκε και προτάθηκε η δημιουργία μιας ομάδος ηθικής όπου η γνώμη της θα είναι δεσμευτική προς αυτούς που έχουν την τελική απόφαση για την αποδοχή ή όχι μιας τεχνολογικής εφαρμογής.

|

10. Η χρηματοδότηση της τεχνητής νοημοσύνης από τον δημόσιο τομέα και το ηθικό περίβλημά των κινητήριων δυνάμεων πίσω από την προώθηση της τεχνητής νοημοσύνης.

10.1. Εισαγωγή

Όπως συνηθίζεται σε όλη την εργασία έτσι και σε αυτό το κεφάλαιο θα πρέπει να θεμελιώσουμε την αναζήτηση σε ένα βασικό ερώτημα. Θα πρέπει λοιπόν να απαντηθεί το ερώτημα, ποιος είναι ο λόγος που προωθείται η τεχνητή νοημοσύνη και πως γίνεται αυτό; Σε αυτό λοιπόν το κεφάλαιο θα αναζητηθούν οι λόγοι που μας οδηγούν στην ανάπτυξη της τεχνητής νοημοσύνης αλλά και στη στήριξη που παρέχεται προς αυτή τη κατεύθυνση μέσω της χρηματοδότησης.

10.2. Η χρηματοδότηση της τεχνητής νοημοσύνης από τον δημόσιο τομέα

Η υλοποίηση μιας δημόσιας δομής όπου η τεχνητή νοημοσύνη θα χρησιμοποιείται ευρέως, προϋποθέτει και την παροχή κονδυλίων από το δημόσιο ώστε να δημιουργηθούν νέες υποδομές και να υλοποιηθούν νέες ιδέες. Η ΕΕ, μέλος της οποίας είναι και η χώρα μας, έχει ήδη προσανατολιστεί προς την κατεύθυνση αυτή και χρηματοδοτεί την τόνωση των επενδύσεων και τον καθορισμό δεοντολογικών κατευθυντήριων γραμμών. Η Ευρωπαϊκή Επιτροπή παρουσίασε σειρά μέτρων προκειμένου να τεθεί η τεχνητή νοημοσύνη στην υπηρεσία των Ευρωπαίων πολιτών και να ενισχυθεί η ανταγωνιστικότητα της Ευρώπης στον εν λόγω τομέα. Η ΕΕ (δημόσιος και ιδιωτικός τομέας) θα πρέπει να αυξήσει τις επενδύσεις για έρευνα και καινοτομία στον τομέα της τεχνητής νοημοσύνης κατά τουλάχιστον 20 δισ. EUR από σήμερα μέχρι το τέλος του 2020. Για να στηρίξει αυτές τις προσπάθειες, η Επιτροπή αυξάνει το ύψος της επένδυσής της σε 1,5 δισ. EUR για την περίοδο 2018-2020 στο πλαίσιο του προγράμματος έρευνας και καινοτομίας Ορίζοντας 2020. Για την επόμενη δεκαετία σχεδιάζει την χρηματοδότηση με 20 δισεκατομμύρια ανά έτος. Επιπλέον η χρηματοδότηση της τεχνητής νοημοσύνης θα ενταχθεί στο πρόγραμμα European Investment Fund. EUR 100 million που θα ενεργοποιηθεί το 2020. Ακόμη μια πηγή χρηματοδότησης θα είναι το European Innovation Council με το ποσό των 100 εκατομμυρίων για τα έτη 2019-2020 που θα δοθούν για έρευνα και καινοτομία κυρίως για ανθρωποκεντρική τεχνητή νοημοσύνη (Human centric AI). Επίσης προτίθεται να ιδρύσει γεωγραφικά διασκορπισμένα Κέντρα Έρευνας για την τεχνητή νοημοσύνη αλλά και Κέντρα Δοκιμών για τις εφαρμογές της, όπως επίσης και τεχνολογικά κέντρα hubs όπου θα συγκεντρώνονται οι ερευνητικές προσπάθειες σε κάθε μέλος.

Παράλληλα η ΕΕ τον Ιανουάριο του 2019 ίδρυσε την πρώτη συνεργατική (79 εταιρείες του χώρου της τεχνητής νοημοσύνης) πλατφόρμα τεχνητής νοημοσύνης AI4EU, για τη δημιουργία “εστιακού σημείου”, στο οποίο θα συγκεντρώνοντα “οι πόροι τεχνητής νοημοσύνης”, όπως αποθετήρια δεδομένων, υπολογιστική ισχύς,

εργαλεία, αλγόριθμοι. Το έργο αυτό θα χρηματοδοτηθεί με το ποσό των 25 εκατομμυρίων ευρώ.

Η χώρα μας παρότι δεν είχε υπογράψει την ιδρυτική «Διακήρυξη για τη Συνεργασία σε θέματα Τεχνητής Νοημοσύνης» (Declaration of Cooperation on Artificial Intelligence) της 10ης Απριλίου 2018, συμμετέχει πλέον στο παραπάνω πρόγραμμα(AI4EU).

Ηγετικό ρόλο στην τεχνητή νοημοσύνη επιθυμεί να έχει η χώρα του Καναδά, γιατί έχει ήδη δαπανήσει το μεγαλύτερο συγκριτικά ποσό για έρευνα γύρω από την τεχνητή νοημοσύνη μέσω του Καναδικού Ινστιτούτου για την Προηγμένη Έρευνα (Canadian Institute for Advanced Research CIFAR): 125 εκατομμύρια δολάρια για το 2017 και 1.7 δισεκατομμύρια για το 2018 και τα επόμενα πέντε έτη.

Εκτός βέβαια από την δημόσια χρηματοδότηση υπάρχουν και ιδιωτικοί φορείς που χρηματοδοτούν νέες πρωτοβουλίες για την τεχνητή νοημοσύνη που με τη σειρά τους μπορούν να φανούν χρήσιμες στο δημόσιο. Μια τέτοια επιχείρηση ‘‘άγγελος επενδυτής’’ είναι και η Zeroth.AI που χρηματοδοτεί νέες επιχειρήσεις με καινοτόμες ιδέες για την τεχνητή νοημοσύνη.

Επίσης στην χρηματοδότηση της τεχνητής νοημοσύνης μπορεί να ενταχθεί και το βραβείο των 5 εκατομμυρίων δολαρίων της XPRIZE, που δίνεται σε δράσεις που θα χρησιμοποιήσουν την τεχνητή νοημοσύνη για την αντιμετώπιση των μεγαλύτερων κοινωνικών προκλήσεων που αντιμετωπίζει η πλανήτης.

10.3. Οι κινητήριες δυνάμεις πίσω από την προώθηση της τεχνητής νοημοσύνης και το ηθικό περίβλημά τους

Αναφέρθηκε σε προηγούμενο κεφάλαιο ότι τα τελευταία χρόνια η Τεχνητή νοημοσύνη προωθείται από τον δημόσιο τομέα μέσω της χρηματοδότησης σχετικών δράσεων. Η χρηματοδότηση στρέφεται είτε προς το ίδιο το δημόσιο είτε προς ιδιωτικές πρωτοβουλίες. Αυτό κυρίως φαίνεται στην απόφαση της ΕΕ να δαπανήσει το ποσό των 25 δισεκατομμυρίων ευρώ για την έρευνα γύρω από την τεχνητή νοημοσύνη. Θα εξεταστεί στη συνέχεια ποια είναι τα πιθανά κίνητρα πίσω από μια απόφαση να ασχοληθεί κάποιος με την τεχνητή νοημοσύνη είτε σαν άτομο είτε σαν ένας ευρύτερος φορέας; Το ερώτημα αυτό ενδιαφέρει την παρούσα έρευνα γιατί εισχωρεί στην σφαίρα της ηθικής του κάθε ατόμου. Το άτομο βρίσκεται μπροστά σε ηθικό δίλλημα για το ποια απόφαση πρέπει να πάρει ζυγίζοντας τις αδυναμίες του με τις ηθικές αρχές του. Έτσι διακρίνουμε τα παρακάτω τέσσερα **κίνητρα** για να προχωρήσει κάποιος σε μια νέα τεχνολογική εφαρμογή:

A. Η ανθρώπινη τάση για νέες ανακαλύψεις

Στην μακρόχρονη πορεία των 200.000 ετών του Homo Sapiens υπήρξε μια μεγάλη τεχνολογική βελτίωση, και μοναδική ανάμεσα στα άλλα έμβια όντα. Πολλές

από τις νέες ανακαλύψεις δεν είχαν κανένα οικονομικό σκοπό και απλά έρχονται σαν αποτέλεσμα μιας εσωτερικής ευγενούς αναζήτησης του ωραίου.

B. Η ανθρώπινη τάση για διευκόλυνση της ζωής

Άλλες πάλι τεχνολογικές ανακαλύψεις ήρθαν σαν μια ανάγκη να αντικατασταθούν επίπονες για τον άνθρωπο εργασίες με αυτοματοποιημένες διαδικασίες, ώστε να διευκολυνθεί η ζωή. Σαν τέτοιες εργασίες διεθνώς χαρακτηρίζονται με τα τρία d: dull, dirty, dangerous. (Patrick Lin, Keith Abney, George Bekey. 2010). Με την αύξηση των δυνατοτήτων των μηχανών, στις εργασίες αυτές προστέθηκαν και εργασίες που οι μηχανές θα μπορούσαν να κάνουν καλύτερα π.χ. με μεγαλύτερη ακρίβεια, όπως μια ρομποτική εγχείρηση.

Γ. Η εξοικονόμηση και η ανθρώπινη οικονομική απληστία

Η εξοικονόμηση χρημάτων ή η αύξηση των κερδών είναι στην σύγχρονη εποχή ένας βασικός σκοπός των περισσότερων ανθρώπων. Τα κίνητρα πίσω από αυτή την επιδίωξη είναι: είτε ευγενή π.χ. όταν κάποιος βρίσκεται σε οικονομική στένωση προσπαθεί να εξοικονομήσει πόρους για να ανταπεξέλθει, είτε είναι απεχθή π.χ. όταν κάποιος με τεράστια οικονομική δύναμη απολύοντας εργαζομένους, αντικαθιστά ανθρώπινη εργασία με αυτοματοποιημένες διαδικασίες, με σκοπό να δημιουργήσει υπερκέρδη.

Δ. Η ανθρώπινη τάση προς το κακό

Αν και υπήρξε τεράστια βελτίωση στο επίπεδο της διαβίωσης του ανθρώπου, σε ένα άλλο επίπεδο αυτό της βιολογικής προόδου εικάζεται από πολλούς ότι δεν υπήρξε μεγάλη αλλαγή (J.Ayala. 2017). Ένα από τα υποσυνείδητα χαρακτηριστικά του ανθρώπου είναι και αυτό της επιδίωξης της επιβολής του είτε στους άλλους είτε στο περιβάλλον του. Υπάρχουν πολλά παραδείγματα που δείχνουν ότι αυτή η διαπίστωση είναι αληθινή. Ένα τέτοιο έντονο παράδειγμα είναι η πραγματικότητα της χρήσης των διηπειρωτικών όπλων μαζικής καταστροφής που είναι μια μεγάλη απόδειξη ύπαρξης αυτού του χαρακτηριστικού. Επίσης η μη αποδοχή από τα ηγετικά μέλη του ΟΗΕ της απαγόρευσης των σημερινών αυτόνομων οπλικών συστημάτων, μαρτυρά ότι σε βιολογικό επίπεδο οι αλλαγές σε διάστημα χιλιάδων ετών δεν ήταν μεγάλες. Η πιο λογική εξήγηση για την ύπαρξη αυτού του χαρακτηριστικού στον άνθρωπο είναι ότι πρόκειται για ένα δευτερογενές συναίσθημα το οποίο πηγάζει από την απληστία, τον φθόνο, και στους αναξιοπαθούντες από τη βίωση των ανισοτήτων (Boris Gershman. 2015). Εικάζεται ότι πρόκειται για μια υποσυνείδητη και άρα μη ελεγχόμενη από τον άνθρωπο κατάσταση, η οποία παραμένει σε μεγάλο βαθμό άγνωστη (James Hollis. 2007).

Μετά την παραπάνω αναφορά των κινήτρων πίσω από την προώθηση μιας τεχνολογικής ανακάλυψης, φτάνει να προσαρμόσουμε κατάλληλα το ερώτημα για την τεχνητή νοημοσύνη: **Γιατί ο δημόσιος τομέας προωθεί την τεχνητή νοημοσύνη, είναι τα κίνητρά του ευγενή ή όχι;** Είναι προφανές ότι οι απαντήσεις

που μπορούν να δοθούν είναι αντικρουόμενες μιας και υπάρχουν εφαρμογές που εξυπηρετούν το κακό όπως τα σημερινά αυτόνομα οπλικά συστήματα ή τα αμφιλεγόμενα ολοκληρωτικά συστήματα κοινωνικής βαθμολόγησης, αλλά και εφαρμογές που εξυπηρετούν ξεκάθαρα το κοινό καλό όπως τα συστήματα πρόβλεψης του καιρού ή διάγνωσης των ασθενειών. Το ποια πλευρά θα επικρατήσει θα καθοριστεί από την σύνεση που θα δείξουν οι άνθρωποι που έχουν στα χέρια τους τις πολιτικές αποφάσεις.

10.4. Περίληψη

Στο κεφάλαιο αυτό επιχειρήθηκε να γίνει μια χαρτογράφηση του χώρου της ηθικής των κινήτρων πίσω από την προώθηση της τεχνητής νοημοσύνης. Αρχικά τεκμηριώνεται η άποψη ότι πράγματι η ΕΕ προωθεί την τεχνητή νοημοσύνη δίνοντας μεγάλη χρηματοδότηση σε ιδιωτικές και δημόσιες δράσεις. Έπειτα γίνεται αναφορά σε τέσσερα κίνητρα που θα μπορούσαν να είναι υπεύθυνα για την προώθηση της τεχνητής νοημοσύνης. Τα κίνητρα αυτά είναι: η ανθρώπινη τάση για νέες ανακαλύψεις, η ανθρώπινη τάση για διευκόλυνση της ζωής, η ανθρώπινη τάση για εξοικονόμηση (ευγενής ή όχι), η ανθρώπινη τάση για το κακό. Επειδή ακόμη η τεχνητή νοημοσύνη δεν έχει αναπτυχθεί πλήρως δεν μπορούμε να πούμε με σιγουριά ποια είναι τα πραγματικά κίνητρα, και το μόνο που μπορούμε να κάνουμε είναι να ευελπιστούμε ότι η καλή χρήση της θα κυριαρχήσει στο μέλλον.

11. Πως πρέπει να προετοιμαστεί ο δημόσιος τομέας και οι πολίτες ώστε να υποδεχθούν την τεχνητή νοημοσύνη;

11.1. Εισαγωγή

Η χώρα μας δεν είναι ανάμεσα σε αυτές τις χώρες που πρωταγωνιστούν στην ανάπτυξη της τεχνητής νοημοσύνης. Αυτό βέβαια εκτός από τα μειονεκτήματα που δίνει, όπως π.χ. ο τεχνολογικός αναχρονισμός, έχει και ένα βασικό πλεονέκτημα: δίνει ένα σημαντικό περιθώριο χρόνου ώστε πολίτες και δημόσιο να προετοιμαστούν για να υποδεχτούν την έλευση της τεχνητής νοημοσύνης. Όσον αφορά το δημόσιο η προετοιμασία αφορά στην δημιουργία των απαραίτητων δομών και την εξακρίβωση των σκοπών του, και όσον αφορά τους πολίτες την ενημέρωση των πολιτών για τα πλεονεκτήματα και μειονεκτήματα.

11.2. Πως πρέπει να προετοιμαστεί ο δημόσιος τομέας για να υποδεχτεί την τεχνητή νοημοσύνη;

Η υιοθέτηση των εφαρμογών της τεχνητής νοημοσύνης από τον δημόσιο τομέα απαιτεί να γίνει μια προσαρμογή του υπάρχοντος συστήματος σε νέες καταστάσεις και νέες συνθήκες. Η προσαρμογή αυτή προϋποθέτει την αλλαγή σε κάποιες λειτουργίες του δημοσίου τομέα.

Αρχικά πρέπει να καθιερωθεί ένα νέο **στρατηγικό πλάνο**. Το πλάνο αυτό θα περιλαμβάνει τα εξής:

-Αναγνώριση όλων των εσωτερικών διαδικασιών και όλων των δημοσίων υπηρεσιών που θα μπορούσαν να επωφεληθούν από την τεχνητή νοημοσύνη.

-Αναγνώριση όλων των υπηρεσιών που θα μπορούσαν να παρέχονται προσωποποιημένα σε κάθε πολίτη ξεχωριστά. αλλά και σε είδη επιχειρήσεων

-Υιοθέτηση κοινών μοντέλων και κοινών προτύπων λειτουργίας σε όλες τις διαδικασίες.

-Συγκέντρωση ποιοτικών δεδομένων και σε μεγάλες ποσότητες τα οποία θα τροφοδοτήσουν μια μηχανή λήψης αποφάσεων.

-Ξεκίνηση της εφαρμογής της τεχνητής νοημοσύνης αρχικά σε κάποια μικρή διαδικασία ώστε να λειτουργήσει πιλοτικά και να εντοπιστούν πιθανά λάθη. Δεν πρέπει να φοβίζει η αποτυχία αλλά να είναι πηγή εμπειριών και μάθησης.

-Καθιέρωση κριτηρίων αξιολόγησης της απόδοσης ενός συστήματος τεχνητής νοημοσύνης, ώστε να απαντηθεί το ερώτημα ποιο είναι το επιθυμητό αποτέλεσμα και πως αυτό θα μετράται.

-Πρόσληψη ειδικών επιστημόνων που θα βοηθήσουν στην εφαρμογή των ανωτέρω. Οι επιστήμονες αυτοί θα είναι ειδικοί στην τεχνητή νοημοσύνη ώστε να εγκαταστήσουν και να επιτηρούν το σύστημα, ειδικοί γλωσσολόγοι ώστε να βοηθήσουν στην ιδιαιτερότητα της Ελληνικής γλώσσας ή άλλων ξένων γλωσσών, ειδικοί στην στατιστική ώστε να αξιολογούν τα αποτελέσματα, ειδικοί στην φιλοσοφία και στην ψυχολογία ώστε να εντοπίζουν αλλαγές στην συμπεριφορά των ωφελούμενων πολιτών.

-Επιμόρφωση των υφισταμένων υπαλλήλων στις καινούργιες συνθήκες με παρακολούθηση ειδικών σεμιναρίων.

-Επένδυση εκ νέου σε όσες εφαρμογές λειτούργησαν καλά και εγκατάλειψη ή βελτίωση όσων δεν λειτούργησαν.

Για την προετοιμασία ενός κράτους ενόψει της χρήσης της τεχνητής νοημοσύνης μια μελέτη ακόμη που έχει εκδοθεί και αναφέρεται στο Καναδικό κράτος είναι αυτή των Avin Gaon και Ian Stedman (Avin Gaon, Ian Stedman. 2019). Στην μελέτη αυτή οι συγγραφείς επισημαίνουν την ανάγκη ένα κράτος να μην είναι θεατής των εξελίξεων αλλά να λάβει ουσιαστικό ρόλο στην υιοθέτηση της τεχνητής νοημοσύνης χαράσσοντας την **κεντρική πολιτική** που θα ακολουθηθεί, αναλύοντας την υπάρχουσα κατάσταση, ψηφίζοντας τους κατάλληλους νόμους, και δίνοντας την κατάλληλη εκπαίδευση στις επόμενες γενιές. Οι νόμοι αυτοί θα προκύψουν αφού ληφθούν υπόψη κάποιες ανησυχίες όπως η πολιορκία της ιδιωτικότητας, η απώλεια στις θέσεις εργασίας, η επιρροή στην εμπιστοσύνη έναντι του δημοσίου, η γιγάντωση του ιδιωτικού τομέα εις βάρος του δημοσίου καλού, η παρωχημένη νομοθεσία κ.αλ. Αυτή η αντιμετώπιση που προτείνεται είναι η ακριβώς αντίθετη από την “πολιτική” που ακολουθείται στην χώρα μας, ότι δηλαδή βρισκόμαστε πάντοτε θεατές των εξελίξεων και λαμβάνουμε μέτρα θεραπείας.

11.3. Πως πρέπει να προετοιμαστούν οι πολίτες για την υποδοχή της τεχνητής νοημοσύνης στον δημόσιο τομέα;

Η τεχνητή νοημοσύνη μπορεί να διαμορφώσει ένα νέο κοινωνικό περιβάλλον και διαφορετικούς τρόπους συμβίωσης που αλλάζουν την έννοια της καθημερινής ζωής. Πολλοί είναι αυτοί που παρομοιάζουν την έλευση της τεχνητής νοημοσύνης σαν μια νέα “βιομηχανική επανάσταση” που θα φέρει ραγδαίες αλλαγές στο κοινωνικό τοπίο.

Επειδή η τεχνητή νοημοσύνη είναι ήδη παρούσα σε πολλές χώρες του εξωτερικού θα πρέπει και στην χώρα μας να γίνει το συντομότερο ένα στρατηγικό πλάνο για την υποδοχή αυτής της τεχνολογίας που θα περιλαμβάνει και την προετοιμασία των πολιτών στο νέο περιβάλλον που θα διαμορφωθεί.

Αρχικά θα πρέπει να τηρηθούν οι επιταγές του συντάγματος και η διαφύλαξη της δημοκρατίας, δηλαδή να ρωτηθούν οι πολίτες αν επιθυμούν ή όχι την νέα

τεχνολογία, αφού πρώτα ενημερωθούν για τα οφέλη και τις ζημιές που θα προκύψουν; Έτσι λοιπόν για παράδειγμα, θα πρέπει να ερωτηθούν οι πολίτες αν στο όνομα της εξοικονόμησης δημοσίων δαπανών θέλουν να συναναστρέφονται και να ομιλούν με chatbots στους δημοσίους χώρους ή να σκανάρονται τα πρόσωπά τους και τα χαρακτηριστικά τους ή ακόμη και να προβλέπεται ο ψυχικός κόσμος τους, και σε ποιο βαθμό;

Προφανώς για να γίνει αυτό θα πρέπει πρωτίστως να γίνει μια εμπειριστατωμένη μελέτη από ειδικούς επιστήμονες που, από την μια, που θα λείπει ποιες είναι οι επιπτώσεις στον ψυχισμό του πολίτη και ποια είναι τα άλλα αρνητικά αποτελέσματα (π.χ. τεχνολογική ανεργία) και από την άλλη, ποια θα είναι τα οφέλη (π.χ. εξοικονόμηση χρημάτων, καλύτερευση συνθηκών εργασίας για τους ήδη εργαζομένους στο δημόσιο). Συγκριμένα θα πρέπει να αναλυθούν τα παρακάτω:

-Ποιες αξίες θα επηρεαστούν από την τεχνητή νοημοσύνη (ισότητα, δικαιοσύνη, ανθρώπινη αξιοπρέπεια, ακεραιότητα του προσώπου, ιδιωτικότητα, κ.αλ.);

-Ποιες ομάδες πολιτών κυρίως θα επηρεαστούν (εργάτες, αγρότες, οδηγοί, δημόσιοι υπάλληλοι που παρέχουν πληροφορίες και παράγουν έγγραφα, εργαζόμενοι στον τομέα υγείας, καταναλωτές, αυτοί που παίρνουν αποφάσεις, στρατιωτικοί, κ.αλ.);






-Τι είδους αντίκτυπος θα υπάρξει, θετικός, αρνητικός, και οι δύο; Στα θετικά θα μπορούσαν να προσμετρηθούν όλα τα επιτεύγματα που έχουν αναφερθεί μέχρι τώρα στην διατριβή και περιλαμβάνουν: καλύτερες διαγνώσεις στην υγεία, ασφάλεια στα αεροδρόμια, καλύτερη κυκλοφορία οχημάτων και γρηγορότερες μεταφορές, καλύτερες δημόσιες υπηρεσίες, έγκαιρες και έγκυρες προγνώσεις καιρού κ.αλ. Στα αρνητικά θα μπορούσαν να προσμετρηθούν: αποδοχή μεγαλύτερου επιπέδου στην ανεργία, αποδοχή συνεχούς επιτήρησης, αποδοχή στην δημιουργία προφίλ μέσω της επιτήρησης, αποδοχή απωλειών από ανθρωποειδή στρατιωτικά ρομπότ, αποδοχή κινδύνου απώλειας δεδομένων, αποδοχή τυχόν διακρίσεων και ανισοτήτων, αποδοχή κινδύνου μιας διεθνούς ολοκληρωτικής διακυβέρνησης, αποδοχή μεγαλύτερων ενεργειακών καταναλώσεων κ.αλ. Στα αμφιλεγόμενα θα μπορούσαν να προσμετρηθούν: αλλαγή στις ανθρώπινες σχέσεις μιας και θα υπάρχει συναναστροφή με ρομπότ, αλλαγή στο νομικό πλαίσιο ώστε η συναναστροφή με ένα αντικείμενο να καλυφθεί νομικά, αποδοχή μιας νέας ηθικής κατάστασης.


Μια προσπάθεια προς τη κατεύθυνση της μελέτης του αντικτύπου έγινε από την εταιρεία Sienna (http://www.sienna-project.eu/digitalAssets/721/c_721882-1_1-k_sienna-infographic-ai-robotics-pdf.pdf) και στο παρακάτω σχήμα (Σχήμα 5. AI's impact.) παρουσιάζονται εν συντομία τα ευρήματά της.


Σχήμα 5. AI's impact (πηγή: http://www.sienna-project.eu/digitalAssets/721/c_721882-1_1-k_sienna-infographic-ai-robotics-pdf.pdf).

AI & ROBOTICS

AI & robotics will have a significant impact on society, values & human rights

EFFECT ON VALUES & PEOPLE		TYPES OF IMPACTS		
				
VALUES AFFECTED	PEOPLE HIGHLY AFFECTED	POSITIVE	DUAL	NEGATIVE
<ul style="list-style-type: none"> • Accountability • Autonomy • Equality • Fairness • Human dignity • Justice • Integrity of the person • Privacy • Security • Solidarity 	<ul style="list-style-type: none"> • Consumers • Disfavoured or 'excluded' people • First responders • Healthcare providers • Inhabitants of poor countries • Patients • Recipients of insurance & social benefits • Regulators & policy-makers • Tenants (e.g. minority ethnic groups) • The elderly • Workers 	<ul style="list-style-type: none"> • Improve medical diagnostics • Advance cybersecurity • Improve decision-making via data analysis • Enhance healthcare • Improve elder care • Advance language translation • Improve voting security • Tenants (e.g. minority ethnic groups) • Introduce new forms of co-operation & inclusion • Reduce repetitive tasks 	<ul style="list-style-type: none"> • Alter human relations • Alter legal frameworks • Alter moral conceptions • Change understandings of personhood • Create unintended consequences • Intensify big data analytics • Escalate surveillance • Increase profiling • Increase leisure 	<ul style="list-style-type: none"> • Harm/threat of harm from autonomous weapons • Loss of control over privacy & personal data • Bias & discrimination • Cyberwarfare • Increase class/wealth domination • Diminish media pluralism • Greater energy consumption





We will help address some of these issues by:

- Reviewing the legal, ethical, social & economic impacts of AI & robotics
- Connecting with citizens to understand concerns via surveys in 11 countries & panels in 5 countries
- Consulting experts and discussing ethical & human rights issues
- Developing ethical codes & protocols

Εφόσον εντοπισθούν οι ζημίες που τυχόν θα υπάρξουν θα πρέπει να οργανωθεί ένα **σχέδιο αντιμετώπισης** τους. Από την αντίθετη πλευρά, αν εντοπισθούν ωφέληματα θα πρέπει να υπολογισθεί **ποιος θα καρπωθεί τα κέρδη** από την υιοθέτηση της τεχνητής νοημοσύνης.

Συγκεκριμένα θα πρέπει το δημόσιο να οργανώσει δράσεις ενημέρωσης των πολιτών και δράσεις θεραπείας των προβλημάτων.

Όσον αφορά την ενημέρωση των πολιτών, θα μπορούσε να γίνει μια ενιαία καμπάνια σε όλη την χώρα, όπου μια ομάδα ειδικών θα μιλήσει στους πολίτες για τις αλλαγές που πρόκειται να συμβούν, και για την στάση που πρέπει να κρατούν οι πολίτες απέναντι στην τεχνολογία αυτή. Επίσης μπορεί να διαμοιραστεί ειδικό έντυπο υλικό για ενημερωθούν καλύτερα.

Όσον αφορά τη Θεραπεία τυχόν προβλημάτων αυτή θα πρέπει να στραφεί σε πολλά μέτωπα:

Είναι σίγουρο ότι κάθε καινοτόμος ενέργεια αλλάζει την κοινωνική ζωή των πολιτών και προκαλεί αλλαγές κυρίως στον **ψυχισμό** των πολιτών (Spyros Makridakis.

2017). Αυτό θα πρέπει να αντιμετωπιστεί από ειδικούς επιστήμονες του χώρου της ψυχικής υγείας. Η κοινωνική **συμβίωση** ανθρώπων και μηχανών τεχνητής νοημοσύνης είναι ακόμη ένα φανταστικό σενάριο που βρίσκεται στο αρχικό επίπεδο μελετών. Η σημερινή αντιμετώπιση ρομπότ από ανθρώπους δεν έχει ξεπεράσει στερεότυπα της εποχής που οι μηχανές είναι σκλάβοι των ανθρώπων, και άρα μπορούν να δέχονται προσβολές και κακή αντιμετώπιση (Antonella De Angeli.2009). Όμως όταν ίσως στο μέλλον η παρουσία των ρομπότ γίνεται περισσότερο πυκνή και αισθητή και η χρησιμότητα τους απαραίτητη, τότε η κοινωνική αντίληψη προς αυτά θα μεταβληθεί. Είναι ακόμη άγνωστο προς ποια κατεύθυνση θα μεταβληθεί η ανθρώπινη συμπεριφορά, φιλική ή εχθρική.

Η παραπάνω πρόκληση της συμβίωσης με ρομπότ αποκτά ιδιαίτερο ενδιαφέρον αν ευοδωθούν οι μελέτες γύρω από τα βιο-τεχνικά ρομπότ (Bio-Syncretic Robots). Αυτά τα ρομπότ θα αποτελούνται από βιολογικό και μη βιολογικό υλικό. Τα ρομπότ αυτά εικάζεται (Chuang Zhanga. Wenxue Wanga, Ning Xic, Yuechao Wanga, Lianqing Liu. 2018) ότι θα έχουν την ικανότητα αυτοεπιδιόρθωσης και της αυτοσυναρμολόγησης. Μέχρι σήμερα, η γνωστή στο κοινό πρόοδος, δεν έχει καταφέρει να δείξει ένα ολοκληρωμένο ρομπότ με ανθρωποειδή μορφή, παρά μόνο έχει να δείξει κάποια επιμέρους επιτεύγματα, όπως κίνηση με ιστούς, βακτήρια που κινούνται με κατανάλωση γλυκόζης κ.αλ. Είναι πολύ νωρίς ακόμη για να προσδιορίσουμε ποια θα είναι η αντιμετώπιση από τον άνθρωπο προς ένα ρομπότ με βιολογικά υλικά. Θα εκλαμβάνεται σαν ένα έμβιο όν ή όχι, και ποια θα είναι η επίδρασή του στην ανθρώπινη ψυχολογία; Μια προσπάθεια προς αυτή την κατεύθυνση έγινε από το ίδρυμα Lifeboat Foundation που είναι μια μη κερδοσκοπική οργάνωση που σκοπό έχει να βοηθήσει την ανθρωπότητα να αποφύγει ρίσκα και κακές χρήσεις της τεχνολογίας, όπως ρίσκα που προέρχονται από την γενετική μηχανική, την ναυτοτεχνολογία, τα ρομπότ και την τεχνητή νοημοσύνη. Σε μια από τις αναλύσεις του στο πρόγραμμα AIShield για την προστασία εναντίον μιας εχθρικής τεχνητής νοημοσύνης, αναφέρει ότι μια επίδραση στην **κοινωνική συμπεριφορά** που μπορεί να προκληθεί από την συμβίωση με τα ρομπότ είναι η μείωση της κοινωνικής ευαισθησίας λόγω της συναναστροφής με άψυχα όντα, με αποτέλεσμα την κοινωνική κατάρρευση. Επίσης επισημαίνεται ο κίνδυνος, όταν η πλειοψηφία των εργασιών θα γίνεται από ρομπότ, να δημιουργηθεί ένας άνθρωπος που δεν χρειάζεται να κάνει κάτι χρήσιμο, κάτι που είναι η απαρχή μιας εξαφάνισης (πηγή <https://lifeboat.com/ex/aishield>).

Έχει εκφραστεί η άποψη ότι αν η τεχνητή νοημοσύνη παράξει αποτελέσματα θαυμαστά για τον άνθρωπο αυτό θα λειτουργήσει σαν αίσθημα κατωτερότητας απέναντι στις μηχανές και θα αποτελέσει ένα “τέλος” για την φιλοσοφική και υπαρξιακή θεώρηση και την ανθρώπινη φαντασία άρα και την αρχή ενός γενικότερου τέλους (Michael L. Anderson. 2005). Η άποψη αυτές αναφέρονται μόνο για την τήρηση μιας αμερόληπτης στάσης απέναντι στην καταγραφή όλων των απόψεων και σε μεγάλο βαθμό δεν υιοθετούνται γιατί η τεχνολογική πραγματικότητα όπως την βιώνουμε καθημερινά δεν έχει δείξει καμιά υλοποίηση τέτοιων σεναρίων.

Ακόμη δυο σημαντικά προβλήματα που πρέπει να αντιμετωπιστούν άμεσα είναι αυτά της **ανισότητας** και της **τεχνολογικής ανεργίας**.

-Αυτό που πρέπει να ληφθεί υπόψη είναι η δημιουργία μιας επιπλέον ανισότητας (Andrew Simms. 2018) στην οικονομική ανάπτυξη των χωρών του κόσμου. Η τεχνητή νοημοσύνη θα δώσει μια έκρηξη παραγωγικότητας στις ήδη ανεπτυγμένες χώρες με αποτέλεσμα το χάσμα με τις υποανάπτυκτες να μεγαλώσει σε τεράστιο βαθμό. Θα πρέπει λοιπόν σε επίπεδο διεθνούς στρατηγικής τα επιτεύγματα της τεχνητής νοημοσύνης να διασκορπιστούν γεωγραφικά με μια ισόρροπη και δίκαιη μοιρασιά. Η ένταση των ανισοτήτων μπορεί να εντοπιστεί ακόμη και σε εθνικό επίπεδο, μιας και στην χώρα μας υπάρχει ήδη διεύρυνση της φτώχειας (σχεδόν 3,5 εκ. σύμφωνα με την ΕΛΣΤΑΤ http://www.statistics.gr/documents/20181/7146186/LivingConditionsInGreece_1118.pdf/a7883ee9-6ce0-4c35-b70c-b5b8618df3b3) και προφανώς στο μέλλον, λόγω οικονομικών προβλημάτων, τα ποσοστά αναφαρβητισμού θα μεγαλώσουν και ο κοινωνικός και οικονομικός αποκλεισμός θα διογκωθεί.

-Είναι ιστορικά διαπιστωμένο και επιστημονικά εξακριβωμένο ότι κάθε τεχνολογική επανάσταση φέρνει ένα τελικό αρνητικό ισοζύγιο στην εργασία, δηλαδή αυτοί που συνολικά εργάζονται μετά την τεχνολογική επανάσταση είναι λιγότεροι από αυτούς που εργαζόταν πριν αυτή να συμβεί. Αλλάζει επίσης η ποιότητα εργασίας προς το καλύτερο, και ταυτόχρονα γίνονται πολλές μετακινήσεις εργατικού δυναμικού μεταξύ κλάδων της οικονομίας. Το φαινόμενο αυτό είναι ευρύτερα γνωστό σαν **Τεχνολογική ανεργία** (Young Joon, Kima Kyungsoo, Kimb Su, Kyoung Lee. 2017), και επιβεβαιώνει άλλωστε και τον σκοπό μιας τεχνολογικής καινοτομίας που είναι η απαλλαγή του ανθρώπου από το βάρος μιας οποιαδήποτε εργασίας. Πραγματικά η απαλλαγή από μια κουραστική εργασία είναι ο σκοπός του ανθρώπου, εφόσον βέβαια αυτή η πράξη δεν τον στέλνει στην ανεργία. Αλλά ακόμη και αν τον στέλνει θα πρέπει να βρεθεί ένας τρόπος ώστε να έχει δικαίωμα στα κέρδη που έφερε η αντικατάστασή του από μια μηχανή.

Έρευνα μιας από τις πιο γνωστές εταιρείες στον χώρο, της McKinsey Global (https://www.mckinsey.com/~media/mckinsey/featured%20insights/future%20of%20organizations/what%20the%20future%20of%20work%20will%20mean%20for%20jobs%20skills%20and%20wages/mgi%20jobs%20lost-jobs%20gained_report_december%202017.ashx) δείχνει ότι το 50% των υπαρχουσών εργασιών έχουν την δυνατότητα να αυτοματοποιηθούν στα επόμενα χρόνια. Το ίδιο επιβεβαιώνει η έρευνα του πανεπιστημίου της Οξφόρδης (Carl Benedikt Frey, Michael A. Osborne. 2013) ότι δηλαδή το 47% της εργασίας στις ΗΠΑ θα χαθεί λόγω της αυτοματοποίησης και των υπολογιστών.

Μια ακόμη έρευνα που επιβεβαιώνει ότι η τεχνητή νοημοσύνη θα προκαλέσει ανεργία αλλά και άνιση κατανομή του εισοδήματος από την τεχνολογία είναι και αυτή του Λευκού Οίκου (

<https://www.whitehouse.gov/sites/whitehouse.gov/files/images/EMBARGOED%20AI%20Economy%20Report.pdf>. Σελ 14).

Επίσης το γεγονός ότι τον Ιανουάριο του 2018 στο γνωστό τεστ του πανεπιστημίου του Stanford (Stanford Question Answering Dataset), όπου άνθρωποι και λογισμικό τεχνητής νοημοσύνης απαντούν σε ερωτήσεις κατανόησης κειμένου, για πρώτη φορά το λογισμικό των εταιρειών Alibaba (BABA) και Microsoft (MSFT) νίκησε τον άνθρωπο, αυτό θα σημαίνει στο μέλλον ότι πολλές θέσεις εξυπηρέτησης του κοινού θα μπορούν να αντικατασταθούν με μηχανές.

Για να αντιμετωπιστεί η τεχνολογική ανεργία θα πρέπει να υπάρξει ένας μακροπρόθεσμος σχεδιασμός για την εργασία, ώστε να δημιουργηθεί εκ των προτέρων μια δομή που θα υποδέχεται ανθρώπους που θα χάσουν την εργασία τους από νέες τεχνολογίες. Διεθνώς υπάρχει η σκέψη να υπάρξει ειδικός φόρος στα ρομπότ, τα έσοδα του οποίου θα δίνονται σε ζημιωθέντες από τον αυτοματισμό. Είναι προφανές ότι στον δημόσιο τομέα οποιαδήποτε εξοικονόμηση χρημάτων μπορεί εύκολα να διαμοιραστεί σαν κοινωνικό μέρισμα από το δημοσιονομικό πλεόνασμα που θα προκύψει. Όμως τι γίνεται με τον ιδιωτικό τομέα; Δυστυχώς στον ιδιωτικό τομέα η μείωση του εργατικού κόστους είναι ένας στόχος και ταυτόχρονα η αύξηση των κερδών δεν είναι περιουσία του δημοσίου που μπορεί να μοιραστεί, αλλά περιουσία του ιδιοκτήτη που πρέπει να μην υπερφορολογείται ώστε να λειτουργεί σαν κίνητρο για την επιχειρηματικότητα. Επειδή ο ιδιωτικός τομέας είναι το συντριπτικό ποσοστό μιας εθνικής οικονομίας, η ανεργία και συνεπώς η φτώχεια που θα έρθει λόγω της τεχνητής νοημοσύνης από εκεί είναι δυνητικά τεράστια και δύσκολο να αντιμετωπιστεί αν δεν υπάρξει μια σκέψη διαμερισμού των υπερκερδών που θα προκύψουν από την εφαρμογή της τεχνητής νοημοσύνης. Αν και το αντικείμενο της εργασίας είναι ο δημόσιος τομέας, βλέπουμε ότι και οι δράσεις του ιδιωτικού τομέα πρέπει να εξουσιάζονται από το δημόσιο γιατί έχουν σοβαρό αντίκτυπο στο δημόσιο. Η σκέψη αυτή προϋποθέτει μια συνολική αλλαγή στην οικονομική θεώρηση της λειτουργίας ενός κράτους. Αν θεωρήσουμε ότι στο μέλλον η αυτοματοποίηση θα αγγίξει το μέγιστο δυνατό όριο αντικατάστασης της ανθρώπινης εργασίας τότε θα αναρωτηθεί κάποιος αν κάθε άνεργος θα έχει δικαίωμα στο να απολαμβάνει αγαθά για τα οποία εργάστηκαν μηχανές είτε ιδιωτικές είτε του δημοσίου; Προς αυτή τη κατεύθυνση της εξισορρόπησης των εισοδημάτων είναι και η τελευταία πρόταση του προέδρου της Επιτροπής για την προετοιμασία της ΕΕ ώστε να καθιερώσει το ελάχιστο εγγυημένο εισόδημα για την μείωση των κοινωνικών ανισοτήτων και της φτώχειας. Μάλιστα ονόμασε το κατώτερο εγγυημένο εισόδημα σαν ένα “επίπεδο υπερηφάνειας” που πρέπει να γίνει σεβαστό. Βέβαια το γεγονός ότι μετά από σαράντα χρόνια περίπου της χώρας μας στην ΕΕ δεν έχει γίνει κάτι προς αυτή τη κατεύθυνση, δείχνει ότι οι προτάσεις αυτές είναι κατευναστικές και τελικά περιπαιχτικές.

Το σημαντικότερο ίσως πρόβλημα που μπορεί να κρύβει μια τεχνολογική εφαρμογή τεχνητής νοημοσύνης από το δημόσιο τομέα είναι η **κακή χρήση** της

δύναμης που συγκεντρώνει μια κυβέρνηση στα χέρια της. Θα αναλυθεί παρακάτω πως αυτό μπορεί να συμβεί:

Κάποιες από τις χρήσεις της τεχνητής νοημοσύνης θα είναι και η απολαβή δημοσίων υπηρεσιών αλλά και η επικοινωνία των πολιτών με την κυβέρνηση μέσω των chatbots (Aggeliki Androutsopoulou, Nikos Karacapilidis, Euripidis Loukisa, Yannis Charalabidis. 2018). Αυτό προϋποθέτει σε πολλές περιπτώσεις την παροχή από τον πολίτη κάποιων προσωπικών δεδομένων στην εφαρμογή της τεχνητής νοημοσύνης. Αυτά τα δεδομένα θα συλλέγονται σε ένα συγκεκριμένο χώρο που επιλέγει η κυβέρνηση. Ένα από τα μεγαλύτερα προβλήματα που κρύβεται πίσω από την τεχνητή νοημοσύνη και μπορεί να δημιουργηθεί μέσω του δημοσίου τομέα είναι ίσως η **συγκέντρωση της δύναμης** που δίνει η κατοχή των πληροφοριών αυτών. Στη συνέχεια όποιος έχει στα χέρια του αυτή τη δύναμη μπορεί να την χρησιμοποιήσει είτε για καλό είτε για κακό σκοπό. Η κακή χρήση είναι ένας κίνδυνος που ελλοχεύει και πριν εκδηλωθεί μοιάζει με ασήμαντη απειλή. Επειδή η διακίνηση των πληροφοριών είναι ελεύθερη και προστατεύεται από το δικαίωμα στην πληροφόρηση, και επιπλέον το ευρωπαϊκό δίκαιο είναι ένα κοινό ενιαίο μοντέλο, είναι επομένως δυνατό **να μεταφερθεί γεωγραφικά** το κέντρο που θα διαχειρίζεται τις πληροφορίες. Έχει ήδη εκφραστεί η επιθυμία της ΕΕ να δημιουργήσει ένα ενιαίο κέντρο δεδομένων το **European Data Space**, το οποίο θα αποτελεί δεξαμενή δεδομένων και για άλλες εφαρμογές που στο μέλλον θα αναπτυχθούν (<https://ec.europa.eu/digital-single-market/en/news/coordinated-plan-artificial-intelligence>, σελ 13 του pdf). Είναι προφανές ότι όποιος έχει πρόσβαση σε αυτές τις πληροφορίες για όλους τους πολίτες έχει στα χέρια του μια μεγάλη πολιτική δύναμη. Ο τρόπος που θα χρησιμοποιηθούν αυτές οι πληροφορίες τελικά, εξαρτάται από την ηθική αυτών που τις κατέχουν. Η χρήση των πληροφοριών θα είναι είτε προς την καλή και δίκαιη μεταχείριση, είτε προς μια αυταρχική και ίσως τυραννική χρήση (Yuting Lin, Andreas Eisingerich, Hersen Doong. 2017). Η συγκέντρωση δύναμης μέσω των τεχνολογιών τεχνητής νοημοσύνης δεν μπορεί διόλου να συγκριθεί με την συγκέντρωση δύναμης που σήμερα είναι γνωστή, αυτό αποδεικνύεται από το γεγονός ότι πρωταγωνιστές στη διεθνή οικονομική σκηνή, όπως ο οικονομολόγος George Soros αναγνωρίζουν ότι βαίνουμε σε μια κατάσταση πρωτόγνωρη συγκρινόμενη με αυτή που έχει περιγράψει ο George Orwell. Αξίζει εδώ να μεταφερθεί ολόκληρη η δήλωση του για το σύστημα κοινωνικής βαθμολόγησης, όπως κατατέθηκε στο παγκόσμιο οικονομικό φόρουμ του Νταβός : “Μια συμμαχία μεταξύ των εξουσιαστικών κρατών και των μεγάλων μονοπωλίων πληροφορικής, πλούσιων σε κατοχή δεδομένων, που φέρνουν σε επαφή τα νεογέννητα συστήματα εταιρικής επιτήρησης με ένα ήδη αναπτυσσόμενο σύστημα επιτήρησης από το κράτος. Αυτό μπορεί να οδηγήσει σε έναν ιστό ολοκληρωτικού ελέγχου, τον οποίο δεν μπορούσε καν να φανταστεί ο George Orwell. (πηγή Remarks delivered at the World Economic Forum Davos. 2019)”

Τα θεμέλια μιας μη δημοκρατικής αντιμετώπισης των πολιτών μπαίνουν ήδη από την ιδρυτική διακήρυξη μιας πολιτικής ένωσης και φανερώσουν ποια πορεία θα ακολουθήσει στο μέλλον. Αυτό φαίνεται στον καθορισμό της διαδικασίας λήψης των

πολιτικών αποφάσεων. Ακόμη και στον θεσμό της ΕΕ, μετά από δεκαετίες ύπαρξης, δεν έχει ωριμάσει η ιδέα της πραγματικής δημοκρατίας. Οι αποφάσεις στην ΕΕ παίρνονται από την Επιτροπή ή το Συμβούλιο και όχι από το κοινοβούλιο των πραγματικών αντιπροσώπων όπως θα απαιτούσε μια δημοκρατικότερη μεταχείριση. Η βούληση των πολιτών περνά μέσα από ένα κανάλι που μπορεί διακόπτεται είτε με την αδιαφορία των ευρωβουλευτών για τους πολίτες, είτε με την αδιαφορία των Επιτρόπων για το κοινοβούλιο. Έτσι βλέπουμε να παίρνονται πολιτικές αποφάσεις στην ΕΕ, π.χ. για το ύψος του ρυθμού ‘‘ανάπτυξης’’, οι οποίες εφαρμόζονται πάνω σε φτωχούς πολίτες με απόλυτη αδιαφορία αν αυτοί έχουν την δυνατότητα να ανταπεξέλθουν. Παράλληλα έχει εγκατασταθεί το κατάλληλο αυτοματοποιημένο σύστημα φορολόγησης, και ίσως στο μέλλον ένα σύστημα τεχνητής νοημοσύνης, που εξασφαλίζει το μέρος της αυτόματης χρέωσης των πολιτών. Η ΕΕ είναι ένα ζωντανό παράδειγμα μιας **μεταφοράς εξουσιών** εκτός των ορίων ενός κράτους. Επίσης το έλλειμμα της δημοκρατίας δίνει την **δυνατότητα** να μετατραπεί η αντιμετώπιση των πολιτών σε τυραννική. Ο ρόλος της τεχνολογίας και ιδιαίτερα των αυριανών chatbots δεν θα είναι μόνον εξυπηρετικός αλλά παράλληλα και ένα **φράγμα** ή αλλιώς ένα **στεγανό** στην προσέγγιση των κέντρων εξουσίας από τους πολίτες. Ένα τέτοιο πρόσφατο παράδειγμα είναι η αντιμετώπιση των πολιτών που έχουν μη εξυπηρετούμενα δάνεια: ενώ πριν η απόφαση για το μέλλον ενός δανείου ήταν στα χέρια ενός δικαστή που κρίνει κατά συνείδηση, με τις νέες ρυθμίσεις η απόφαση θα λαμβάνεται από μια ηλεκτρονική εφαρμογή με ντετερμινιστικά κριτήρια (http://www.keyd.gov.gr/hr_ocw_aithsh-link/), και συνεπώς ο πολίτης θα έχει απέναντί του μια μηχανή και δεν θα μπορεί να απευθυνθεί σε έναν συνειδητό άνθρωπο που είναι υπεύθυνος για την λήψη των αποφάσεων.

Έτσι λοιπόν η τεχνητή νοημοσύνη μπορεί να δημιουργήσει τις συνθήκες της μεταφοράς της κυβέρνησης εκτός μιας χώρας. Βέβαια η απόσταση από μια παγκόσμια ή ευρωπαϊκή διακυβέρνηση είναι ίσως μεγάλη. Η ετερογένεια πολλών κρατών (Viktor Mayer-Schönberger, David Lazer. 2007) στέκεται εμπόδιο σε αυτή την επιδίωξη. Επίσης όσο μεγαλώνουν οι κοινωνικές ανισότητες και όσο η φτώχεια θα αγγίζει περισσότερο πληθυσμό η αντίδραση των πολιτών σε αυτή την ανισόροπη ανάπτυξη του δημοσίου θα αυξάνει.

Κλείνοντας, θα πρέπει να απαντηθεί το κυρίαρχο ηθικό δίλλημα: αν εν όψει ενός μεγάλου κινδύνου μιας τυραννικής χρήσης των πληροφοριών που θα συλλέγονται από μηχανές τεχνητής νοημοσύνης, θα πρέπει να μην επιτραπεί η ανάπτυξη τέτοιων εφαρμογών ή να αποδεχτούμε τον κίνδυνο λαμβάνοντας κάποιου είδους αντιμετρών; Σε σχετική έρευνα (France Bélanger, Lemuria Carter. 2008) φαίνεται ότι από τους πολίτες υπάρχει **αμφισβήτηση των προθέσεων** που έχουν οι πολιτικοί και στον τρόπο που λειτουργούν τα πολιτικά κέντρα λήψης αποφάσεων σχετικά με την προώθηση της ηλεκτρονικής διακυβέρνησης και άρα των μέσων που αυτή θα χρησιμοποιήσει για να εδραιωθεί. Πολλοί είναι αυτοί που πιστεύουν ότι με τα κατάλληλα μέτρα, όπως η διαφάνεια στην λήψη των αποφάσεων από τα αυτοματοποιημένα συστήματα, θα εξαλειφτεί ο κίνδυνος μιας κακής χρήσης. Όμως ο

θεμελιώδης στόχος όπως έχει ήδη αναλυθεί είναι η βελτίωση της ποιότητας της ζωής του πολίτη, και όταν η ποιότητα της ζωής βαίνει στο χειρότερο είναι δύσκολο να αποδείξει κανείς τις επικαλούμενες προσπάθειες για επίτευξη πολλαπλών στόχων (βελτίωση δημοσίων υπηρεσιών, ανοιχτή διακυβέρνηση, βελτίωση εμπιστοσύνης,) από την υιοθέτηση της ηλεκτρονικής διακυβέρνησης (Damascene Twizeyimanaa, Annika Andersson.2019).

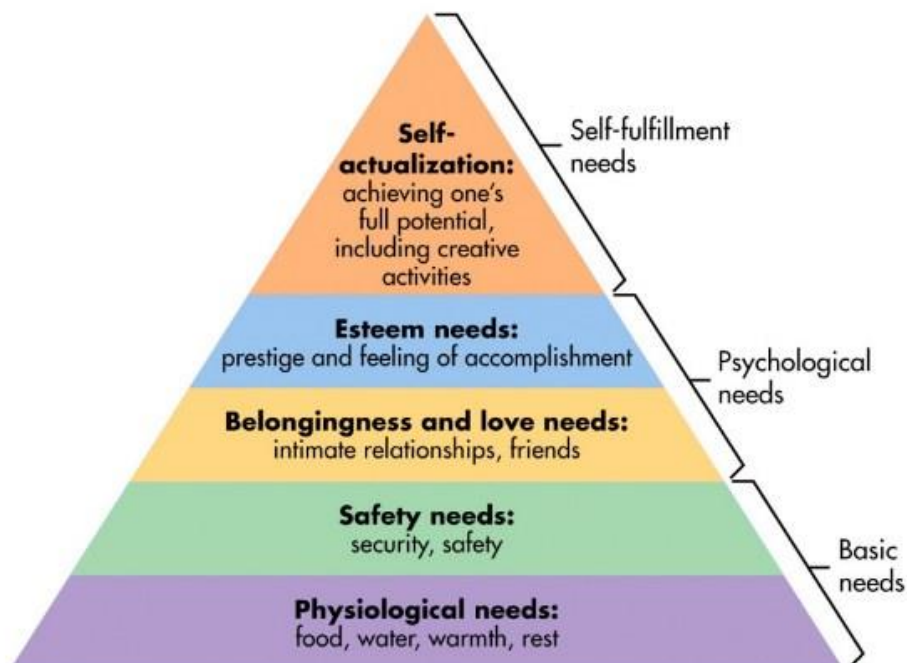
11.4. Έρευνες για την αποδοχή της τεχνητής νοημοσύνης από τους πολίτες

Εν κατακλείδι γεννάτε το ερώτημα αν θα μπορούσαμε να πούμε ότι μια εφαρμογή τεχνητής νοημοσύνης θα γίνει αποδεκτή από τους ανθρώπους και ποια θα είναι η επίδραση τους στο γενικό καλό; Στη συνέχεια θα γίνει μια αναφορά σε δύο έρευνες που εξετάζουν την αποδοχή από τους πολίτες σε εφαρμογές τεχνητής νοημοσύνης.

A. Ένα πείραμα που παρουσιάζεται από τον IEEE και αφορά κάποιες περιπτώσεις αυτόνομων συστημάτων (ένα μικρό ρομποτικό άκρο, ένα ρομπότ λιανικής πώλησης, ένα chatbot για την ψυχική υγεία, και ένα ρομπότ συντροφιάς) και την αλληλεπίδρασή τους με τους ανθρώπους, δείχνει τις αντιδράσεις των ανθρώπων στην θέαση αυτών των μηχανών. Οι αντιδράσεις αυτές φαίνονται στα σχήματα 7 έως 10.

Οι αντιδράσεις μετρήθηκαν με βάση την εκδοχή του Maslow για την ιεράρχηση των αναγκών του ανθρώπου, όπως αυτές φαίνονται στο παρακάτω σχήμα (Σχήμα 6. Ιεράρχηση των αναγκών του ανθρώπου.):

Σχήμα 6. Ιεράρχηση των αναγκών του ανθρώπου (πηγή: <https://www.simplypsychology.org/maslow.html>).



Ενδείξεις: 0=καθόλου αντίκτυπος, - = αρνητικός αντίκτυπος, + = θετικός αντίκτυπος, ?= άγνωστος αντίκτυπος.

Σχήμα 7. Αντιδράσεις ανθρώπων σε αυτόνομα συστήματα (πηγή: IEEE).

Small Factory Arm	Individual Direct	Individual Indirect	Environment Direct	Individual Indirect	Social Direct	Social Indirect
Basic Needs	+	+	0	-	+	+
Safety	?	?	?	-	?	+
Belonging	-	-	0	0	0	0
Esteem	-	-	0	0	0	0
Self-Actualization	0	0	0	0	0	0
Overall Impact	Mild -	Mild -	Nil	Mild -	Mild +	Mild +

Σχήμα 8. Αντιδράσεις ανθρώπων σε αυτόνομα συστήματα (πηγή: IEEE)..

Retail Kiosk Robot	Individual Direct	Individual Indirect	Environment Direct	Individual Indirect	Social Direct	Social Indirect
Basic Needs	0	0	0	–	+	?
Safety	?	?	?	–	+	?
Belonging	+	?	+	?	+	?
Esteem	+	?	0	0	+	?
Self-Actualization	?	?	0	0	?	?
Overall Impact	Mild +	Unknown	Very Mild +	Mild –	Strong +	Unknown

Σχήμα 9. Αντιδράσεις ανθρώπων σε αυτόνομα συστήματα (πηγή: IEEE)..

Mental Health Chatbot	Individual Direct	Individual Indirect	Environment Direct	Individual Indirect	Social Direct	Social Indirect
Basic Needs	0	0	0	0	0	0
Safety	+	0	0	0	?	+
Belonging	+	?	0	0	?	–
Esteem	+	?	0	0	?	–
Self-Actualization	?	0	0	0	0	0
Overall Impact	Strong +	Unknown	Nil	Nil	Unknown	Mild –

Σχήμα 10. Αντιδράσεις ανθρώπων σε αυτόνομα συστήματα (πηγή: IEEE).

Companion Robot like Paro	Individual Direct	Individual Indirect	Environment Direct	Individual Indirect	Social Direct	Social Indirect
Basic Needs	0	0	0	-	0	0
Safety	+	?	0	0	0	0
Belonging	+	?	0	0	?	-
Esteem	+	?	0	0	?	-
Self-Actualization	?	0	0	0	0	0
Overall Impact						

Σχολιάζοντας τα παραπάνω αποτελέσματα μπορούμε να πούμε ότι, με εξαίρεση τις αντιδράσεις στο ρομπότ λιανικής πώλησης, οι αντιδράσεις των ανθρώπων κλίνουν προς την ουδέτερη και αρνητική πλευρά. Αυτό σημαίνει ότι η αποδοχή των αυτόνομων συστημάτων βρίσκεται σε ένα στάδιο αρχικής διερεύνησης και δεν έχει ακόμη γίνει μια καθημερινή συνήθεια των ανθρώπων.

B. Μια ακόμη έρευνα σχετικά με την αποδοχή των αυτόνομων συστημάτων που έγινε μεταξύ των ετών 2012-2017 σε Ευρωπαϊκά κράτη (Timo Gnambsa, Markus Appel. 2018), έδειξε ότι η συμπεριφορά των ανθρώπων απέναντι στα ρομπότ γίνεται ολοένα και πιο αρνητική ιδιαίτερα σε αυτά που δρουν βοηθητικά στην εργασία. Μόνο οι άνθρωποι μεγαλύτερων ηλικιών βλέπουν τα ρομπότ με περισσότερη συμπάθεια. Γενικότερα διαπιστώθηκε ένα αυξανόμενο κίνημα συντηρητισμού απέναντι στα ρομπότ.

11.5. Περίληψη

Η καθυστέρηση που έχει η χώρα μας σε σχέση με άλλες ευρωπαϊκές χώρες στον τομέα της τεχνητής νοημοσύνης, της δίνει τον απαραίτητο χρόνο για να δημιουργήσει το κατάλληλο περιβάλλον ώστε στην έλευση της τεχνητής νοημοσύνης, πολίτες και δημόσιο να είναι κατάλληλα προετοιμασμένοι. Το δημόσιο αφού χαράξει ένα στρατηγικό πλάνο πλεύσης θα πρέπει να δει στο εσωτερικό του και να αναγνωρίσει ποιες διαδικασίες και ποιες υπηρεσίες μπορεί να αντικαταστήσει με την τεχνητή νοημοσύνη, και έπειτα να δει με τρόπο αυτό θα γίνει, ποιες αλλαγές πρέπει να γίνουν και ποιες συνέπειες αυτό το εγχείρημα θα έχει. Οι πολίτες αφού πρώτα ενημερωθούν για τα οφέλη και τις αρνητικές συνέπειες της τεχνητής νοημοσύνης στην ζωή τους θα πρέπει να αποφασίσουν αν θα δεχτούν αυτή την αλλαγή στη ζωή τους. Όταν αυτό δεν είναι εφικτό θα πρέπει να δημιουργηθεί από την

πολιτεία το πλαίσιο που θα θεραπεύει όσο τον δυνατό τις συνέπειες (αντιμετώπιση ανεργίας, μείωση οικονομικής ανισότητας, αποφυγή συγκέντρωσης δύναμης κ.αλ.). Στο τέλος αυτού του κεφαλαίου παρουσιάζονται και δύο έρευνες για να φανερωθεί ποια είναι εν τέλει η γνώμη των πολιτών για την τεχνητή νοημοσύνη. Τα αποτελέσματα της έρευνας κλίνουν περισσότερο προς την αρνητική στάση.

12. Λογισμικό εταιρειών για ανάπτυξη νευρωνικών δικτύων και εφαρμογών τεχνητής νοημοσύνης

12.1. Εισαγωγή

Η ανάπτυξη εφαρμογών τεχνητής νοημοσύνης βασίζεται κυρίως στην διαδικασία της μάθησης μηχανών. Αυτό σημαίνει ότι είναι απαραίτητη η ανάπτυξη κατάλληλου λογισμικού. Ο δημόσιος τομέας έχει την οικονομική δυνατότητα είτε να αναπτύξει δικό του λογισμικό είτε να αγοράσει έτοιμα πακέτα από σχετικές εταιρείες. Σε αυτό τα κεφάλαιο θα γίνει μια αναφορά στον χώρο της ανάπτυξης λογισμικού τεχνητής νοημοσύνης.

12.2 Λογισμικό εταιρειών για ανάπτυξη νευρωνικών δικτύων και εφαρμογών τεχνητής νοημοσύνης

Αναφέρθηκε στα προηγούμενα κεφάλαια ότι η τεχνητή νοημοσύνη βασίζεται στην διαδικασία της μάθησης που με την σειρά της βασίζεται στα νευρωνικά δίκτυα. Εν συντομία, και επειδή η βαθύτερη ανάλυση δεν είναι αντικείμενο της διατριβής, θα μπορούσαμε να πούμε μόνον λίγα λόγια για το τι είναι ένα νευρωνικό δίκτυο. Ένα νευρωνικό δίκτυο συντίθεται από απλά στοιχεία που είναι παράλληλα συνδεδεμένα μεταξύ τους. Κάθε σύνδεση μεταξύ δύο στοιχείων έχει ένα “βάρος”, δηλαδή μια χαρακτηριστική ιδιότητα της σύνδεσης. Κάθε φορά που εισάγονται νέα δεδομένα στο δίκτυο, αυτό παράγει μια έξοδο, στη συνέχεια συγκρίνει την έξοδο με έναν στόχο και αναπροσαρμόζει τα βάρη. Αυτή η αναπροσαρμογή είναι ουσιαστικά και η διαδικασία της μάθησης (<https://www.mathworks.com/products/deep-learning.html>).

Αν κάποιος οργανισμός θέλει να ασχοληθεί μόνος του με την δημιουργία νευρωνικών δικτύων μπορεί είτε να χρησιμοποιήσει δωρεάν λογισμικό όπως:

1. Tensor flow της Google,

2. CNTK της Microsoft.,

είτε λογισμικό που πωλείτε όπως:

. 3 Matlab Neural Network.

Εφόσον η κυβέρνηση ή κάποιος κρατικός οργανισμός θέλει να χρησιμοποιήσει στις διαδικασίες του μια ολοκληρωμένη και έτοιμη λύση στην τεχνητή νοημοσύνη έχει να επιλέξει μεταξύ κάποιων εταιρειών που πωλούν έτοιμα πακέτα λογισμικού τεχνητής νοημοσύνης ή κάποιων εταιρειών που λειτουργούν μη κερδοσκοπικά.

Μια ηγετική εταιρεία που έχει αναπτύξει τέτοιο λογισμικό είναι η IBM. Το λογισμικό ονομάζεται IBM Watson. Καύχημα της εταιρείας είναι η χρήση της πιο σύγχρονης μηχανής μάθησης που δεν χρειάζεται μεγάλο όγκο δεδομένων. Η εταιρεία

υπόσχεται ότι πετυχαίνει αύξηση παραγωγικότητας, αλλαγή στον τρόπο που εξυπηρετούνται οι πολίτες, εμπιστοσύνη στο σύστημά της, δημιουργία chatbot, προσαρμογή στις απαιτήσεις των πολιτών, διαφάνεια στον τρόπο λήψης των αποφάσεων.

Επιπλέον η ίδια εταιρεία έχει παράγει ένα εξειδικευμένο εργαλείο τεχνητής νοημοσύνης για τον τομέα της υγείας, το Watson Health. Η εφαρμογή αυτή δέχεται δεδομένα που αφορούν διάγνωση και θεραπεία καρκίνου, σε οποιαδήποτε γλώσσα και οποιαδήποτε χρονολογία. Επίσης παίρνει και όλα τα νέα δεδομένα που δημοσιεύονται. Το σύστημα για έναν δεδομένο ασθενή εκδίδει μια θεραπεία συνδυάζοντας τα δεδομένα στην βάση του με τα δεδομένα του ασθενή, όπως: το γονιδίωμα του, το ιατρικό ιστορικό του, την απεικόνιση και την παθολογική του κατάσταση. Σκοπός των ανθρώπων της IBM είναι το σύστημα αυτό να λειτουργεί σαν βοηθητικό και υποστηρικτικό στους θεράποντες ιατρούς στην λήψη αποφάσεων για μια σωστή διάγνωση και όχι στην αντικατάστασή τους. Πρέπει να αναφερθεί ότι το 2017 η MD Anderson Clinic διέκοψε την συνεργασία της με την IBM λόγω του μεγάλου κόστους λειτουργίας και λόγω μετανάστευσης δεδομένων μεταξύ ιατρικών εγγραφών.

Μια ακόμη μεγάλη εταιρεία που ανέπτυξε πακέτο τεχνητής νοημοσύνης για επιχειρήσεις (και προφανώς μπορεί να χρησιμοποιηθεί και από το δημόσιο, όπως αναφέρει και η ίδια) είναι η Microsoft με το Microsoft AI (<https://www.microsoft.com/en-us/ai>). Μια εφαρμογή της για τον τομέα της υγείας που θα μπορούσε να φανεί χρήσιμη στο δημόσιο είναι το Ochsner Health System που δρα βοηθητικά προς τους γιατρούς ώστε να αναλύσουν χιλιάδες ιατρικά δεδομένα.

Επιπλέον το Microsoft AI κάνει λόγο για την αντιμετώπιση των ηθικών αλλαγών που διαφαίνονται, καθιερώνοντας ένα πλαίσιο έξι αρχών: δικαιοσύνη, αξιοπιστία, διαφάνεια, ασφάλεια και προστασία της ιδιωτικότητας, αποκλειστικότητα, και καταλογισμό.

Κάτω από το ίδιο πρόγραμμα αναπτύχθηκαν κάποιες εφαρμογές με κοινωνικό χαρακτήρα, και θα μπορούσαν να υιοθετηθούν από το δημόσιο, όπως: Η εφαρμογή Seeing AI η οποία βοηθά τους ανθρώπους με δυσκολία στην όραση να σκιαγραφούν ανθρώπους και αντικείμενα που είναι γύρω τους καθώς και να διαβάζουν κείμενα. Επίσης αναπτύχθηκε μια εφαρμογή αναγνώρισης προσώπου ώστε να βοηθά παιδιά που χρειάζονται πλαστική επέμβαση. Στο ίδιο πλαίσιο λειτουργεί η εφαρμογή Premonition για την προστασία της βιοποικιλότητας.

Εκτός από τις εταιρείες που πωλούν έτοιμα πακέτα τεχνητής νοημοσύνης υπάρχει και η μη κερδοσκοπική εταιρεία OpenAI, της οποίας αποστολή είναι παραγωγή ανοιχτού λογισμικού.

Όσον αφορά την δημιουργία ενός chatboat η εταιρεία Python έχει αναπτύξει το ελεύθερο λογισμικό Chatterbot, που μπορεί να κληθεί με κατάλληλη εντολή (pip install chatterbot) μέσω των γραμμών κώδικα της γλώσσας Python.

12.3. Περίληψη

Όταν ένας δημόσιος φορέας θέλει να χρησιμοποιήσει την τεχνητή νοημοσύνη στις λειτουργίες του μπορεί, είτε να δημιουργήσει από το μηδέν μια εφαρμογή χρησιμοποιώντας δωρεάν ή με πληρωμή λογισμικό δημιουργίας νευρωνικών δικτύων, είτε να αγοράσει ένα έτοιμο πακέτο εφαρμογής από ιδιωτικές εταιρείες που πωλούν τέτοιο λογισμικό-υλικό, είτε να χρησιμοποιήσει λογισμικό από εταιρείες που το παρέχουν δωρεάν.

13. SWOT analysis

13.1. Εισαγωγή

Έχουν εκφραστεί κατά καιρούς πολλές απόψεις από ανθρώπους που δημιουργούν εφαρμογές τεχνητής νοημοσύνης ότι πρέπει άμεσα να ληφθούν μέτρα για τον περιορισμό της ανάπτυξης της τεχνητής νοημοσύνης προς κάποιες κατευθύνσεις (Paul J.Werbos. 2019). Πώς όμως θα ληφθεί μια τέτοια απόφαση;

Σε προηγούμενο κεφάλαιο αναφέρθηκε η έννοια της μεσότητας και η έννοια της αποδοχής ενός ρίσκου με σκοπό να βελτιωθεί η ζωή μας προς τα καλύτερο. Όταν πρόκειται λοιπόν να ληφθεί μια απόφαση, πρέπει πριν να προηγηθεί μια εκτίμηση για την επιτυχία ή την αποτυχία. Μια από τις γνωστότερες μεθόδους για την βοήθεια στην στρατηγική θεώρηση και την επιτυχή λήψη αποφάσεων είναι η SWOT analysis, η οποία δίνει μια ολοκληρωμένη εικόνα ενός προβλήματος. Αν και η εφαρμογή της αφορά κυρίως επιχειρήσεις, αυτό δεν αποκλείει και την χρήση της είτε σε οργανισμούς είτε στον ευρύτερο δημόσιο τομέα. Στη προκειμένη περίπτωση θα εξεταστεί αν ο δημόσιος τομέας μετά την χρήση εφαρμογών της τεχνητής νοημοσύνης θα μεταβεί σε μια κατάσταση ηθικότερη της προηγούμενης ή όχι. Έτσι εξετάζοντας αφενός μεν την εσωτερική κατάσταση του δημοσίου τομέα, δηλ. τις αδυναμίες του και τις δυνατότητες του και αφετέρου δε τα εξωτερικό περιβάλλον δηλ. τις ευκαιρίες και τους κινδύνους, θα προκύψει ένα γενικό συμπέρασμα.

13.2. SWOT analysis

Στο παρακάτω σχήμα (Σχήμα 11. Τα πεδία της SWOT analysis.) φαίνονται τα πεδία της ανάλυσης που θα ακολουθήσει.

Σχήμα 11. Τα πεδία της SWOT analysis.

	Βοηθητικό στο να επιτευχθεί ο στόχος	Εμπόδιο στο να επιτευχθεί ο στόχος
Εσωτερική ανάλυση	Δυνατότητες	Αδυναμίες
Εξωτερική ανάλυση	Ευκαιρίες	Κίνδυνοι

Παρακάτω θα αναλυθεί ξεχωριστά το καθένα πεδίο.

13.2.1. Δυνατότητες

Η κυρίαρχη δυνατότητα του δημοσίου τομέα είναι η άμεση και καθολική εφαρμογή ενός νέου προγράμματος σε όλη την επικράτεια ενός κράτους ή μιας ένωσης κρατών όπως η Ευρωπαϊκή Ένωση. Η δυνατότητα αυτή πηγάζει αφενός μεν από τα μεγάλη άνεση οικονομικής χρηματοδότησης ακριβών προγραμμάτων του δημοσίου τομέα και αφετέρου δε από την ήδη επαρκώς εγκατεστημένη υποδομή (τεχνική αλλά και σε ανθρώπινους πόρους) που θα υποστηρίξει και θα εφαρμόσει τα νέο πρόγραμμα. Τα οικονομικά μεγέθη του δημοσίου είναι ασύγκριτα με αυτά του ιδιωτικού τομέα και μπορούν να χρηματοδοτήσουν πολύ μεγάλα προγράμματα, επίσης το εργατικό δυναμικό του δημοσίου αγγίζει σχεδόν το 10% του πληθυσμού, πράγμα που δίνει την σιγουριά ότι υπάρχουν αρκετοί και εξειδικευμένοι πολίτες για να υποστηρίξουν το πρόγραμμα. Η τήρηση ενός προγράμματος βασίζεται στον επιτακτικό χαρακτήρα των αποφάσεων του δημοσίου προς τους πολίτες του, αλλά και την δυνατότητα περιφρούρησης της εφαρμογής του. Η οδός λοιπόν του δημοσίου τομέα είναι η ταχύτερη και η αποτελεσματικότερη για την εφαρμογή ενός νέου προγράμματος τεχνητής νοημοσύνης. Βέβαια αυτή η δυνατότητα μπορεί να αποβεί μοιραία για την ευημερία των πολιτών όταν η εξουσία έλθει στα χέρια ανήθικων κυβερνώντων που δεν θα σέβονται την δημοκρατικά κατοχυρωμένη λαϊκή κυριαρχία, και θα αποφασίζουν ενάντια στην γνώμη της πλειοψηφίας. Τότε, λόγω της καθολικότητας, κανένας πολίτης δεν θα μπορεί να αντιδράσει ενάντια στις μη δημοκρατικές αποφάσεις και εν τέλει να τις αποτρέψει. Βλέπουμε λοιπόν ότι μια δυνατότητα του δημοσίου τομέα κάτω από διαφορετικές συνθήκες να μετατρέπεται

στην μεγαλύτερη αδυναμία του, και ταυτόχρονα να αποτελεί έναν κίνδυνο για τους πολίτες. Αυτή είναι και η μεγαλύτερη απειλή μιας ηλεκτρονικής απομακρυσμένης διακυβέρνησης, ότι δηλαδή θα γίνει ένα χρήσιμο εργαλείο για να απομακρυνθεί και άλλο μια ήδη ανήθικη κυβέρνηση από τα πραγματική ζωή των πολιτών, οι οποίοι δεν θα μπορούν να απευθυνθούν και να εισακουστούν από αυτή με κανένα τρόπο. Αυτή βέβαια φαίνεται πλέον, με τις αποφάσεις που παίρνονται ερήμην της βούλησης των πολιτών, να αποτελεί στόχο και όχι μια πιθανή απειλή. Πρόσφατο παράδειγμα μιας τέτοιας αδιαφορίας προς την κυρίαρχη βούληση των πολιτών είναι η μη λήψη υπόψη από την κυβέρνηση του δημοψηφίσματος για την αποδοχή νέων μέτρων έναντι νέου δανεισμού, αλλά και η αδιαφορία για την κυρίαρχη άποψη για την ονοματοδοσία του γειτονικού μας κράτους με το αρχαίο Ελληνικό όνομα της Μακεδονίας. Όπως αναφέρθηκε και από την Επιτροπή Ηθικής της ΕΕ (EGE), η συγκέντρωση της δύναμης σε χέρια λίγων αποτελεί κίνδυνο για την κοινωνία. Σε παρόμοιες περιπτώσεις όταν το αντικείμενο που διαχειρίζονται κάποιοι είναι επικίνδυνο πρέπει αυτοί που το διαχειρίζονται να μην μπορούν να το αποκτήσουν, άρα και να το χρησιμοποιήσουν μόνο προς δικό τους όφελος. Για παράδειγμα όταν μια εφαρμογή τεχνητής νοημοσύνης απαιτεί τα γεωμετρικά χαρακτηριστικά των πολιτών θα πρέπει οι διαχειριστές στον κεντρικό έλεγχο του συστήματος να μην έχουν πρόσβαση σε αυτά ώστε να αποτραπεί μια μελλοντική κατάχρησή τους με ανήθικους σκοπούς. Δεν πρέπει δηλαδή να επιβεβαιωθεί τα ρητό ‘‘ ο επικίνδυνος άνθρωπος γίνεται ακόμη πιο επικίνδυνος αν στα χέρια του περιέλθουν πολύτιμες πληροφορίες’’.

13.2.2. Αδυναμίες

Η κυριότερη αδυναμία του δημοσίου τομέα είναι η παγιωμένη νοοτροπία μέρους των υπαλλήλων του να θεωρούν ότι μπορούν να είναι αργόσχολοι, χωρίς αυτό να έχει αντίκτυπο στις αποδοχές τους αλλά και στην δυνατότητα άρσης της μονιμότητας της εργασίας τους. Αυτό έχει σαν αποτέλεσμα την καθυστερημένη παροχή υπηρεσιών και εν τέλει την οικονομική ζημία και του ίδιου του δημοσίου αλλά και όσων εμπλέκονται με αυτό. Βέβαια ένα τεχνολογικό σύστημα τεχνητής νοημοσύνης μπορεί να παρακάμψει αυτή τη νοοτροπία, με την προϋπόθεση ότι οι κατασκευαστές του το επιθυμούν αυτό και δεν θέλουν να διαιωνίζεται μια τέτοια αναποτελεσματική κατάσταση.

Μια ακόμη κακή νοοτροπία, που έχει παρατηρηθεί τα τελευταία χρόνια της οικονομικής κρίσης, είναι και η ανάπτυξη από μέρους των δημοσίων υπαλλήλων μιας **ελιτίστικης συμπεριφοράς**, που προσομοιάζει με παρασιτική επιβίωση εις βάρος των άλλων πολιτών, αφού αυτοί έχουν καταλάβει ότι είναι πλέον οι μόνοι που θέλουν τον επιπλέον δανεισμό και την συνεπαγόμενη υπερφορολόγηση των πολιτών, προκειμένου να μην χάσουν την ευημερία τους, κάτι που έχει συμβεί ήδη στον ιδιωτικό τομέα. Βέβαια αυτή η νοοτροπία στα χέρια κάποιου ανήθικου κυβερνήτη είναι χρήσιμη γιατί θα βρει μια ομάδα πολιτών ‘‘πρόθυμη’’ να στραφεί εναντίων των υπολοίπων. Αυτή η νοοτροπία με την είσοδο μιας εφαρμογής τεχνητής νοημοσύνης

ίσως να ενισχυθεί μιας και ένας μηχανισμός π.χ. που θα εισπράττει φόρους αυτόματα από τραπεζικούς λογαριασμούς, θα είναι προς την κατεύθυνση που επιθυμεί η ελίτ. Να επισημανθεί εδώ ότι η EGE, στην ίδια γνώμη που εκφράζει τον κίνδυνο της συγκέντρωσης δύναμης στα χέρια λίγων, κάνει λόγο και για τον κίνδυνο δημιουργίας μιας ελίτ πολιτών που θα έχουν αυξημένα προνόμια έναντι των άλλων πολιτών, σαν αντάλλαγμα για να εξυπηρετούν ένα ανήθικο σύστημα. Να σημειωθεί εδώ ότι το Σύνταγμά μας (άρθρο 4ο παρ.1&2) αν και αναφέρει ότι όλοι οι πολίτες είναι ίσοι και έχουν ίσα δικαιώματα αυτό συχνά καταστρατηγείται θέτοντας εξαιρέσεις που αποβαίνουν μεροληπτικές εναντίον κάποιων πολιτών. Πρόσφατο παράδειγμα που επιβεβαιώνει του λόγου το αληθές είναι η επιπλέον μοριοδότηση υποψηφίων μαθητών ιδιωτικών μόνον σχολείων για την εισαγωγή στα πανεπιστήμια της χώρας ή η ιδιαίτερη δικαστική μεταχείριση των αδικημάτων που έκαναν βουλευτές και υπουργοί.

13.2.3. Ευκαιρίες

Οι ευκαιρίες αναφέρονται στο εξωτερικό περιβάλλον του δημοσίου. Στην συγκεκριμένη διατριβή μας ενδιαφέρουν οι δυνατότητες που δίνει τομέας των εφαρμογών της τεχνητής νοημοσύνης, άρα οι ευκαιρίες που υπάρχουν θα διερευνηθούν από την σκοπιά των δυνατοτήτων που δίνει η τεχνητή νοημοσύνη.

Μεταξύ άλλων οι δυνατότητες που δίνει η τεχνητή νοημοσύνη στον δημόσιο τομέας μπορούν να ανευρεθούν στο παρακάτω σχήμα (Σχήμα 12. Οι δυνατότητες που μπορεί να δώσει τεχνητή νοημοσύνη στον δημόσιο τομέα.) που αποτελεί μια σύνοψη από την εταιρεία QCM TECHNOLOGIES INC. Αναλυτικότερα βλέπουμε ότι το δημόσιο με την βοήθεια της τεχνητής νοημοσύνης μπορεί να:

1. Να αναγνωρίζει έγκαιρα μεταδοτικές ασθένειες με την βοήθεια επιτευγμάτων στον τομέα της ιατρικής.
2. Να δημιουργεί δυναμικούς πίνακες πληροφόρησης όπου οι πολίτες θα βρίσκουν τις πληροφορίες που αναζητούν.
3. Να δημιουργήσει μηχανισμούς όπου η κυβέρνηση θα “παρακολουθεί” ή θα “ακούει” τις δημοσιεύσεις στα κοινωνικά δίκτυα, με σκοπό να τις χρησιμοποιεί στην λήψη αποφάσεων.
4. Να προλαμβάνει κυβερνοεπιθέσεις αναγνωρίζοντας εγκαίρως κακόβουλες συμπεριφορές.
5. Να διευκολύνει τις δημόσιες συγκοινωνίες με την δημιουργία αυτόνομων λεωφορείων που θα μεταφέρουν τους επιβάτες από και προς τους σταθμούς.
6. Να παρακολουθεί τα κοινωνικά δίκτυα με σκοπό την γρήγορη επισήμανση επειγουσών καταστάσεων.

7. Να διαβιβάζει με δυναμικό τρόπο αναφορές πεπραγμένων σε τμήματα πληροφορικής.
8. Να αναγνωρίζει εγκαίρως ψευδείς αξιώσεις χρηματικών παροχών.
9. Να εγκαταστήσει ένα αξιόπιστο σύστημα προγραμματισμού συναντήσεων.
10. Να δημιουργεί συνθετικά δεδομένα (δηλ. όχι απευθείας παρατηρήσιμα ή μετρήσιμα, αλλά παράγωγα) για την γρήγορη τροφοδοσία τους σε αλγόριθμους μηχανών μάθησης.
11. Να βοηθήσει την χειρονακτική εργασία, του ανθρώπινου δυναμικού, σε έγγραφα.
12. Να εντοπίζει εύκολα πιθανές παραβιάσεις του ιατρικού κώδικα δεοντολογίας.
13. Να εποπτεύει και να καταμετρά το πλήθος των πολιτών σε μια συγκεκριμένη περιοχή.
14. Να φέρει τα μέλη των μειονοτικών ομάδων (ηλικιωμένοι, παιδιά, ΑΜΕΑ, άστεγοι κλπ) εγγύτερα στις δημόσιες υπηρεσίες.
15. Να προτείνει στους πολίτες που αιτούνται, κάποιες πρόσθετες υπηρεσίες και πηγές.
16. Να συλλέγει και να αναλύει γρήγορα υλικό γεγονότων που συγκεντρώνεται από τις κάμερες τις αστυνομίας.
17. Να καθοδηγούν τους πολίτες στην επιθυμητή υπηρεσία ανεξάρτητα από τον τρόπο σύνταξης μιας ερώτησης του πολίτη (δηλ. τέλεια αναγνώριση του λεξιλογίου και του συντακτικού μια γλώσσας από τα chatbots)
18. Να προβλέπει και να αποτρέπει την πιθανή σύγκρουση οχημάτων στις δημόσιες οδούς.
19. Να αξιολογεί τα αποτελέσματα που έφερε η ακολουθούμενη παιδική ιατρική μέριμνα και να προτείνει αλλαγές.
20. Να διαχειρίζεται αυτοματοποιημένα τις περισσότερο ζητούμενες δημόσιες υπηρεσίες.
21. Να προβλέπει τις μελλοντικές ανάγκες σε συντήρηση και αντικατάσταση των δημοσίων υποδομών.
22. Να προβλέπει το έγκλημα προτείνοντας την περιπολία αστυνομικής δύναμης.
23. Να αναλύει την κυκλοφορία των οχημάτων με σκοπό τον προληπτικό σχεδιασμό αποφυγής συνωστισμού.
24. Να προβλέπει και να αποφεύγει την βλάβη των υποδομών παροχής υδάτων.

Σχήμα 12. Οι δυνατότητες που μπορεί να δώσει τεχνητή νοημοσύνη στον δημόσιο τομέα (πηγή www.qcmtech.com).



List sources: Daniel Castro, Steve Nichols and Eric Ellis (quoted above), Adobe Chief Information Officer Cynthia Stoddard and Government Technology reporting.

Συμπληρώνοντας στο παραπάνω σχήμα κάποιες επιπλέον δυνατότητες που θα μπορούσε να δώσει η τεχνητή νοημοσύνη μπορούμε να πούμε ότι:

25. θα αυξηθεί η δυνατότητα να αποδίδονται ευθύνες στους αρμόδιους υπαλλήλους,
26. θα δίνεται η δυνατότητα να παίρνονται αποφάσεις με ισχυρή τεκμηρίωση (informed decisions),
27. θα δίνεται η δυνατότητα για συνεργασίες με άλλους τομείς μέσα στο δημόσιο αλλά και τον ιδιωτικό τομέα,
28. θα μπορεί να δοθεί η δυνατότητα στους πολίτες στην άμεση συμμετοχή στη λήψη των αποφάσεων.

13.2.4. Απειλές

Πριν προχωρήσουμε στην αναφορά των πιθανών κινδύνων που κρύβουν οι εφαρμογές της τεχνητής νοημοσύνης για τον δημόσιο τομέα, θα πρέπει να γίνει ένας σημαντικός διαχωρισμός ως προς την προέλευση αυτών των κινδύνων.

Οι πηγές λοιπόν των κινδύνων είναι δύο:

-Πρώτον η ίδια η τεχνολογία και

-δεύτερον οι άνθρωποι, πίσω από την τεχνολογία, που λαμβάνουν αποφάσεις για το πώς αυτή θα χρησιμοποιηθεί.

Να σημειωθεί εδώ ότι αυτή η διάκριση των πηγών παύει να υπάρχει εάν το επίπεδο εξέλιξης της τεχνητής νοημοσύνης φτάσει σε σημείο να υπάρχει πλήρης αυτονομία λήψης αποφάσεων από τις ίδιες τις τεχνολογικές εφαρμογές, ώστε η ευθύνη λήψης αποφάσεων και των συνεπειών τους να μην βαρύνει τον άνθρωπο αλλά τις ίδιες τις μηχανές.

Στη συνέχεια θα γίνει μια σύντομη αναφορά στους κινδύνους μιας και για αυτούς έχει γίνει σε προηγούμενο κεφάλαιο αναλυτική παρουσίαση.

Ξεκινώντας λοιπόν την αναφορά στους κινδύνους μπορούμε να αναφέρουμε την **ανεργία** σαν έναν από τους μεγαλύτερους κινδύνους. Οι συνθήκες που ακολουθούν μια τεχνολογική εξέλιξη είναι γνωστές στον επιστημονικό χώρο σαν “τεχνολογική ανεργία”. Είναι βεβαιωμένο από την ιστορία ότι κάθε τεχνολογικό επίτευγμα προκαλεί συνολικά μια μείωση του συνολικού αριθμού των εργαζομένων στην οικονομία. Αναλυτικότερα έχουμε μια μείωση των εργαζομένων στον κλάδο που εφαρμόζεται η νέα τεχνολογική εφαρμογή (π.χ. απόλυση εργαζομένων που εργάζονται στην παροχή πληροφοριών σε ένα δημόσιο γραφείο και αντικατάστασή τους με chatbots) και παράλληλα μια μικρότερη αύξηση στον κλάδο που η εφαρμογή δημιουργεί ανάγκη εργασίας (π.χ. στην κατασκευή chatbots). Το αρνητικό αυτό

ισοζύγιο δεν αμφισβητείται γιατί ο σκοπός μιας νέας εφαρμογής είναι η απαλλαγή του ανθρώπου από εργασίες βαριές, μονότονες και μή δημιουργικές. Μετά βέβαια την απαξίωση της θέσης εργασίας γενάτε το ερώτημα αν ο απολυμένος εργαζόμενος έχει μερίδιο στο οικονομικό όφελος που θα προκύψει από την νέα εφαρμογή;

Σε αυτό το σημείο αναδεικνύεται ένας ακόμη κίνδυνος που είναι η **κάρπωση των οικονομικών κερδών** που θα προέλθουν από την τεχνητή νοημοσύνη. Να επισημανθεί εδώ ότι οι δημόσιες επιχειρήσεις στη χώρα μας έχουν επί το πλείστον απαξιωθεί και πολλές έχουν αποκρατικοποιηθεί, πράγμα που κάνει την πιθανότητα χάραξης μιας κοινωνικής πολιτικής μέσω διαμοιρασμού των επιπλέον κερδών απ την υιοθέτηση εφαρμογών τεχνητής νοημοσύνης πολύ δύσκολη, ταυτόχρονα η ιδέα αυτή είναι σχεδόν ανύπαρκτη στην σκέψη των ιδιοκτητών επιχειρήσεων στον ιδιωτικό τομέα.

Ένας από τους σημαντικότερους κινδύνους που πρέπει να αντιμετωπιστεί είναι η **συγκέντρωση της δύναμης** της κατοχής πληροφοριών στα χέρια λίγων ανθρώπων.

Έχει επισημανθεί και από την Ομάδα Ηθικής της ΕΕ (EGE) ότι τα συστήματα που συγκεντρώνουν πληροφορίες στα χέρια λίγων ανθρώπων αποτελούν κίνδυνο για την κοινωνία. Επειδή η τεχνητή νοημοσύνη και ιδιαίτερα εφαρμογές που θα χρησιμοποιηθούν από τον δημόσιο τομέα, απαιτούν σαν απαραίτητη προϋπόθεση λειτουργίας την συλλογή πληροφοριών (προσωπικών δεδομένων ή άλλων), θα μπορούσε να χρησιμοποιηθεί από την πολιτική διοίκηση προς το κακό των πολιτών. Παραδείγματος χάριν, εφόσον η ψηφιακή ταυτότητα (digital id) γίνει υποχρεωτική για όλους ώστε τα Chatbots του δημοσίου να αναγνωρίζουν την εικόνα των πολιτών θα μπορούσε αυτό να χρησιμοποιηθεί για την κρυφή παρακολούθηση ενός ανθρώπου από αυτούς που θα έχουν πρόσβαση στον κεντρικό έλεγχο του συστήματος των ψηφιακών ταυτοτήτων.

Ο κίνδυνος από την **αυτονόμηση** των μηχανών είναι ακόμα ένας φανταστικός κίνδυνος που θα εμφανιστεί κάτω από προϋποθέσεις πλήρους ανάπτυξης της τεχνητής νοημοσύνης.

Έχει ήδη γίνει αναφορά για τον κίνδυνο για την **ψυχική υγεία** των πολιτών που μπορεί να έλθει όταν το πλήθος των ρομπότ φτάσει σε τέτοιο σημείο που ο άνθρωπος να αισθάνεται ότι πολιορκείται από αυτά και στην χειρότερη περίπτωση ότι παραμερίζετε από αυτά.

Ο κίνδυνος της **κοινωνικής ανισότητας** είναι υπαρκτός μιας και σε κρατικό επίπεδο οι περισσότερο ανεπτυγμένες χώρες θα χρησιμοποιήσουν πρώτες την τεχνητή νοημοσύνη με αποτέλεσμα να διευρυνθεί το οικονομικό χάσμα με τις φτωχότερες χώρες με αποτέλεσμα να μεγεθυνθούν και οι κοινωνικές ανισότητες. Όσον αφορά τις ανισότητες σε προσωπικό επίπεδο μέσα στο ίδιο κράτος η εφαρμογή της τεχνητής νοημοσύνης θα δώσει μια εξισορρόπηση μεταξύ διάφορων κοινωνικών στρωμάτων γιατί όλοι θα απολαμβάνουν τις ίδιες υπηρεσίες από το δημόσιο, όμως λόγω του ότι ο

ιδιωτικός τομέας προχωρά πάντοτε πιο γρήγορα (π.χ. στον τομέα της υγείας) ακόμη και σε προσωπικό επίπεδο θα διευρυνθούν οι κοινωνικές ανισότητες.

Έχει γίνει ήδη αναφορά σε προηγούμενα κεφάλαια ότι θα υπάρξουν κίνδυνοι για μαζικές κλοπές σε τραπεζικούς λογαριασμούς και άλλα προσωπικά δεδομένα.

Αφού λοιπόν έχουν αναφερθεί οι ευκαιρίες, οι δυνατότητες, οι αδυναμίες και οι κίνδυνοι φτάνει να γίνει από τους αρμόδιους για την λήψη μιας απόφασης (στην συγκεκριμένη περίπτωση είναι οι κυβέρνηση σαν πρώτη αρχή για το δημόσιο) μια εκτίμηση για το αν οι δυνατότητες θα φέρουν τέτοια αποτελέσματα που θα δικαιολογούν την λήψη ρίσκου και την αποδοχή κάποιου κίνδυνου να επαληθευτούν.

13.3. Περίληψη

Σε αυτό το κεφάλαιο, προς την βοήθεια λήψης μιας απόφασης για την χρήση ή μη της τεχνητής νοημοσύνης, συγκεντρώθηκαν σε μια SWOT ανάλυση, οι δυνατότητες, οι αδυναμίες του δημοσίου αλλά και οι ευκαιρίες και απειλές που δίνει η τεχνητή νοημοσύνη. Αυτή η ανάλυση θα μπορούσε να είναι το μόνο υλικό που θα έχει στα χέρια του ο υπεύθυνος λήψης αποφάσεων, ώστε να λάβει την πολιτική απόφαση αν ο δημόσιος τομέας υιοθετήσει την τεχνητή νοημοσύνη, αλλά παράλληλα θα είναι και ένα εγχειρίδιο προς τους πολίτες για να ενημερωθούν για το μέλλον του δημοσίου τομέα υπό το πρίσμα της υιοθέτησης της τεχνητής νοημοσύνης.

14. Οι κίνδυνοι από την χρήση της τεχνητής νοημοσύνης και η σχέση τους με τον νέο κανονισμό της προστασίας των προσωπικών δεδομένων

14.1. Εισαγωγή

Έχει αναφερθεί σε προηγούμενες αναφορές ότι η βάση της ανάπτυξης μιας εφαρμογής της τεχνητής νοημοσύνης είναι η συλλογή και επεξεργασία μεγάλης ποσότητας δεδομένων. Τα δεδομένα αυτά είτε είναι προσωπικά δεδομένα π.χ. τα γεωμετρικά χαρακτηριστικά του προσώπου ενός πολίτη, είτε είναι άλλα δεδομένα π.χ. οι θερμοκρασίες του περιβάλλοντος.

Αν και η επεξεργασία και η χρήση των δεδομένων που δεν είναι προσωπικά θα μπορούσαν κάτω από συγκεκριμένες συνθήκες να κρύβουν κινδύνους για τον άνθρωπο π.χ. ο έλεγχος των καιρικών συνθηκών με σκοπό την καταστροφή αγροτικών καλλιεργειών, η παρούσα έρευνα θα επικεντρωθεί στον αμεσότερο κίνδυνο της πολιορκίας των θεμελιωδών δικαιωμάτων του πολίτη.

Αφού γίνει μια ανασκόπηση των κινδύνων, στη συνέχεια θα ερευνηθεί το πώς ο νέος κανονισμός για την προστασία των προσωπικών δεδομένων αντιμετωπίζει αυτούς τους κινδύνους.

14.2. Οι κίνδυνοι από την χρήση της τεχνητής νοημοσύνης

Επισημάνθηκε στα προηγούμενα κεφάλαια ότι η χρήση εφαρμογών τεχνητής νοημοσύνης συνδέεται άμεσα με κάποιους κινδύνους για τον άνθρωπο. Συνοψίζοντας μπορούμε να επαναλάβουμε ότι οι κίνδυνοι αυτοί είναι οι:

1. Η αύξηση του επιπέδου της ανεργίας λόγω του φαινομένου της ‘‘τεχνολογικής ανεργίας’’.
2. Η διεύρυνση του οικονομικού χάσματος μεταξύ οικονομικά αναπτυγμένων και υποανάπτυκτων χωρών λόγω της ανισόροπης ανάπτυξης.
3. Η αλλαγή στην αντίληψη της κοινωνικής συμβίωσης λόγω της παρουσίας ανθρωπόμορφων μηχανών.
4. Η αλλαγή στην έννοια της εμπιστοσύνης προς τα τεχνολογικά επιτεύγματα.
5. Ο κίνδυνος κακής χρήσης της τεχνολογίας για συγκέντρωση πολιτικής δύναμης.
6. Η δημιουργία ενός ελιτίστικου κοινωνικού φράγματος.
7. Ο κίνδυνος κακής χρήσης των υποδομών της τεχνητής νοημοσύνης με σκοπό την εγκατάσταση τυραννικών καθεστώτων.
8. Η εύκολη μεταφορά του κέντρου της διακυβέρνησης εκτός των ορίων ενός κράτους

9. Η χρήση της τεχνητής νοημοσύνης στα αυτόνομα όπλα.

10. Η πιθανή αυτονόμηση των μηχανών.

14.3. Η σχέση των κινδύνων με τον νέο κανονισμό της προστασίας των προσωπικών δεδομένων

Εφόσον, όπως αναφέρθηκε σε προηγούμενο κεφάλαιο, δεν υπάρχει στην χώρα μας κάποια επιτροπή που θα απαντήσει αν οι κίνδυνοι που αναφέρθηκαν είναι υπαρκτοί και αν θα πρέπει να ληφθούν κάποια μέτρα, αφήνουμε όλες τις προσδοκίες μας στην ανώτερη επιτροπή που σήμερα έχει την ευθύνη να επισημαίνει τους κινδύνους, και αυτή είναι η EGE.

Δυστυχώς τα αρμόδια όργανα της ΕΕ, αν και γνωρίζουν αυτούς τους κινδύνους, από τις προηγούμενες συστάσεις της EGE αλλά και άλλων ιδιωτικών φορέων, **αποκεντρώνουν και διαχέουν την ευθύνη τους**, μεταθέτοντας την ανάγκη επισήμανσης και αντιμετώπισης των κινδύνων και της προληπτικής δράσης, στον καθένα ξεχωριστά υπεύθυνο, είτε πρόκειται για υπεύθυνο επεξεργασίας προσωπικών δεδομένων, είτε πρόκειται για υπεύθυνο για την εφαρμογή νέων τεχνολογιών. Αυτή η διαπίστωση προκύπτει από τον νέο κανονισμό των προσωπικών δεδομένων και συγκεκριμένα στο άρθρο 35. Έτσι όσον αφορά τα προσωπικά δεδομένα θεσμοθετήθηκε η μελέτη εκτίμησης αντικτύπου, ενώ όσον αφορά τις νέες τεχνολογίες προτάθηκε η Εκτίμηση των Επιπτώσεων στην Προστασία της ιδιωτικής ζωής και των δεδομένων (ΕΕΠ). Παρακάτω θα γίνει εκτενής αναφορά για τις δύο αυτές περιπτώσεις (Α,Β).

A. Η μελέτη εκτίμησης αντικτύπου

Στον νέο κανονισμό προστασίας των προσωπικών δεδομένων της Ευρωπαϊκής Ένωσης (ΚΑΝΟΝΙΣΜΟΣ (ΕΕ) 2016/679 ΤΟΥ ΕΥΡΩΠΑΪΚΟΥ ΚΟΙΝΟΒΟΥΛΙΟΥ ΚΑΙ ΤΟΥ ΣΥΜΒΟΥΛΙΟΥ της 27ης Απριλίου 2016 για την προστασία των φυσικών προσώπων έναντι της επεξεργασίας των δεδομένων προσωπικού χαρακτήρα και για την ελεύθερη κυκλοφορία των δεδομένων αυτών και την κατάργηση της οδηγίας 95/46/ΕΚ (Γενικός Κανονισμός για την Προστασία Δεδομένων)) που ισχύει στη χώρα μας από τον Μάιο του 2018 (τέθηκε σε ισχύ κατά την διάρκεια συγγραφής αυτής της διατριβής) γίνεται αναφορά (Άρθρο 35, Εκτίμηση αντικτύπου σχετικά με την προστασία δεδομένων) για πρώτη φορά στην υποχρέωση που έχει κάθε υπεύθυνος επεξεργασίας, ιδίως όταν χρησιμοποιούνται νέες τεχνολογίες, και όταν η επεξεργασία ενδέχεται να επιφέρει υψηλό κίνδυνο για τα δικαιώματα και τις ελευθερίες των φυσικών προσώπων, να διενεργεί πριν την επεξεργασία, εκτίμηση των επιπτώσεων των σχεδιαζόμενων πράξεων επεξεργασίας στην προστασία δεδομένων προσωπικού χαρακτήρα.

Η εκτίμηση αντικτύπου σχετικά με την προστασία δεδομένων απαιτείται ιδίως στην περίπτωση:

α) συστηματικής και εκτενούς αξιολόγησης προσωπικών πτυχών σχετικά με φυσικά πρόσωπα, η οποία βασίζεται σε αυτοματοποιημένη επεξεργασία, περιλαμβανομένης της κατάρτισης προφίλ, και στην οποία βασίζονται αποφάσεις που παράγουν έννομα αποτελέσματα σχετικά με το φυσικό πρόσωπο ή ομοίως επηρεάζουν σημαντικά το φυσικό πρόσωπο,

β) μεγάλης κλίμακας επεξεργασίας των ειδικών κατηγοριών δεδομένων που αναφέρονται στο άρθρο 9 παράγραφος 1 ή δεδομένων προσωπικού χαρακτήρα που αφορούν ποινικές καταδίκες και αδικήματα που αναφέρονται στο άρθρο 10 του ίδιου κανονισμού ή

γ) συστηματικής παρακολούθησης δημοσίως προσβάσιμου χώρου σε μεγάλη κλίμακα.

Η κάθε εθνική εποπτική αρχή καταρτίζει και δημοσιοποιεί κατάλογο με τα είδη των πράξεων επεξεργασίας που υπόκεινται στην απαίτηση για διενέργεια εκτίμησης αντικτύπου σχετικά με την προστασία των δεδομένων. Η εποπτική αρχή ανακοινώνει τον εν λόγω κατάλογο στο Συμβούλιο Προστασίας Δεδομένων (οι αρμοδιότητες του οποίου αναφέρονται στο άρθρο 68 του ίδιου κανονισμού). Επίσης η κάθε εθνική εποπτική αρχή δύναται να καταρτίζει και να δημοσιοποιεί κατάλογο με τα είδη των πράξεων επεξεργασίας για τα οποία δεν απαιτείται εκτίμηση αντικτύπου σχετικά με την προστασία των δεδομένων. Η εποπτική αρχή ανακοινώνει τον εν λόγω κατάλογο στο Συμβούλιο Προστασίας Δεδομένων.

Όσον αφορά το περιεχόμενο μιας εκτίμησης αντικτύπου θα πρέπει αυτή κατ' ελάχιστον να περιέχει τα εξής:

α) συστηματική περιγραφή των προβλεπόμενων πράξεων επεξεργασίας και των σκοπών της επεξεργασίας, περιλαμβανομένου, κατά περίπτωση, του έννομου συμφέροντος που επιδιώκει ο υπεύθυνος επεξεργασίας,

β) εκτίμηση της αναγκαιότητας και της αναλογικότητας των πράξεων επεξεργασίας σε συνάρτηση με τους σκοπούς,

γ) **εκτίμηση των κινδύνων** για τα δικαιώματα και τις ελευθερίες των υποκειμένων των δεδομένων που αναφέρονται στην παράγραφο 1 του οικείου άρθρου του κανονισμού και

δ) τα προβλεπόμενα **μέτρα αντιμετώπισης των κινδύνων**, περιλαμβανομένων των εγγυήσεων, των μέτρων και μηχανισμών ασφάλειας, ώστε να διασφαλίζεται η προστασία των δεδομένων προσωπικού χαρακτήρα και να αποδεικνύεται η συμμόρφωση προς τον παρόντα κανονισμό, λαμβάνοντας υπόψη τα δικαιώματα και τα έννομα συμφέροντα των υποκειμένων των δεδομένων και άλλων ενδιαφερόμενων προσώπων.

Είναι προφανές ότι η επεξεργασία δεδομένων των πολιτών από το δημόσιο τομέα εντάσσεται χωρίς αμφισβήτηση στην κατηγορία της επεξεργασίας μεγάλης

κλίμακας διότι γίνεται σε πολλά υποκείμενα και σε μεγάλη γεωγραφική έκταση αλλά και σε μεγάλο όγκο δεδομένων. Επομένως η μελέτη εκτίμησης αντικτύπου είναι υποχρεωτική για τους δημόσιους φορείς που παρέχουν υπηρεσίες προς πολίτες.

Δυστυχώς τα μέτρα αντιμετώπισης τα οποία αναφέρονται έχουν πεδίο μόνο την προστασία των προσωπικών δεδομένων, και όχι, καθώς θα έπρεπε, ένα ευρύτερο πεδίο. Έτσι για παράδειγμα όταν ο επαπειλούμενος κίνδυνος είναι η **συγκέντρωση της δύναμης της πληροφορίας στα χέρια λίγων**, όπως αναφέρει η EGE, δεν είναι ένας άμεσος κίνδυνος και δεν προκαλεί την στιγμή που γίνεται κακό σε κανένα προσωπικό δεδομένο, είναι όμως κάτι που πρέπει να αντιμετωπιστεί, με μέτρα τα οποία υπερβαίνουν την δύναμη και το μέγεθος ενός υπεύθυνου επεξεργασίας, και ταυτόχρονα επιδρούν στο γενικό σύνολο. Έτσι για παράδειγμα αν σε ευρωπαϊκό επίπεδο η συγκέντρωση βιομετρικών στοιχείων όλων των πολιτών χαρακτηριστεί “επικίνδυνη συγκέντρωση δύναμης στα χέρια λίγων”, ποιος θα ήταν εκείνος που θα μπορούσε, σαν μέτρο αντιμετώπισης, να σταματήσει την συλλογή τους; Προφανώς η απάντηση είναι ότι αν και κάθε υπεύθυνος επεξεργασίας γνωρίζει αυτό τον κίνδυνο δεν θα μπορέσει να εναντιωθεί σε μια υπάρχουσα γιγαντιαία και στρεβλή κατάσταση.

Επειδή ο νέος κανονισμός προστασίας δεδομένων είναι προϊόν μακράς μελέτης και συλλογής πολλών επιστημονικών απόψεων και εμπειριστατωμένων γνώμων δεν δικαιολογείται να αφήσει τέτοιου είδους πεδίο υποψίας στοχευμένου λάθους ή εκ του πονηρού πράξεων, εις βάρος του. Οι κίνδυνοι που υπάρχουν για τους πολίτες είναι ήδη γνωστοί, από τους πιο εγκυρότερους θεσμούς σε θέματα ηθικής (π.χ. από την EGE), σε αυτούς που συνέταξαν τον κανονισμό, και παρόλα αυτά, αυτοί επέλεξαν να υποβαθμίσουν αυτούς τους κινδύνους και να επωμισθεί η αντιμετώπισή τους από το κατώτερο επίπεδο του υπευθύνου επεξεργασίας. Ενώ θα έπρεπε να αντιμετωπιστούν με το αρμόζων μέγεθος που θα είναι μια διεθνής ή μια εθνική επιτροπή ηθικής. Κατά του γνώμη του συγγραφέα της παρούσας διατριβής ο νέος κανονισμός εσκεμμένα κλείνει τα μάτια σε εγνωσμένους κινδύνους για την ανθρώπινη αξιοπρέπεια, και επιτρέπει να εγκατασταθεί ένα επικίνδυνο για τον άνθρωπο πλέγμα πληροφοριακής επιφυλακής και οικονομικού στραγγαλισμού, συγκρινόμενο με μυθική Οργουελική κοινωνία, παραπέμποντας στο μέλλον την αντιμετώπιση “τυχόν” προβλημάτων που θα εμφανιστούν, εξυπηρετώντας το ήδη ωφελημένο διεθνές χρηματοπιστωτικό κατεστημένο (<https://www.hba.gr/FinacialLaw/List?type=NationalOnlineBanking>, <https://diavgeia.gov.gr/doc/7%CE%A3%CE%A83%CE%97-%CE%97%CE%A99?inline=true>), μιας και όπως επισημάνθηκε στο Ελληνικό δημόσιο οι πληροφοριακές τεχνολογικές εξελίξεις αφορούν κυρίως φοροεισπρακτικούς στόχους που περνούν υποχρεωτικά μέσα από το τραπεζικό σύστημα .

B. Εκτίμηση των Επιπτώσεων στην Προστασία της ιδιωτικής ζωής και των δεδομένων (ΕΕΠ)

Αν και ο κανονισμός προστασίας προσωπικών δεδομένων είναι σχετικά πρόσφατος, η έννοια της μελέτης για τυχόν κινδύνους για τον άνθρωπο έχει αναφερθεί πολύ νωρίτερα από επίσημους θεσμούς της Ευρωπαϊκής Ένωσης. Έτσι η Ομάδα του Άρθρου 29, η οποία είναι μια συμβουλευτική προς την Ευρωπαϊκή Επιτροπή ομάδα σε μια από τις παλαιότερες γνώμες (Opinion No 9/2011) που είχε εκδώσει και αφορούσε νέες τεχνολογίες που χρησιμοποιούν αναγνώριση μέσω ραδιοσυχνικών ετικετών, πρότεινε προς την επιτροπή ότι κάθε ενδιαφερόμενη εταιρεία που θέλει να εγκαταστήσει μια εφαρμογή με τέτοια τεχνολογία θα πρέπει, ανάλογα με το επίπεδο ταξινόμησης, να εκπονεί Εκτίμηση των Επιπτώσεων στην Προστασία της ιδιωτικής ζωής και των δεδομένων (ΕΕΠ). Σύμφωνα με αυτή θα πρέπει να ακολουθείται η διαδικασία:

1. χαρακτηρισμός εφαρμογής (επίπεδα επικινδυνότητας: 0-κατώτερο έως 3-ανώτερο).
2. εντοπισμός κινδύνων.
3. συσταση για ελέγχους για τους κινδύνους.
4. τεκμηρίωση αποτελεσμάτων της ΕΕΠ, έλεγχος για παραμένοντες κινδύνους.

Έτσι λοιπόν, για παράδειγμα, όταν ένας πολίτης προσέρχεται σε μια δημόσια υπηρεσία και απευθύνεται σε ένα chatbot με σκοπό να απολάβει μια υπηρεσία θα πρέπει, πριν χρησιμοποιηθεί ένα σύστημα ταυτοποίησης του πολίτη από το chatbot, να έχει **προηγηθεί** μια μελέτη για τους κινδύνους που μπορεί να προκαλέσει μια τέτοια πράξη.

Βλέπουμε λοιπόν ότι ήδη από το 2011 έχει επισημανθεί ότι οι νέες τεχνολογίες, να μεν είναι ευεργετικές προς τον άνθρωπο, όμως κρύβουν κινδύνους οι οποίοι μπορούν να επιδράσουν αντίθετα από τον προσδοκώμενο στόχο. Να επισημανθεί εδώ ότι η πρόταση αυτή δεν υιοθετήθηκε ούτε λήφθηκε υπ' όψιν έως σήμερα από την Ευρωπαϊκή Επιτροπή.

Είναι προφανές ότι αυτή η τακτική της αποκέντρωσης της ευθύνης της λήψης μέτρων, προς τον υπεύθυνο επεξεργασίας και εκτιμητή του αντικτύπου, ή προς τον υπεύθυνο Εκτίμησης των Επιπτώσεων στην Προστασία της ιδιωτικής ζωής και των δεδομένων, εξυπηρετεί την **διάχυση της ευθύνης**. Είναι εύλογο το ερώτημα του πώς κάποιος υπεύθυνος επεξεργασίας που θα έχει υπό τον έλεγχό του π.χ. ένα chatbot που συγκεντρώνει εικόνες με γεωμετρικά χαρακτηριστικά προσώπων, θα έχει την δύναμη να εκφράσει ότι αυτή η ενέργεια μπορεί να αποτελεί κίνδυνο για την δημοκρατία και ποια είναι τα μέτρα που αυτός μπορεί να λάβει σε κάτι που έχει τεράστιες κοινωνικές διαστάσεις; Για να μην αδικηθούν πλήρως οι ευρωπαϊκοί θεσμοί πρέπει να ανφέρουμε ότι ο ευρωπαϊκός νμοθέτης αναγνωρίζει ότι υπάρχουν ρίσκα για τα δεδομένα και άλλα σημαντικά και οικονομικά μειονεκτήματα, επίσης η ομάδα Article 29 Data Protection Working Party και η εποπτική αρχή European Data

Protection Supervisor προτείνουν μια βαθύτερη ανάλυση των κοινωνικών και ηθικών συνεπειών από την χρήση των δεδομένων (Alessandro Montelero. 2018).

14.4. Περίληψη

Μεσούσης της συγγραφής της παρούσας εργασίας τέθηκε σε ισχύ και στην χώρα μας ο νέος κανονισμός προστασίας των προσωπικών δεδομένων. Στο 35^ο άρθρο του κανονισμού γίνεται λόγος για την ανάγκη σύνταξης μελέτης για τον αντίκτυπο της επεξεργασίας των προσωπικών δεδομένων όταν αυτά φορούν συστηματική και εκτενή αξιολόγηση προσωπικών πτυχών ή επεξεργασία δεδομένων μεγάλης κλίμακας. Από την έρευνα σε παρελθούσες γνώμες της EGE ή και άλλων ιδιωτικών φορέων που έχουν γνωστοποιήσει τις γνώμες στους στα αρμόδια όργανα της ΕΕ, είναι σαφές ότι **οι κίνδυνοι που κρύβει η τεχνητή νοημοσύνη ήταν ήδη γνωστοί στην ΕΕ**, άρα η μετάθεση της ευθύνης της σύνταξης μελέτης αντικτύπου στον υπεύθυνο επεξεργασίας γίνεται για να δώσει τον απαραίτητο χρόνο σε τεχνολογίες που κινούνται στα όρια της προσβολής της ανθρώπινης αξιοπρέπειας να αναπτυχθούν ανενόχλητα. Η επιφόρτιση του υπευθύνου επεξεργασίας με την υποχρέωση να προτείνει μέτρα αντιμετώπισης των κινδύνων από την επεξεργασία προσωπικών δεδομένων είναι ένα βαρύ φορτίο, δυσανάλογο των δυνατοτήτων του. Πώς θα μπορούσε ένας υπεύθυνος επεξεργασίας του δημοσίου να επισημάνει και τελικά να εμποδίσει την μεταφορά της δύναμης που δίνει η πληροφόρηση, ιδιαίτερα όταν αυτή η μεταφορά επιτρέπεται νομικά π.χ. η μεταφορά πληροφοριών για την οικονομική κατάσταση των Ελλήνων στην ευρωπαϊκή κεντρική τράπεζα; Σ αυτήν αλλά και σε άλλες περιπτώσεις (κλωνισμός εμπιστοσύνης, κίνδυνος εγκαθίδρυσης τυραννικού καθεστώτος, τεχνολογική ανεργία, κ.αλ.) όπου ο υπεύθυνος επεξεργασίας επωμίζεται ένα τεράστιο φορτίο, η ΕΕ μάλλον διέπραξε ένα λάθος που κανείς δεν θα ήθελε να σκεφτεί ότι έγινε εσκεμμένα.

15. Ο ρόλος του δημοσίου τομέα στην συγκέντρωση δεδομένων μεγάλης κλίμακας (Big Data) και η σχέση του με την τεχνητή νοημοσύνη

15.1. Εισαγωγή

Τα τελευταία χρόνια ένας νέος όρος έχει εισχωρήσει στην επιστήμη της πληροφορικής, στις εταιρείες που παρέχουν επικοινωνία, αλλά και στις συζητήσεις μεταξύ των πολιτών. Πρόκειται για τα δεδομένα μεγάλης κλίμακας (Big Data). Σύμφωνα με τον νέο κανονισμό προστασίας των δεδομένων αλλά και την ερμηνεία αυτού, δεδομένα μεγάλης κλίμακας θεωρούνται τα δεδομένα που:

- Αφορούν μεγάλο πληθυσμό πολιτών,
- Αφορούν μεγάλο σύνολο δεδομένων,
- Αφορούν μεγάλη γεωγραφική περιοχή.
- Αφορούν σε μεγάλη διάρκεια ή μόνιμο χαρακτήρα της δραστηριότητας επεξεργασίας δεδομένων (Α Π Ο Φ Α Σ Η 65/2018 ΑΠΔΠΧ).

Σε αυτό το κεφάλαιο θα διερευνηθεί η σχέση του δημοσίου τομέα με την συλλογή δεδομένων σε μεγάλη κλίμακα.

15.2. Ο ρόλος του δημοσίου τομέα στην συγκέντρωση δεδομένων μεγάλης κλίμακας (Big Data)

Ο δημόσιος τομέας είναι αυτός που έχει υπό τον έλεγχό του το σύνολο του πληθυσμού μιας χώρας σε όλο το γεωγραφικό μήκος της. Αυτό σημαίνει ότι μπορεί άμεσα, εφόσον υπάρχει η υποδομή, να δημιουργήσει ένα τεράστιο σύνολο δεδομένων για κάθε έναν ξεχωριστά πολίτη, το μεγαλύτερο δυνατό από αυτό που μπορεί κάθε ιδιωτική πρωτοβουλία να κάνει. Άρα συνεπάγεται ότι ο δημόσιος τομέας είναι η μοναδική οδός για κάποιον που επιθυμεί να έχει στα χέρια του τα δεδομένα όλων των πολιτών σε όλη τη γεωγραφική έκταση μιας χώρας. Επίσης η δυνατότητα της επιβολής του δημοσίου στους πολίτες δίνει την δυνατότητα το σύνολο των δεδομένων να συγκεντρωθεί υποχρεωτικά, ακόμη και χωρίς την συναίνεση του πολίτη, χωρίς κανείς να μπορεί να αρνηθεί δίχως να έχει συνέπειες. Όλες αυτές τις δυνατότητες δεν τις έχει ο ιδιωτικός τομέας, που ναι μεν μπορεί να δημιουργήσει μια βάση με δεδομένα μεγάλης κλίμακας αλλά όχι στο μέγεθος και την ποιότητα που το δημόσιο μπορεί να κάνει. Η αίσθηση ότι μια τέτοια μελλοντική κοινωνία είναι ήδη στους στόχους αυτών που κυβερνούν ενισχύεται από την πορεία που ακολουθεί η ΕΕ: στην τελευταία αναφορά της αρμόδιας ομάδας για την ηθική, Ethics Advisory Group, εκφράζεται η άποψη ότι “οι μεμονωμένοι πολίτες θα κυβερνούνται ολοένα και περισσότερο με βάση τα δεδομένα που αυτοί θα παράγουν” (πηγή https://edps.europa.eu/sites/edp/files/publication/18-01-25_eag_report_en.pdf). Αυτού του είδους η διακυβέρνηση προϋποθέτει την ύπαρξη μιας υποδομής που

συλλέγει διαρκώς δεδομένα από τους πολίτες, παρόμοια με αυτή του Κινεζικού κράτους όπως έχει ήδη αναφερθεί.

15.3. Η σχέση των δεδομένων μεγάλης κλίμακας με την τεχνητή νοημοσύνη

Ποιος θα είναι όμως ο ρόλος της τεχνητής νοημοσύνης σε ένα τέτοιο εγχείρημα; Η τεχνητή νοημοσύνη θα βρίσκεται πίσω από την **συγκέντρωση**, την **ανάλυση** των δεδομένων μεγάλης κλίμακας και εν τέλη στην **αυτοματοποιημένη λήψη αποφάσεων**. Ακόμη και πριν την ενέργεια του νέου κανονισμού είχαν εκδηλωθεί ανησυχίες για τον αντίκτυπο της επεξεργασίας τέτοιων δεδομένων στις κοινωνικές και ηθικές αξίες και ήδη έχουν διατυπωθεί ειδικά μοντέλα ανάλυσης και διαχείρισης του κινδύνου αυτού (Alessandro Mantelero. 2018).

Ωστόσο, η συγκέντρωση των δεδομένων μεγάλης κλίμακας μπορεί να επιτευχθεί με την σύμπραξη δημοσίου και ιδιωτικού τομέα, δηλαδή ο δημόσιος τομέας να χρησιμοποιήσει δεδομένα που θα συλλέγονται από ιδιωτικές επιχειρήσεις. Στην ερώτηση γιατί ένα κράτος, παρόλη την δύναμη που έχει να δημιουργεί μεγάλες δομές, να απευθυνθεί σε έναν ιδιωτικό φορέα για να αντλήσει από εκεί δεδομένα, θα μπορούσαν να δοθούν δύο ενδεχόμενες απαντήσεις: πρώτον γιατί η αγορά των δεδομένων είναι πιο συμφέρουσα και δεύτερον γιατί ενδεχομένως υπήρξαν κάποιου είδους πολιτικά ανταλλάγματα. Με δεδομένο ότι ήδη στην αγορά υπάρχουν κολοσσιαίες επιχειρήσεις που συγκεντρώνουν προσωπικά δεδομένα εκατομμυρίων χρηστών από όλο το πλανήτη μια τέτοιου είδους συναλλαγή είναι εφικτή. Για επιμέρους θέματα και εξειδικευμένα δεδομένα υπάρχουν ήδη τέτοιες συνεργασίες, όμως ακόμη δεν έχουμε φτάσει σε σημείο όπου μια μόνο εταιρεία θα αναλάβει την διαχείριση των προσωπικών δεδομένων για όλους τους πολίτες ενός κράτους.

Επίσης υπάρχει και η πραγματικότητα της επόπτευσης από κυβερνητικούς φορείς των μέσων κοινωνικής δικτύωσης για λόγους αστυνόμευσης και πρόληψης εγκλημάτων (Euripidis Loukis, Yannis Charalabidis, Aggeliki Androutsopoulou. 2017). Τέτοιο παράδειγμα είναι η χρησιμοποίηση των συστημάτων τεχνητής νοημοσύνης του κοινωνικού δικτύου Twitter από το Καναδικό κράτος για την ανίχνευση εστιατορίων με κακές συνθήκες υγιεινής.

Συμπερασματικά μπορούμε να πούμε ότι η τεχνητή νοημοσύνη είναι ένα εργαλείο στα χέρια του δημοσίου τομέα ώστε να λειτουργεί και να εποπτεύει μια κοινωνία που θα κυβερνάται από τα δεδομένα που θα παράγει.

15.4. Περίληψη

Τα τελευταία χρόνια γίνεται συχνά λόγος για την συγκέντρωση δεδομένων μεγάλης κλίμακας. Σαν δεδομένα **μεγάλης κλίμακας** θεωρούνται αυτά που αφορούν: μεγάλο πληθυσμό πολιτών, μεγάλο σύνολο δεδομένων, μεγάλη γεωγραφική περιοχή, μεγάλη διάρκεια ή μόνιμο χαρακτήρα της δραστηριότητας επεξεργασίας δεδομένων. Η συγκέντρωση δεδομένων από το δημόσιο πληροί πάντοτε ένα από τα παραπάνω κριτήρια, και συνεπώς είναι συγκέντρωση δεδομένων μεγάλης κλίμακας. Ο ρόλος που θα παίζει στο μέλλον η τεχνητή νοημοσύνη είναι η συγκέντρωση, η επεξεργασία αυτών των δεδομένων και ίσως και η αυτοματοποιημένη λήψη αποφάσεων.

16. Προτάσεις για περαιτέρω έρευνα και βελτίωση

Η μελέτη γύρω από την τεχνητή νοημοσύνη αλλά και την ηθική γέννησε κάποιες ιδέες που θα μπορούσαν στο μέλλον να υλοποιηθούν. Προϋπόθεση βέβαια της χρήσης μιας ιδέας είναι να βρεθεί μια χρυσή τομή με το πεδίο της προστασίας των δικαιωμάτων του ανθρώπου. Αυτές οι ιδέες (Α έως ΣΤ) παρουσιάζονται παρακάτω.

A. Πρόταση για τον τομέα της άμυνας

Η θεώρηση του τρόπου που λειτουργούν σήμερα τα οπλικά συστήματα γέννησε την ιδέα για την ανάγκη τήρησης ενός ηθικού πρωτοκόλλου στην χρήση των οπλικών συστημάτων. Ακόμη και στην συνθήκη για τα συμβατά όπλα (Convention on Conventional Weapons) αλλά και στο Πρωτόκολλο I τη συνθήκης της Γενεύης, δεν θα βρει κανείς τον τρόπο που τα συμβατά όπλα πρέπει να χρησιμοποιούνται αλλά μόνον γενικές απαγορεύσεις (π.χ. άρθρα 51, 54 του Πρωτοκόλλου I: οι επιθέσεις σε πολίτες μη στρατιωτικούς είναι παράνομες).

Προς τη κατεύθυνση μιας συνειδητής χρήσης των οπλικών συστημάτων είναι και η συμφωνία που έγινε στην συνδιάσκεψη των Ηνωμένων Εθνών που αφορούσε την συμφωνία για τα Συμβατά Όπλα (Convention on Certain Conventional Weapons) συμφωνήθηκε η αρχή “meaningful human control for LAWS”. Σαν έννοια για πρώτη φορά έχει προταθεί από την Βρετανική μη κυβερνητική οργάνωση Article 36 (πηγή <https://aticdn.s3-eu-west-1.amazonaws.com/2018/03/140318-Ai-ethics-and-the-law-public-panel-report.pdf>). Στην συμφωνία αυτή επισημαίνεται ότι πάντοτε και σε κάθε ξεχωριστή δράση ο άνθρωπος πρέπει να κρατά τον ακριβή και ουσιαστικό-συνειδητό έλεγχο των αυτόνομων οπλικών συστημάτων.

Όμως ακόμη μια συμφωνία παρέμεινε σε γενικό επίπεδο και δεν προχώρησε στην απάντηση πως αυτή θα υλοποιηθεί στην πράξη. Μια μικρή αναφορά στον τρόπο που σήμερα πλήττεται ένας στόχος, ιδιαίτερα από απόσταση με πυραύλους που προκαλούν καταστροφές μεγάλου μεγέθους και πλήττουν αδιάκριτα, δείχνει το ηθικό κενό που επικρατεί: το επιτιθέμενο μέλος αποφασίζει να πλήξει έναν στόχο και το κάνει χωρίς να κρατείται κάποιου είδους ιστορικού και δέσμευσης για το πώς επιλέχθηκε ο στόχος και ποιο ήταν το επιδιωκόμενο αποτέλεσμα και το πιο σημαντικό με ποιο τρόπο η “διάκριση” μεταξύ αμάχων και στρατιωτικά εμπλεκόμενων εξασφαλίζεται όπως οι νόμοι ορίζουν. Έτσι βλέπουμε πολλές φορές αντίπαλες δυνάμεις να αλληλοκατηγορούνται για την πατρότητα ενός πυραύλου που έπληξε αμάχους το οποίο απαγορεύεται από τον Οργανισμό Ηνωμένων Εθνών και την Δήλωση για την προστασία γυναικών και παιδιών σε έκτακτες συνθήκες πολέμου (1974 United Nations, Declaration on the Protection of Women and Children in Emergency and Armed Conflict).

Για αυτό το λόγο προτείνεται η εξής ηθική αρχή: **να υπάρχει δημοσιότητα και διαφάνεια στον τρόπο που πλήττεται ένας στόχος και να κρατείται, από ουδέτερο τρίτο, ένα ιστορικό λήψης αποφάσεων, στη συνέχεια να συγκρίνονται**

τα αποτελέσματα της επίθεσης με τα επιδιωκόμενα αποτελέσματα. Όσο τα αυτόνομα οπλικά συστήματα προχωρούν η ανιχνευσιμότητα αυτή θα είναι πολύ πιο εύκολο να επιτευχθεί, μιας και η λήψη αποφάσεων μπορεί αυτόματα να καταγράφεται, αλλά ταυτόχρονα να καταγράφονται και τα αποτελέσματα μέσω μιας πιθανής ενοποίησης των ραντάρ με κατάλληλο λογισμικό τεχνητής νοημοσύνης, κάτι που δεν είναι εφικτό όταν οι αποφάσεις παίρνονται από ανθρώπους. Αν υιοθετηθεί αυτός ο τρόπος από το σύνολο των κρατών θα λειτουργεί παράλληλα και σαν νομικό εφόδιο, γιατί καμιά αδίκως επιτιθέμενη δύναμη δεν θα απευθυνόταν στον ουδέτερο τρίτο πριν επιτεθεί. Το σημερινό πρωτόκολλο εμπλοκής αρκείται μόνο στην “κήρυξη” ή την “δήλωση” ενός πολέμου και όχι στην διαφάνεια της στρατηγικής και τακτικής μιας επιτιθέμενης δύναμης. Επίσης υπήρξαν πόλεμοι που έγιναν χωρίς να κυρηχθούν όπως η επίθεση στη Λιβύη. Δηλαδή η κήρυξη του πολέμου αναφέρεται στο τι πρέπει να γίνεται σε έναν πόλεμο αλλά δεν αναφέρεται στο αν και στο πως πρέπει αυτό να γίνεται. Να σημειωθεί εδώ ότι το νομικό πλαίσιο γύρω απ τους κανόνες του πολέμου είναι σχεδόν χαοτικό και αποσπασματικό, αφού συντίθεται από συνθήκες και δηλώσεις που υιοθετήθηκαν και υπάρχουν ακόμη και από το 1856 έως σήμερα. Επιπλέον δίνει την δυνατότητα σε κράτη να μην υπογράφουν κάποιες από τις συνθήκες χωρίς αυτό να επιφέρει συνέπειες για αυτά. Σήμερα δεν υπάρχει ένα ενιαίο νομικό πλαίσιο στο οποίο μπορεί κάποιος να βρει απαντήσεις σε νομικά ζητήματα γύρω από τους πολέμους.

Είναι προφανές ότι το ξεκαθάρισμα των ηθικών αρχών και την προστασία τους μέσω της νομικής περιφρούρησης βαρύνει το δημόσιο τομέα και στην προκειμένη περίπτωση έναν διεθνή οργανισμό που θα φέρει την αποδοχή όλων των κρατών, όπως π.χ. τον ΟΗΕ ή το ΝΑΤΟ. Παράλληλα θα πρέπει όλα τα αυτόνομα οπλικά συστήματα να βρίσκονται κάτω από την επιτήρηση ενός μοναδικού φορέα, πράγμα το οποίο θα αντιμετωπίσει πολλά εμπόδια στην υλοποίηση του λόγω της σημερινής πραγματικότητας.

Η πραγματικότητα γύρω από τα αυτόνομα οπλικά συστήματα είναι διαφορετική: Πρώτον τα αυτόνομα οπλικά συστήματα ήδη παράγονται και χρησιμοποιούνται (π.χ. μη επανδρωμένα κατασκοπευτικά ή βομβαρδιστικά αεροσκάφη) και δεύτερον η τελευταία σύσκεψη για τα αυτόνομα οπλικά συστήματα (ονομαζόμενα και killer robots) που έγινε τον Νοέμβριο του 2018 στην Γενεύη δεν κατέληξε σε μια ομόφωνη συμφωνία, αλλά παρέπεμψε τις αποφάσεις στη νέα χρονιά, επειδή η αποδοχή μιας συμφωνίας απαιτεί την συνολική αποδοχή μιας συνθήκης από όλα τα κράτη μέλη. Τα κράτη Ρωσία, Ισραήλ, Αυστραλία, Νότια Κορέα, και οι Ηνωμένες Πολιτείες Αμερικής δήλωσαν ότι δεν μπορούν να υποστηρίξουν μια νέα συμφωνία του CCW και έτσι η σύσκεψη απέβη άκαρπη. Αυτή η στάση των ισχυρών κρατών να αρνούνται την υπακοή σε ηθικές αρχές δείχνει και την πορεία που θα ακολουθήσει η τεχνητή νοημοσύνη στο μέλλον, ότι δηλαδή τα μελλοντικά οπλικά συστήματα δεν είναι σίγουρο ότι θα φέρουν ένα ηθικό φίλτρο και πιθανώς να είναι προγραμματισμένα (άρα και μη αυτόνομα) να εξυπηρετούν σκοπούς που προσβάλλουν τα ανθρώπινα δικαιώματα και εισπράττουν μια διεθνή διαμαρτυρία.

B. Πρόταση για την δικαιοσύνη

Αναφέρθηκε και σε προηγούμενο κεφάλαιο ότι η σημερινή τεχνολογία έχει φτάσει στο σημείο να μπορεί να διεισδύει στον ψυχικό κόσμο ενός ανθρώπου, και να αναγνωρίζει τα συναισθήματά του. Εφόσον συντρέχουν οι κατάλληλες προϋποθέσεις η δυνατότητα αυτή θα μπορούσε να φανεί χρήσιμη στην απονομή της δικαιοσύνης. Η ανίχνευση συναισθημάτων μέσω της κατόπτευσης θα μπορούσε να δώσει στους δικαστές μια επιπλέον ένδειξη για την αλήθεια ή το ψεύδος των λεγομένων ενός διαδίκου. Η τεχνητή νοημοσύνη θα είναι ο νέος ‘‘όρος της αλήθειας’’. Οι απαραίτητες προϋποθέσεις θα πρέπει να είναι η συγκατάβαση του υποκειμένου αλλά και να εξασφαλίζεται η διαφανής λειτουργία της εφαρμογής και η διαφανής διαδικασία λήψης της απόφασης από την εφαρμογή, ώστε να μην έχουμε το φαινόμενο που εμφανίστηκε στα δικαστήρια των ΗΠΑ (κρυφή και αδιαφανής χρήση λογισμικού). Η χρήση αυτής της εφαρμογής θα είναι επιθυμητή όχι μόνο από τους δικαστές αλλά και από αυτούς που επιθυμούν να αποδείξουν ότι πράγματι δεν βαρύνονται με κάποιο αδίκημα. Επίσης ο ρόλος της εφαρμογής δεν θα πρέπει να είναι δεσμευτικός προς τους δικαστές αλλά μόνον συμβουλευτικός ώστε να μην φτάσουμε στο γεγονός μηχανές να δικάζουν ανθρώπους.

Γ. Πρόταση για ανοιχτή εκλογική δημοκρατική διαδικασία

Εφόσον η διαδικασία της δημιουργίας μιας ψηφιακής ταυτότητας για όλους τους πολίτες προχωρήσει, θα πρέπει αυτό να χρησιμοποιηθεί και προς την καλή κατεύθυνση. Δεν θα πρέπει μόνο η ψηφιακή ταυτότητα να χρησιμοποιείται από τις κυβερνήσεις για την υλοποίηση φορολογικών και άλλων top-down προγραμμάτων αλλά και για την υλοποίηση μιας δομής citizen to government. Η τεχνητή νοημοσύνη θα μπορούσε να φανεί χρήσιμη στην αληθινή δημοκρατία, αφού ο κάθε ένας ξεχωριστά πολίτης θα μπορεί σε μια ενιαία πληροφοριακή δομή να κάνει τις παρακάτω ενέργειες:

Πρώτον, να αναδεικνύει προς ψήφιση την δική του πρόταση που θα είναι ορατή σε όλους τους πολίτες και

Δεύτερον, να συμμετέχει σε όλες τις ψηφοφορίες που άλλοι ή η κυβέρνηση ανέδειξαν/ε και

Τρίτον, η ψηφοφορία που θα γίνεται μέσα από αυτή τη δομή να γίνεται σεβαστή και να αποτελεί και νόμο της πολιτείας, και

Τέταρτον η γνώμη του κάθε ενός θα είναι διαφανής ώστε δεν θα μπορεί κανείς να κρύβεται πίσω από την μυστικότητα της ψήφου.

Με αυτό τον τρόπο η γνώμη της πλειοψηφίας των πολιτών θα γίνεται πάντοτε σεβαστή και παράλληλα δεν θα υπάρχει το πρόβλημα της ελλιπούς αντιπροσώπευσης και της λήψης αποφάσεων που δεν πέρασαν ποτέ από την κρίση των πολιτών, όπως

γίνεται σήμερα. Η λειτουργία ενός τέτοιου συστήματος προϋποθέτει την ύπαρξη ενός πληροφοριακού συστήματος με αυστηρή ασφάλεια και διαφάνεια.

Δ. Πρόταση για την σύσταση μιας επιτροπής ηθικής

Έχει ήδη αναφερθεί και αναλυθεί λεπτομερώς σε προηγούμενο κεφάλαιο η ιδέα της δημιουργίας μιας επιτροπής ηθικής που θα είναι η πρώτη αρχή με την ανώτατη εξουσία να δέχεται ή να απαγορεύει την ένταξη μιας εφαρμογής της τεχνητής νοημοσύνης σε ένα κράτος. Προφανώς πρόκειται για μια πρόταση γενικότερης στρατηγικής και όχι για μια τεχνική υλοποίηση όπως οι άλλες προτάσεις. Αναφέρεται ξανά σε αυτό το κεφάλαιο μόνο για λόγους μιας συγκεντρωμένης παρουσίασης των προτάσεων.

Ε. Πρόταση για χρήση της τεχνητής νοημοσύνης στην καταπολέμηση της τρομοκρατίας

Έχει αναφερθεί σε προηγούμενο κεφάλαιο η δυνατότητα που δίνει πλέον η τεχνολογία για την αναγνώριση των συναισθημάτων μέσω της ανάλυσης φωνής και της ανάλυσης του προσώπου. Μια ακόμη χρήση αυτής της τεχνολογίας είναι η χρήση της για την ενίσχυση της ασφάλειας εναντίον της τρομοκρατίας. Εδώ το ηθικό δίλλημα είναι μεγάλο και εγείρει πολλά ηθικά ζητήματα: αξίζει μια παγκόσμια ηλεκτρονική ταυτοποίηση όλου του πληθυσμού μέσω των αστυνομικών αρχών για χάρη πρόληψης και σύλληψης μερικών ελαχίστων επικίνδυνων ατόμων; Μήπως η καταπολέμηση της τρομοκρατίας είναι μια πρόφαση που πίσω της κρύβει ένα μεγαλεπήβολο σχέδιο μιας καχέκτυπης παγκόσμιας ηλεκτρονικής διακυβέρνησης; Κατά πόσον η παγκόσμια ταυτοποίηση θίγει την προσωπική ελευθερία; Επιπλέον, είναι επικίνδυνο όλος αυτός ο όγκος χρήσιμων πληροφοριών να είναι συγκεντρωμένος στα χέρια λίγων ατόμων; Μια λύση που θα μπορούσε να προταθεί ώστε να μην γίνεται διαρκώς σε όλους τους πολίτες παρακολούθηση μέσω των καμερών, θα ήταν να γίνεται μόνο για μια φορά ένας ανωνυμοποιημένος έλεγχος μόνον κατά την είσοδο ενός ταξιδιώτη στην χώρα.

ΣΤ. Πρόταση για την διαφάνεια σε οικονομικά θέματα

Τα τελευταία χρόνια στην χώρα μας ολοένα και περισσότερο οι πολίτες ωθούνται είτε από τον ιδιωτικό τομέα (π.χ. με παροχή εκπτώσεων στις τιμές των προϊόντων με τη χρήση καρτών) είτε από το δημόσιο (παροχή έκπτωσης φόρου για ηλεκτρονικές πληρωμές) στην αύξηση της χρήσης του ηλεκτρονικού χρήματος εις βάρος του φυσικού χρήματος. Επειδή οι ηλεκτρονικές υπογραφές, με τις οποίες υπογράφονται τα αντίγραφα του ηλεκτρονικού χρήματος έχουν το ρίσκο του έμπιστου τρίτου και επιπλέον περνούν μέσω της εποπτείας μιας ιδιωτικής τράπεζας

(της Τράπεζας της Ελλάδος) θα πρέπει να δημιουργηθεί ένας δημόσιος φορέας που θα εποπτεύει όλη την κίνηση του ηλεκτρονικού χρήματος. Η δυνατότητες που δίνει η τεχνητή νοημοσύνη για την ανάλυση τεραστίων ποσοτήτων δεδομένων θα μπορούσε να χρησιμοποιηθεί ώστε, όταν όλο το χρήμα γίνει ηλεκτρονικό, να δίνει ανά πάσα στιγμή το πόθεν έσχες όλων των πολιτών. Επιπλέον, αν σε αυτό το σύστημα δοθεί διαφάνεια, θα μπορούμε να εντοπίσουμε κάθε παράνομη συναλλαγή, αλλά και την ιστορία του πως απέκτησε ο καθένας την περιουσία του. Αν όμως το σύστημα αυτό έχει σκοτεινά σημεία όπως, το να αφήνεται ο έλεγχος σε ιδιωτικές τράπεζες ή να μην μπορεί ο κάθε πολίτης να ελέγξει την περιουσία του άλλου, τότε το χρηματοοικονομικό σύστημα θα γίνει περισσότερο επικίνδυνο από ότι είναι σήμερα.

17. Επίλογος

17.1 Σύνοψη και συμπεράσματα

Είναι αδιαμφισβήτητο ότι η τεχνητή νοημοσύνη θα φέρει τεράστιες αλλαγές στην ζωή μας λόγω των τεράστιων δυνατοτήτων που υπόσχεται. Η διαδικασία της μάθησης που ακολουθεί και προσομοιάζει αυτή των ανθρώπων είναι ο θεμέλιος λίθος για μια ραγδαία ανάπτυξη. Στη συνέχεια θα γίνει μια προσπάθεια να δούμε αν υπάρχει μια ισόρροπη ανάπτυξη σε όλους τους τομείς που μπορεί να αναπτυχθεί η τεχνητή νοημοσύνη, υπό την αιγίδα του δημοσίου τομέα.

Σύμφωνα με την παρουσίαση που έγινε στα προηγούμενα κεφάλαια οι εφαρμογές με τον μεγαλύτερο όγκο δεδομένων και την χρήση αυτών για έλεγχο των πολιτών εντοπίζονται σε κράτη όπου τα πολιτεύματα αποκλίνουν πολύ από την δημοκρατία, όπως στο κράτος της Κίνας και της Ινδίας. Ενώ παράλληλα ελπιδοφόρες εφαρμογές όπως η πρόγνωση ασθενειών, η ηλεκτρονική συμμετοχή στις πολιτικές αποφάσεις, δεν έχουν αναπτυχθεί σε ικανοποιητικό βαθμό, γιατί δεν τους έχει δοθεί η πρωτεύουσα σημασία. Επιπλέον, σε καμιά οργανωμένη πολιτεία δεν έχει ξεκινήσει μια σοβαρή συζήτηση για το ποιος θα καρπωθεί τα οικονομικά οφέλη που θα προκύψουν από την υιοθέτηση της τεχνητής νοημοσύνης. Η απάντηση σε αυτές τις διαπιστώσεις ίσως να βρίσκεται στην πρωιμότητα της μελέτης της ανάπτυξης της τεχνητής νοημοσύνης αλλά ίσως και στην έλλειψη ενός ηθικού μανδύα αυτών που είναι αρμόδιοι για την λήψη πολιτικών αποφάσεων.

Επιχειρώντας ένα γενικό συμπέρασμα σαν αποτέλεσμα της έρευνας αυτής θα μπορούσαμε να πούμε ότι ο δημόσιος τομέας πριν υιοθετήσει μια εφαρμογή τεχνητής νοημοσύνης πρέπει να απαντήσει στο βασικό ερώτημα “αυτή η εφαρμογή θα κάνει την ζωή των πολιτών καλύτερη ή όχι;”. Σε περίπτωση που μια εφαρμογή έχει και καλά και κακά αποτελέσματα θα πρέπει να αποφασιστεί προς ποια πλευρά θα δείξουμε ανοχή μετρώντας βέβαια και τις ανάλογες συνέπειες. Οι αποφάσεις αυτές θα πρέπει να βαρύνουν είτε μια αρμόδια επιτροπή ηθικής ή τους ίδιους τους πολίτες αφού πρώτα πρέπει να ενημερωθούν για τις συνέπειες της απόφασής τους.

17.2 Όρια και περιορισμοί της έρευνας

Κατά τον χρόνο που γράφεται αυτή η έρευνα υπάρχει μια μεγάλη κινητικότητα γύρω από την τεχνητή νοημοσύνη και κυρίως γύρω από την δημιουργία μιας ηθικής διεπαφής. Τα τελευταία δυο χρόνια συστάθηκαν πολλές ομάδες που έχουν σαν πεδίο ευθύνης την περιφρούρηση της ηθικής από τις εφαρμογές της τεχνητής νοημοσύνης, με κυριότερες, και σε ανώτερο βαθμό, αυτές τις ΕΕ. Βρισκόμαστε δηλαδή στην στιγμή που ακόμη υπάρχει μια δυναμική διαδικασία γύρω από την μελέτη της ηθικής αλλά ακόμη δεν έχουν παρθεί πολιτικές αποφάσεις για την στάση που πρέπει να κρατήσουμε απέναντι σε εφαρμογές που εγείρουν σοβαρά ερωτήματα για την ζωή μας. Αυτό το συμπέρασμα είναι μια καταληκτική κατάσταση

η οποία αποτελεί και ένα τέρμα στην αναζήτηση των ευθυνών αλλά και των υπευθύνων. Όμως δεν αποτελεί και το τέλος της έρευνας γύρω από την ανάπτυξη της τεχνητής νοημοσύνης, πράγμα που δίνει την ευκαιρία να αναπτυχθούν εφαρμογές τεχνητής νοημοσύνης με αμφιλεγόμενη ηθική υπόσταση. Με δεδομένη αυτή την κατάσταση που ακόμη υπάρχει μια μη οριοθετημένη νομικά περιοχή, είναι δύσκολη η πλήρης χαρτογράφηση των δυνατοτήτων που μπορεί να έχει η τεχνητή νοημοσύνη μιας και η μυστικότητα μπορεί να καλύπτει πολλές προθέσεις και άλλες τόσες εφαρμογές που δεν αποσκοπούν στο γενικό καλό. Προγράμματα όπως η διεπαφή χρηστών –υπολογιστή της εταιρείας κοινωνικής δικτύωσης Facebook αλλά και προγράμματα όπως το Brain της Google, αλλά και τα ASIST και μBRAIN της στρατιωτικής Darpa, κινούνται στο μεταίχμιο της ηθικής αλλά παράλληλα αποτελούν και μια μεγάλη τεχνολογική πρόοδο, που ο δημόσιος τομέας θα μπορούσε στο μέλλον να δει σαν ένα χρήσιμο προϊόν.

17.3 Μελλοντικές Επεκτάσεις

Αναφέρθηκε σε προηγούμενο κεφάλαιο ότι η έννοια της τεχνητής νοημοσύνης είναι μια δυναμική έννοια. Αυτό σημαίνει ότι αυτό που σήμερα αντιλαμβανόμαστε σαν τεχνητή νοημοσύνη σίγουρα στο μέλλον θα είναι μια ξεπερασμένη έννοια. Έτσι η τεχνολογία δεν θα πάψει να προκαλεί το ενδιαφέρον όλων των πολιτών και να αλλάζει την κοινωνική ζωή.

Το βήμα προς το μέλλον έχει ήδη γίνει με την συζήτηση, αν και αρνητική προς το παρόν, που έχει ξεκινήσει γύρω από την πιθανή ανάπτυξη **συνείδησης** από τις μηχανές (Ray Kurzweil, Viking Adult, DrewMc_Dermott. 2006). Αυτό θα αποτελέσει και ένα μεγάλο πεδίο έρευνας που πιθανόν στο μέλλον να δώσει μια νέα τεχνολογική επανάσταση παρόμοια με αυτή που σήμερα επικαλείται η τεχνητή νοημοσύνη. Στην διεθνή έρευνα γύρω από την τεχνητή νοημοσύνη γίνεται ήδη λόγος ότι η τεχνητή νοημοσύνη αφού έχει ξεπεράσει το πρώτο στάδιο ανάπτυξης την Narrow Inteligence βρίσκεται ήδη στο στάδιο της General Inteligence και το βλέμα της στρέφεται στην κατάκτηση της Super Inteligence, και την ανάπτυξη συνείδησης (RomanOrmandy. 2019).

Εκτός βέβαια από τον πρώτο παράγοντα εξέλιξης, την νοημοσύνη των μηχανών, υπάρχει και ο παράγοντας άνθρωπος, σαν φορέας κάποιων ηθικών αρχών. Είναι λοιπόν και η αντίληψη της ηθικής που μπορεί να εξελιχθεί και να δώσει μια αλλαγή στην σχέση της με την τεχνολογία.

Επιπλέον, το χρονικό διάστημα που μεσολαβεί από τον χρόνο που υιοθετείται μια εφαρμογή μέχρι τον χρόνο που φαίνονται οι μακροπρόθεσμες επιρροές της στον άνθρωπο μπορεί να είναι και μερικές δεκάδες χρόνια. Αυτό σημαίνει ότι η ηθική αντίληψη μπορεί να μεταβληθεί εφόσον ένα αποτέλεσμα κριθεί κακό ή να παραμείνει ίδια και να βελτιωθεί αν ένα αποτέλεσμα κριθεί καλό για την κοινωνική ζωή, π.χ. η χρήση πλαστικών δοχείων τροφίμων έχει αρχίσει να προκαλεί ανησυχία για την

υγεία των χρηστών κάτι που αναπτύχθηκε δεκαετίες μετά την πρώτη γνωριμία. Επομένως στο μέλλον ίσως μετά από μια αναθεώρηση της χρησιμότητας μιας τεχνολογικής εφαρμογής να οπισθοδρομήσουμε ή να αντίθετα να προχωρήσουμε σε κάτι πιο ανεπτυγμένο.

Δεν είναι μόνον όμως οι παράγοντες τεχνητή νοημοσύνη και ηθική που μπορούν να αλλάξουν στο μέλλον. Ακόμη και η έννοια του δημοσίου μπορεί να μεταβληθεί είτε μέσω μιας πολιτικής ενοποίησης π.χ. της ΕΕ είτε μέσω μιας διάσπασης, είτε μέσω μιας αλλαγής του πολιτεύματος. Μια μεταβολή λοιπόν της έννοιας του δημοσίου και το πώς αυτή θα επηρεάσει τη ηθική και την χρήση της τεχνολογίας, είναι ένα μεγάλο πεδίο έρευνας που στο μέλλον μπορεί να απασχολήσει το ενδιαφέρον της επιστήμης.

Αν θεωρήσουμε ότι πράγματι η τεχνητή νοημοσύνη είναι μια νέα τεχνολογική επανάσταση θα πρέπει να σκεφτούμε την προσαρμογή στις σημερινές συνθήκες της ανησυχίας του οικονομολόγου Wassily Leontief, όπου το 1983 είπε ότι αν κάποιος νομίζει ότι μετά την βιομηχανική επανάσταση θα βρει εύκολα εργασία είναι σαν να πιστεύει ότι τα άλογα θα ξαναβρούν την εργασία τους μετά την ανάπτυξη των οχημάτων. Αφού λοιπόν η τεχνητή νοημοσύνη πολιορκεί τον άνθρωπο ποιο θα είναι το ανάλογο των αλόγων της τότε εποχής ή το ανάλογο των κινδύνων;

Παράρτημα 1. Εξακρίβωση επαναληψιμότητας

Ο οδηγός που ακολουθήθηκε για την συγγραφή της διτριβής (Richard T. Watson, Jane Webster 2002)) επιβάλλει την καταγραφή όλων των αναζητήσεων που έγιναν, επίσης το πού έγιναν και ποιά ήταν τα αποτελέσματα των αναζητήσεων. Για αυτό το λόγο στον επόμενο πίνακα θα παρουσιαστούν όλα τα λήμματα που αναζητήθηκαν και τα αποτελέσματα που βρέθηκαν καθώς και που έχει γίνει αναφορά στο κείμενο της εργασίας. Για λόγους διατήρησης του μεγέθους σε λογικά επίπεδα δεν ήταν δυνατό να αναφερθούν ευρήματα που δεν χρησιμοποιήθηκαν, αυτό θα ήταν αδύνατο (π.χ μόνον στο λήμμα “artificial intelligence” η βάση sciencedirect δίνει 131 αποτελέσματα). Ο πίνακας αυτός είναι ο Πίνακας Επαναληψιμότητας γιατί εξυπηρετεί την ανάγκη να μπορεί ο αναγνώστης να επαναλάβει τις αναζητήσεις.

Πίνακας επαναληψιμότητας.

Λήμμα.	Που αναζητήθηκε;	Ευρήματα.	Αναφορά στην διατριβή.
History of Artificial intelligence	https://www.aaai.org	https://www.aaai.org/ojs/index.php/aimagazine/article/view/1848/1746	Ναι
Ιλιάδα	διαδίκτυο	http://ebooks.edu.gr/modules/ebook/show.php/DSGYM-B108/208/1441,4841/	ναι
Οδύσσεια	διαδίκτυο		ναι
Understanding the fear anxiety	διαδίκτυο	https://www.ncbi.nlm.nih.gov/pubmed/27609244	ναι
Fear of new technologies, technophobia	διαδίκτυο	https://en.wikipedia.org/wiki/Technophobia	όχι
Λήμμα.	Που αναζητήθηκε;	Ευρήματα.	Αναφορά στην διατριβή.
Writing a Literature Review	scholar.google.gr	http://www.jstor.org/stable/4132319?seq=1#page_scan_tab_contents	ναι
		https://www.lib.uoguelph.ca/get-assistance/writing/specific-types-papers/writing-literature-review	όχι
Λήμμα.	Που αναζητήθηκε;	Ευρήματα.	Αναφορά στην διατριβή.
ethics on artificial intelligence	sciencedirect	https://www.sciencedirect.com/science/article/pii/S0267364916300814	ναι
	διαδίκτυο	file:///C:/Users/user/Downloads/934-4279-1-PB.pdf	ναι
What ethics is	scholar.google.gr	https://books.google.gr/books?hl=el&lr=&id=doxKfg6gR1AC&oi=fnd&pg=PP1&dq=what++ethics+is&ots=5Dgrc0GAHT&sig=HRo88rdiyMiGqCQ4k6BwS0GyyI&redir_esc=y#v=onepage&q=what%20%20ethics%20is&f=false	ναι
ethics	scienced	https://ac.els-cdn.com/S0142694X17300662/1-s2.0-S0142694X17300662-	

	irect	main.pdf?_tid=2874596f-893f-4410-91f2-d9fc051a81e2&acdnt=1521228509_f2de9a99ab9ed55bdf5b25d24a3f48a2	
ethics	sciencedirect	https://ac.els-cdn.com/S0265964616300170/1-s2.0-S0265964616300170-main.pdf?_tid=0466615c-2b19-4659-8b37-55a739943b04&acdnt=1521230203_7cda4996d2587993f61dfdbfe57d0655	vai
intelligence	sciencedirect	https://www.sciencedirect.com/science/article/pii/S0160289679900187	vai
Animal on us	διαδίκτυο	https://press.princeton.edu/titles/8240.html	vai
googlebrain	googlesc holar	https://www.nature.com/polopoly_fs/1.14481!/menu/main/topColumns/topLeftColumn/pdf/505146a.pdf?origin=ppub	vai
biomimicry	sciencedirect	https://ac.els-cdn.com/S0734975014001517/1-s2.0-S0734975014001517-main.pdf?_tid=5ca70ee5-b78a-4404-9ea9-8a402a947d87&acdnt=1521362897_033be97b52ca657830140c721d74e94a	vai
ethics on artificial intelligence on the public sector	scopus	Κανένα αποτέλεσμα	
ethics on artificial intelligence on the public sector	thomson reuters	Κανένα αποτέλεσμα	
ethics on artificial intelligence on the public sector	Science direct	https://www.sciencedirect.com/science/article/pii/S0267364916300814	vai
ethics on artificial intelligence on the public sector	scholar.g oogle.gr	https://www.intelligentzia.ch/blog/wp-content/uploads/The_Ethical_Frontiers_of_Robotics.pdf	
Centralization of political power	Science direct	https://www.sciencedirect.com/science/article/pii/S0264275117310016	vai
Ethics in public sector	sciencedirect	https://www.sciencedirect.com/science/article/pii/S2212567115005663	vai
ai ethics on public sector	διαδίκτυο	https://ash.harvard.edu/files/ash/files/artificial_intelligence_for_citizen_services.pdf	vai
Speech recongnition	sciencedirect	https://www.sciencedirect.com/science/article/pii/S1474667017491981	vai
artificial speech	sciencedirect	https://www.sciencedirect.com/science/article/pii/S2212683X17301214	vai
artificial speech	sciencedirect	https://www.sciencedirect.com/science/article/pii/S2212683X17301032	vai
artificial speech	sciencedirect	https://www.sciencedirect.com/science/article/pii/S1051200417302439	vai
technological unemployment	sciencedirect	https://www.sciencedirect.com/science/article/pii/S0016328716302063	vai
technological unemployment	διαδίκτυο	https://www.mckinsey.com/global-themes/future-of-organizations-and-work/what-can-history-teach-us-about-technology-and-jobs	vai
Digital identity	sciencedirect	https://www.sciencedirect.com/science/article/pii/S0267364915000035	vai
Digital identity	sciencedirect	https://www.sciencedirect.com/science/article/pii/S1353485815300696	vai
Trust AI	διαδίκτυο	http://www.research.ibm.com/software/IBMResearch/multimedia/AIEthics_Whitepaper.pdf	vai
Trust AI	διαδίκτυο	https://www.ibm.com/blogs/insights-on-business/government/can-public-sector-trust-artificial-intelligence/	vai
Chatbots	διαδίκτυο	https://www.telegraph.co.uk/technology/2017/08/01/facebook-shuts-robots-invent-	vai

overcontrol	o	language/	
Ethics robots	διαδίκτυο	https://www.cc.gatech.edu/ai/robot-lab/online-publications/arkin-rev.pdf	ναι
Why people kill	sciencedirect	https://ac.els-cdn.com/S1359178915001536/1-s2.0-S1359178915001536-main.pdf?_tid=52d43190-1227-42d8-91a3-82dd52c9ae84&acdnat=1523440668_7d8ffe355f9b2b9cfaa4acef940fbf71	όχι
Why people kill	sciencedirect	https://www.sciencedirect.com/science/article/pii/S1359178914001323	ναι
AI in health	sciencedirect	https://www.sciencedirect.com/science/article/pii/S0883540318302158	ναι
hate Speech recognition	sciencedirect	https://ac.els-cdn.com/S1877050913008727/1-s2.0-S1877050913008727-main.pdf?_tid=26c5af64-c55c-4b75-8f4d-f1199446960a&acdnat=1523775998_1ad118623e84efc7de799e2f03fdac9f	ναι
Ethics framework	διαδίκτυο	https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/524298/Data_science_ethics_framework_v1.0_for_publication_1_.pdf	ναι
AI in health	sciencedirect	https://ac.els-cdn.com/S1877042814030055/1-s2.0-S1877042814030055-main.pdf?_tid=83a39203-8d0e-41a9-9b7d-0b3c9e2b4916&acdnat=1523950483_d4a4607e99b2ae3109d0eb069f80d0b3	ναι
Viruses for Machine Deep learning	διαδίκτυο	https://arxiv.org/pdf/1802.03043.pdf	ναι
AI foe weather forecasting	διαδίκτυο	https://web.stanford.edu/class/cs221/2017/restricted/p-final/cvo9/final.pdf	ναι
Genetic engineering	sciencedirect	https://www.sciencedirect.com/science/article/pii/S1751721415300063	ναι
AI in agriculture	διαδίκτυο	https://www.techemergence.com/ai-agriculture-present-applications-impact/	Όχι
AI in agriculture	διαδίκτυο	https://www.ibm.com/blogs/watson/2016/12/five-ways-agriculture-benefit-artificial-intelligence/	Ναι
AI for traffic prediction	sciencedirect	https://www.sciencedirect.com/science/article/pii/S2405896315014202	Ναι
AI for traffic prediction	διαδίκτυο	http://analytics-magazine.org/big-data-helps-city-of-dublin-improve-public-bus-transportation-and-reduce-congestion/	ναι
Predictive policing	sciencedirect	https://www.sciencedirect.com/science/article/pii/B9780124095489096287	Ναι
Augmented intelligence	sciencedirect	https://www.sciencedirect.com/science/article/pii/S0007681318300387	ναι
Data ethics	The royal society	https://royalsocietypublishing.org/doi/full/10.1098/rsta.2016.0360#sec-1	ναι
AI will fail	Google scholar	https://books.google.gr/books?hl=el&lr=&id=e9W9m_4q4pYC&oi=fnd&pg=PR9&dq=Hubert+Dreyfus+says+that+AI+is+impossible&ots=FjLGFaXOF1&sig=rfz7LbOATpzK3FypkLB_nWg&redir_esc=y#v=onepage&q=Hubert%20Dreyfus%20says%20that%20AI%20is%20impossible&f=false	ναι
		https://books.google.gr/books?id=v6IOPDO473EC&pg=PA95&pg=PA95&dq=John+Searle+says+that+the+idea+of+a+non-biological+machine+being+intelligent+is+incoherent&source=bl&ots=5fA7AKH8LV&sig=14mwmAe1H835xsfQY6_YfM9O7Z0&hl=el&sa=X&ved=2ahUKEwj9nKGzV9HfAhVC94MKHaK0BwMQ6AEwAEOECC8QAQ#v=onepage&q=John%20Searle%20says%20that%20the%20idea%20of%20a%20non-biological%20machine%20being%20intelligent%20is%20incoherent&f=false	Ναι
Public sector effectiveness measurement	sciencedirect	https://www.sciencedirect.com/science/article/pii/S1877042815058991	ναι
trust	sciencedirect	https://www.sciencedirect.com/science/article/pii/S0176268017306122	Ναι
what is neural network?	google scholar	file:///C:/Users/user/Downloads/MATLAB_-_Neural_Network_Toolbox_5_-_Users_Guide_2007.pdf	Ναι
Urgency in AI	elsevier	https://www.elsevier.com/connect/an-air-of-urgency-why-we-need-ethical-governance-of-ai	Ναι
AI Impact in mental	sciencedirect	https://www.sciencedirect.com/science/article/pii/S0924933818301627	ναι

health			
Φτώχεια στη Ελλάδα	ΕΛΣΤΑΤ	http://www.statistics.gr/documents/20181/7146186/LivingConditionsInGreece_1118.pdf/a7883ee9-6ce0-4c35-b70c-b5b8618df3b3	Ναι
technological unemployment	sciencedirect	https://www.sciencedirect.com/science/article/pii/S0016328716302063	Ναι
Communication between government and citizens	sciencedirect	https://www.sciencedirect.com/science/article/pii/S0740624X17304008	Ναι
Living with robots	sciencedirect	https://www.sciencedirect.com/science/article/pii/S0747563218305806	Ναι
Threats and Solutions in AI	ssnr	https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3183238	Ναι
Ethical agency	Διαδίκτυο	https://pdfs.semanticscholar.org/20a1/6a9eead8d868f0f24269bade24a359161c34.pdf	Ναι
Ethical conversational agents	google scholar	https://www.researchgate.net/profile/Antonella_De_Angeli/publication/220168786_Ethical_implications_of_verbal_disinhibition_with_CAs/links/54a6a5f80cf267bdb909db3e.pdf#page=49	Ναι
Living with robots	sciencedirect	https://www.sciencedirect.com/science/article/pii/S2095809917307555	Ναι
Human Progress	sciencedirect	https://www.sciencedirect.com/science/article/pii/B9780124201903000338	Ναι
Evil inside us	sciencedirect	https://www.sciencedirect.com/science/article/pii/S016726811400314X	Ναι
	google scholar	https://books.google.gr/books?hl=el&lr=&id=PUzYWfbf-qkC&oi=fnd&pg=PT7&dq=why+people+do+bad+things&ots=rgmKj65PUe&sig=frRi1RNg9RzSU-7kcWYICwZAg8&redir_esc=y#v=onepage&q=why%20people%20do%20bad%20things&f=false	Ναι
E-gov is a tyranny?	google scholar	https://link.springer.com/chapter/10.1007/978-3-319-64248-2_2	Ναι
	google scholar	https://books.google.gr/books?hl=el&lr=&id=xhd-9HzSqdMC&oi=fnd&pg=PA17&dq=egovernment+is+a+new+dictatorships&ots=XJGXFEHTng&sig=Lrga7cc81wkZkt2mF4ccebSzl1M&redir_esc=y#v=onepage&q=egovernment%20is%20a%20new%20dictatorships&f=false	Ναι
Risk in e-gov	sciencedirect	https://www.sciencedirect.com/science/article/abs/pii/S0963868707000637	Ναι
		https://www.sciencedirect.com/science/article/pii/S0740624X1730196X	Ναι
E-gov uses AI	sciencedirect	https://www.sciencedirect.com/science/article/pii/S0740624X17304008	Ναι
AI and propaganda	Διαδίκτυο	http://blogs.oii.ox.ac.uk/politicalbots/wp-content/uploads/sites/89/2017/06/Comprop-Canada.pdf	Ναι
AI and fake news	sciencedirect	file:///C:/Users/user/Desktop/final/fake%20news.pdf	Ναι
		https://reader.elsevier.com/reader/sd/pii/S1877050918318210?token=AC153413F22286D3738F4DAF4F6F7A2109FCEDE297308CB683C2359C48F77DB9A75B132A1E1923A9441705E565B7ABB	Ναι
AI for good	ssnr	https://poseidon01.ssrn.com/delivery.php?ID=903006029020064077117006127077104029024042049020026087093008118026067012125089117112098012030034041017014080084097122126127019040081009023072031115028099104091099029084016037117094120068088019123121102102125086020104006065074099097065004066110086123082&EXT=pdf	Ναι
	Διαδίκτυο	http://unpan1.un.org/intradoc/groups/public/documents/CAIMED/UNPAN028992.pdf	Ναι
Blockchain	sciencedirect	https://www.sciencedirect.com/science/article/abs/pii/S0308596117302483	Ναι
	Διαδίκτυο	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5814166/	Ναι
Λήμμα.	Που αναζητήθηκε;	Ευρήματα.	Αναφορά στην διατριβή.
Evil robots	sciencedirect	https://www.sciencedirect.com/science/article/pii/S1071581916301768	Ναι

	sciencedirect	https://ac.els-cdn.com/S0010027712001278/1-s2.0-S0010027712001278-main.pdf?_tid=01fc5a32-ee75-4b81-9176-4097b5c139fe&acdnat=1525337453_335c4e67440169e5ecc4840ae151bac	Ναι
	sciencedirect	https://reader.elsevier.com/reader/sd/E1C21217D4AD4CFC80B0155FA5A5AF5061985A3FB63BE2BC45610E338F9E08DFB23781604A71211A8FE883FF2EFBF830	Ναι
	sciencedirect	https://www.sciencedirect.com/science/article/pii/S1474667016434324	Ναι
Scaling rights	sciencedirect	https://www.sciencedirect.com/science/article/pii/S0049089X16000065	Ναι
Χρηματοδότηση τεχνητής νοημοσύνης	Διαδίκτυο	file:///C:/Users/user/Downloads/IP-18-3362_EL.pdf	Ναι
EGE	Διαδίκτυο	https://ec.europa.eu/research/ege/index.cfm	Ναι
EGE statement	Διαδίκτυο	https://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf	Ναι
Oath on AI	Διαδίκτυο	https://techcrunch.com/2018/03/14/a-hippocratic-oath-for-artificial-intelligence-practitioners/	Ναι
Ethical frame principles	διαδίκτυο	http://standards.ieee.org/develop/indconn/ec/ead_v2.pdf	ναι
Ethical frame principles	διαδίκτυο	https://www.turing.ac.uk/research/interest-groups/data-ethics-group	ναι
Ιεραρχία αναγκών	διαδίκτυο	https://www.simplypsychology.org/maslow.html	ναι
AAAI	διαδίκτυο	https://aaai.org/Papers/AAAI/2006/AAAI06-292.pdf	ναι
		file:///C:/Users/user/Downloads/8308-38149-1-PB.pdf	ναι
declaratin for Ethis in AI	διαδίκτυο	https://www.montrealdeclaration-responsibleai.com/	Ναι
Future of life principles	διαδίκτυο	https://futureoflife.org/ai-principles/	ναι
UN weapon ethics	Διαδίκτυο	http://www.unidir.org/files/publications/pdfs/considering-how-meaningful-human-control-might-move-the-discussion-forward-en-615.pdf	Ναι
AI oath	Διαδίκτυο	https://techcrunch.com/2018/03/14/a-hippocratic-oath-for-artificial-intelligence-practitioners/	Ναι
Ethics and law, principles	Διαδίκτυο	https://aticdn.s3-eu-west-1.amazonaws.com/2018/03/140318-Ai-ethics-and-the-law-public-panel-report.pdf	ναι
Ο νόμος στη Σπάρτη	Διαδίκτυο	https://youtu.be/hrFzL60HpRw	ναι
	Διαδίκτυο	http://ebooks.edu.gr/modules/ebook/show.php/DSGYM-A113/349/2342,8930/	Ναι
Moral certainty	Sciencedirect	https://www.sciencedirect.com/science/article/pii/0277953686900043	ναι
Λήμμα.	Που αναζητήθηκε;	Ευρήματα.	Αναφορά στην διατριβή.
living with robots	Sciencedirect	https://www.sciencedirect.com/science/article/pii/S0747563218305806	Ναι
Λήμμα .	Που αναζητήθηκε;	Ευρήματα.	Αναφορά στην διατριβή.
AI in education	sciencedirect	https://reader.elsevier.com/reader/sd/pii/S1877050918315382?token=932B7C8B698BC8F3F9E0EE99B556B7A70136926685BC1985DF48C0281234E0F0897176BFE112F8DA319D5BDB67B9E767	Ναι

Βιβλιογραφία

A (Very) Brief History of Artificial Intelligence Bruce G. Buchanan (Alan Turing, Computing machinery and intelligence.Mind, 1950).

Οδύσσεια (θ555).

Ιλιάδα (Σ376).

Anxious: Using the Brain to Understand and Treat Fear and Anxiety Paperback. Joseph LeDoux, Center for the Neuroscience of Fear and Anxiety, New York University, – August 23, 2016.

Richard T. Watson και Jane Webster (Webster & Watson/Guest Editorial, MIS Quarterly Vol. 26 No. 2/June 2002.

Ethics and the Limits of Philosophy.B Williams – 2011.

Ηθικά Νικομάχεια, Αριστοτέλης, βιβλίο Β΄.

Towards an ethics of life. Christopher Ketcham. Elsevier Ltd. 2016.

The Mill on the Floss (1860), p. 628. Eliot G. Penguin Books, New York, 1979.

What is data ethics? Luciano Floridi and Mariarosaria Taddeo. 2016.

Artificial intelligence: past and future. M.Y. Vardi. s.l.: association for computer machinery, 2012.

Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. Mohammad Hossein Jarrahi. 2018.

Stifling artificial intelligence: Human perils. Gonenc Gurkaynak, Ilay Yilmaz, Gunes Haksever. Computer law & security review 32. 2016.

The singularity is near: When humans transcend biology. Kurzweil, R. 2005.

A Safe Ethical System for Intelligent Machines, Mark R. Waser. 2009.

Our Final Invention: Artificial Intelligence and the End of the Human Era. James Barat. 2013.

Primates and Philosophers.How Morality Evolved. Frans de Waal, 2006.

Mind over machine. Hubert L. Dreyfus, Stuart L. Dreyfus, Free Press New York. 2000.

The rebirth of Mankind Homo Evolutis, Trend Goodbaudy, 2012.

The Concept of Intelligence ULRIC NEISSER.Cornell University 1979.

THE LEARNING MACHINES. Using massive amounts of data to recognize photos and speech,deep-learning computers are taking a big step towards true artificial intelligence. NICOLA JONES. 2004.

Product and technology innovation: What can biomimicry inspire? Research review paper.Elena Lurie-Luke. 2014.

Homo Deus: A brief history of tomorrow. Harari, Y. N. 2016.

Axioms and Tests for the Presence of Minimal Consciousness in Agents. Igor Aleksander and Barry Dunmall, 2003.

A Safe Ethical System for Intelligent Machines. Mark R. Waser. 2009.

The Aspects of Performance Measurement in Public Sector Organization, Ingrida Balaboniene, Giedre Veserskiene. 2015.

Centralization or decentralization? A review on the effects of information and communication technology on urban spatial structure. Hashem Dadashpoor, Zahed Yousefi.2018.

Ethics Management in Public Sector – Background and Tools. Silvia Puiu. 2nd GLOBAL CONFERENCE on BUSINESS, ECONOMICS, MANAGEMENT and TOURISM, 2014.

Artificial Intelligence for Citizen Services and Government. Hila Mehr.Harvard Ash Center Technology & Democracy Fellow. August 2017.

What can history teach us about technology and jobs? McKinsey Global Institute. 2018.

The rise of technological unemployment and its implications on the future macroeconomic landscape. Young Joon, Kima Kyungsoo, KimbSu KyoungLee Futures Volume 87, March 2017.

Non-Linear System Identification using Neural Networks Oriented Speech.. LiuHuaqiang, Dai Guanzhong, Xu Naiping. 1993.

Ethical implications of verbal disinhibition with conversational agents. Antonella De Angeli. 2009.

Emotion recognition from Marathi speech database using adaptive artificial neural network. Raviraj Vishwambhar, Darekara Ashwinikumar, Panjabrao Dhande, 2017.

An android architecture for bio-inspired honest signalling in Human-Humanoid Interaction..RosarioSorbelloa, SalvatoreTramontea, CarmeloCalíb, MarcelloGiardinaa, ShuichiNishioc, HiroshiIshigurocd, AntonioChella, 2018.

Prominence features: Effective emotional features for speech emotion recognition. Shaoling Jing, Xia Mao, Lijiang Chen, 2017.

The digital identity issue. Paul Simmonds. Network Security, 2015.

Digital identity and French personality rights – A way forward in recognising and protecting an individual's rights in his/her digital identity. Clare Sullivana, Sophie Stalla-Bourdillon. Computer Law & Security Review, 2015.

Are robots becoming unpopular? Changes in attitudes towards autonomous robotic systems in Europe. Timo Gnambsab, Markus Appel. 2018.

Learning to trust artificial intelligence systems. Accountability, compliance and ethics in the age of smart machines. Dr. Guruduth Banavar. Chief Science Officer, Cognitive Computing Vice President, IBM Research. 2016.

Trust no more? On the lasting effects of corruption scandals. Albert Solé Ollé, Pilar Sorribas-Navarro, European Journal of Political Economy, 2018.

Transforming the communication between citizens and government through AI-guided chatbots. Aggeliki Androutsopoulou, Nikos Karacapilidis, Euripidis Loukis, Yannis Charalabidisa. 2018.

Determinants of User Adoption of e-Government Services in Greece and the role of Citizen Service Centres. Anastasia Voutinioti. 2013.

“An air of urgency” – why we need ethical governance of AI, ORBIT, 2019.

Ethical governance is essential to building trust in robotics and artificial intelligence systems. Alan F. T. Winfield, Marina Jirotko. 15 October 2018.

Ethical guidelines for a superintelligence. Ernest Davis. 2015.

PoTrojan: powerful neural-level trojan designs in deep learning models. Minhui Zou, Student Member, IEEE, Yang Shi, Student Member, IEEE, Chengliang Wang, Member, IEEE, Fangyu Li, WenZhan Song, Senior Member, IEEE and Yu Wang, Senior Member, IEEE. 2018.

Δημοσίευμα στον τύπο “Facebook shuts down robots after they invent their own language” πηγή <https://www.telegraph.co.uk/technology/2017/08/01/facebook-shuts-robots-invent-language/>

Liability for damages caused by artificial intelligence. Paulius Čerkaa, Jurgita Grigienė, Gintarė Širbikytė. 2015.

The ethical primate. Mary Midgley, 2002.

AI, ethics and the law. What challenges and what opportunities? The Alan Turing Institute. 2018.

Οι δωριστί αρμονίες, Το αληθινό πρόσωπο της αρχαίας Σπάρτης. Δ. Νικολακάκης, 1998.

Does counterterrorist legislation hurt human rights practices? A longitudinal cross-national analysis. Shor Eran, Ina Filkobskia, Pazit Ben-Nun Bloomb, Hayder Alkilabia, William Su. 2016.

Artificial Intelligence, Machine Learning, Deep Learning, and Cognitive Computing: What Do These Terms Mean and How Will They Impact Health Care? Stefano A.BiniMD. 2018.

Medical Data Processing System Based on Neural Network and Genetic Algorithm. Azamat Amirov, Olga Gerget, Dmitry Devjatyh b, Arstan Gazaliev.2013.

Societal Impacts and Ethics of GIS. Jeremy W. Crampton, Eric M. Huntley, Emily C. Kaufman. 2017.

AIC – An AI-system for Combination of senses. Anne Håkansson. 2013.

Weather Monitoring Using Artificial Intelligence, IEEE, Computational Intelligence and Networks (CINE), 2016 2nd International Conference

Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. Andreas Kaplan, Michael Haenlein. 2019.

A comparison of selected Western and Chinese smart governance: The application of ICT in governmental management, participation and collaboration. Yanliu Lin, 2018.

Computational Propaganda in Canada: The Use of Political Bots. Fenwick McKelvey, Elizabeth Dubois. 2017.

The current state of fake news: challenges and opportunities. Álvaro Figueira, Luciana Oliveira.2017.

Detecting Fake News in Social Media Networks. Monther Aldwairi, Ali Alwahedi. 2018.

Blockchain's roles in strengthening cybersecurity and protecting privacy. NirKshetri. 2017.

Converging blockchain and next-generation artificial intelligence technologies to decentralize and accelerate biomedical research and healthcare. Polina Mamoshina, Lucy Ojomoko, Yury Yanovich, Alex Ostrovski, Alex Botezatu, Pavel Prikhodko, Eugene Izumchenko, Alexander Aliper, Konstantin Romantsov, Alexander Zhebrak, Iraneus Obioma Ogu, Alex Zhavoronkov. 2017.

From e-government to m-government: towards a new paradigm in public administration? Ignace Snellen, Marcel Thaens. 2008.

Killer decisions: The role of cognition, affect and ‘expertise’ in homicide. Fiona Brookman. 2015.

Ethical Robots in Warfare. RONALD C. ARKIN. 2012.

Analytical Stochastic Model of Traffic.Serge L.Shishk.

Ethical implications of verbal disinhibition with conversational agents. Antonella De Angeli. 2009.

Modelling social action for AI agents. Cristiano Castelfranchi. 1998.

PRINCIPLES AND THE SEARCH FOR MORAL CERTAINTY LARRY R. CHURCHILL and Jo& JORGE SIMON, Social and Administrative Medicine, University of North Carolina, U.S.A.1986.

(MedEthEx:A Prototype Medical Ethics Advisor, Michael Anderson,, Susan Leigh Anderson,, Chris Armen. Published in AAAI 2006.

GenEth: A General Ethical Dilemma Analyzer, Michael Anderson, Susan Leigh Anderson. Published in Paladyn 2014.

Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions.Thomas King. 2018.

Ethical Decision Making in Robots: Autonomy,Trust and Responsibility. Fahad Alaiერი and Andr´e Vellino. Published in ICSR 2016.

Robot ethics. Mapping the issues for a mechanized world. Patrick Lin, Keith Abney, George Bekey. 2010.

On Human Nature. Biology, Psychology, Ethics, Politics, and Religion. Chapter 33 - Human Evolution and Progress. J.Ayala. 2017.

The economic origins of the evil eye belief. Boris Gershman. 2015.

Why Good People Do Bad Things: Understanding Our Darker Selves. James Hollis. 2007.

A CALL TO ACTION: MOVING FORWARD WITH THE GOVERNANCE OF ARTIFICIAL INTELLIGENCE IN CANADA. Aviv Gaon και Ian Stedman. 2019.

The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms Professor Spyros Makridakis, 2017.

Ethical implications of verbal disinhibition with conversational agents. Antonella De Angeli.2009.

Development and Future Challenges of Bio-Syncretic Robots. Chuang Zhanga. Wenxue Wanga, Ning Xic, Yuechao Wanga, Lianqing Liu.2018.

Why is AI so scary? Michael L. Anderson. 2005.

Beware an AI-fuelled world. Andrew Simms. 2018.

The rise of technological unemployment and its implications on the future macroeconomic landscape. Young Joon, Kima Kyungsoo, Kimb Su, Kyoung Lee. 2017.

THE FUTURE OF EMPLOYMENT: HOW SUSCEPTIBLE ARE JOBS TO COMPUTERISATION? Carl Benedikt Frey, Michael A. Osborne. 2013.

Transforming the communication between citizens and government through AI-guided chatbots. Aggeliki Androutsopoulou, Nikos Karacapilidis, Euripidis Loukisa, Yannis Charalabidis. 2018.

Tyrant Leaders as e-Government Service Promoters: The Role of Transparency and Tyranny in the Implementation of e-Government Service. Yuting Lin, Andreas Eisingerich, Hersen Doong. 2017.

Governance and Information Technology: From Electronic Government to Information Government. Viktor Mayer-Schönberger, David Lazer. 2007.

Trust and risk in e-government adoption. France Bélanger, Lemuria Carter. 2008.

The public value of E-Government – A literature review. Jean Damascene Twizeyimana, Annika Andersson. 2019.

Are robots becoming unpopular? Changes in attitudes towards autonomous robotic systems in Europe. Timo Gnamsa, Markus Appel. 2018.

Artificial Intelligence in the Age of Neural Networks and Brain Computing 2019, Pages 161-190. Chapter 8 - The New AI: Basic Concepts, and Urgent Risks and Opportunities in the Internet of Things. Paul J. Werbos. 2019.

ΚΑΝΟΝΙΣΜΟΣ (ΕΕ) 2016/679 ΤΟΥ ΕΥΡΩΠΑΪΚΟΥ ΚΟΙΝΟΒΟΥΛΙΟΥ ΚΑΙ ΤΟΥ ΣΥΜΒΟΥΛΙΟΥ της 27ης Απριλίου 2016 για την προστασία των φυσικών προσώπων έναντι της επεξεργασίας των δεδομένων προσωπικού χαρακτήρα και για την ελεύθερη κυκλοφορία των δεδομένων αυτών και την κατάργηση της οδηγίας 95/46/ΕΚ.

AI and Big Data: A blueprint for a human rights, social and ethical impact assessment. Alessandro Mantelero. 2018.

Promoting open innovation in the public sector through social media monitoring. Euripidis Loukis, Yannis Charalabidis, Aggeliki Androutsopoulou. 2017.

Kurzweil's argument for the success of AI The Singularity Is Near: When Humans Transcend Biology, Ray Kurzweil, Viking Adult. Drew McDermott. 2006.

Artificial Intelligence in the Age of Neural Networks and Brain Computing 2019.

Big Ideas for Little Kids: Teaching Philosophy through Children's Literature. Thomas E. Wartenberg. 2014.

Accuracy of early GFS and ECMWF Sandy (2012) track forecasts: Evidence for a dependence on cumulus parameterization. Nick P. Bassill. 2014.

Firebird: Predicting Fire Risk and Prioritizing Fire Inspections in Atlanta. Michael Madaio, Shang-Tse Chen, Oliver L. Haimson, Wenwen Zhang, Xiang Cheng, Matthew Hinds-Aldrich, Duen Horng Chau, Bistra Dilkina. 2016.

Σχήματα

Σχήμα 1. Οι κύκλοι της ηθικής. Η σχέση των συνόλων ηθικής και νομιμότητας.

Σχήμα 2. AI History 2018.

Σχήμα 3. Οι ηθικές αρχές της X2AI.

Σχήμα 4. Η διαδικασία αποφάσεων στην εφαρμογή τεχνητής νοημοσύνης MedEthEx.

Σχήμα 5. AI's impact.

Σχήμα 6. Ιεράρχηση των αναγκών του ανθρώπου.

Σχήμα 7. Αντιδράσεις ανθρώπων σε αυτόνομα συστήματα.

Σχήμα 8. Αντιδράσεις ανθρώπων σε αυτόνομα συστήματα.

Σχήμα 9. Αντιδράσεις ανθρώπων σε αυτόνομα συστήματα.

Σχήμα 10. Αντιδράσεις ανθρώπων σε αυτόνομα συστήματα.

Σχήμα 11. Τα πεδία της SWOT analysis.

Σχήμα 12. Οι δυνατότητες που μπορεί να δώσει τεχνητή νοημοσύνη στον δημόσιο τομέα.