



ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ
ΣΧΟΛΗ ΚΟΙΝΩΝΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΚΟΙΝΩΝΙΟΛΟΓΙΑΣ

**ΜΕΡΟΛΗΨΙΕΣ ΚΑΙ ΜΑΥΡΑ ΚΟΥΤΙΑ ΤΗΣ ΤΕΧΝΗΤΗΣ
ΝΟΗΜΟΣΥΝΗΣ ΣΕ ΣΥΝΘΗΚΕΣ ΚΟΙΝΩΝΙΑΣ ΤΗΣ
ΔΙΑΚΙΝΝΔΥΝΕΥΣΗΣ:**

**Μια κοινωνιολογική προσέγγιση των αλγορίθμων στη
δημόσια διακυβέρνηση.**

ΤΟΥ : ΧΑΡΑΛΑΜΠΟΠΟΥΛΟΥ ΙΩΑΝΝΗ ΝΙΚΟΛΑΟΥ

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ: ΚΟΥΡΟΥΤΖΑΣ ΧΡΗΣΤΟΣ

Ευχαριστίες

Θα ήθελα να ευχαριστήσω θερμά τον διδάσκοντα και επιβλέποντα καθηγητή Χρήστο Κουρούτζα για την επιλογή, την εμπιστοσύνη και την συνεισφορά του καθ' όλη την διάρκεια της εργασίας. Επίσης το συγκεκριμένο εγχείρημα δεν θα ήταν εφικτό δίχως την συνδρομή του Κρικόρ, της Σοφίας, της Κάτιας και της οικογένειάς μου.

Περίληψη

Η παρούσα εργασία συνιστά μια κοινωνιολογική προσέγγιση των μεροληψιών και των μαύρων κουτιών των τεχνολογιών της τεχνητής νοημοσύνης στην κοινωνία, εμβαθύνοντας στην περίπτωση του μεροληπτικού μαύρου κουτιού του αλγορίθμου των ολλανδικών παροχών παιδικής μέριμνας. Συλλέγοντας και εξετάζοντας πηγές από τον χώρο των σπουδών της Επιστήμης και Τεχνολογίας, της Κοινωνιολογίας της Τεχνητής Νοημοσύνης, των Κοινωνιών του Μαύρου Κουτιού και των Κοινωνιών της Διακινδύνευσης, εξετάζει την δυναμική σχέση αλληλεπίδρασης της τεχνητής νοημοσύνης και των σύγχρονων (μετα)νεωτερικών κοινωνιών. Από την μελέτη των θέσεων των STS αναδεικνύεται η πορεία των δύο ξεχωριστών μέχρι τότε πεδίων της επιστήμης και της τεχνολογίας προς την κατανόησή τους ως ένα ενιαίο, αυτού της τεχνοεπιστήμης χάρη στην επιρροή του κοινωνικού κονστρουκτιβισμού. Έτσι οι τεχνολογίες τεχνητής νοημοσύνης συγκροτούνται κοινωνικά με αποτέλεσμα να μην είναι ουδέτερες λαμβάνοντας συχνά μεροληπτικές αποφάσεις κατά των αδύναμων και Υποεκπροσωπούμενων κοινωνικών ομάδων εξαιτίας της μετατροπής τους σε μαύρα κουτιά. Καθοριστική σημασία για την διάνοιξή τους έχουν οι κριτικές σπουδές αλγορίθμων οι οποίες αντιλαμβάνονται τις τεχνολογίες τεχνητής νοημοσύνης ως κοινωνικοτεχνικά συστήματα καθώς αυτά όχι μόνο επηρεάζονται αλλά και επηρεάζουν τις κοινωνίες στις οποίες χρησιμοποιούνται. Από την ανάλυση συμπεραίνεται ότι η αλόγιστη αξιοποίηση τέτοιων συστημάτων, όπως στην περίπτωση των ολλανδικών παροχών μέριμνας, αποτελεί μια σύγχρονη μορφή διακινδύνευσης που απειλεί τις σύγχρονες κοινωνίες με την αναπαραγωγή μεροληψιών και την άνιση κατανομή της εξουσίας.

Λέξεις κλειδιά

Κοινωνίες της Τεχνητής Νοημοσύνης, Κοινωνιολογία της Τεχνητής Νοημοσύνης, Μεροληψίες της Τεχνητής Νοημοσύνης, Μαύρα Κουτιά, Τεχνητή Νοημοσύνη στην Δημόσια Διακυβέρνηση, Toeslagenaffaire.

Abstract

This paper constitutes a sociological approach to the biases and black boxes of AI technologies in society, delving into the case of the biased black box of the Dutch childcare benefits algorithm. By collecting and examining sources from the fields of Science and Technology Studies, Sociology of AI, Black Box Societies and Risk Societies, it examines the dynamic interaction between AI and contemporary (post)modern societies. The study of the STS positions highlights the course of the two hitherto separate fields of science and technology towards their understanding as a unified one, that of technoscience thanks to the influence of social constructivism. Thus, AI technologies are socially constructed so that they are not neutral, often taking biased decisions against weak and underrepresented social groups because of their turn into black boxes. Critical studies of algorithms that understand AI technologies as socio-technical systems are crucial for their opening up, as they not only affect but also influence the societies in which they are used. The analysis concludes that the indiscriminate use of such systems, as in the case of Dutch care benefits, is a modern form of risk that threatens (meta)modern societies by reproducing biases and unequal distribution of power.

Keywords

AI Societies, Sociology of AI, AI Biases, Black Boxes, AI in Public Governance, Toeslagenaffaire

Περιεχόμενα

Περίληψη.....	3
Λέξεις κλειδιά	3
Abstract	4
Keywords.....	4
Εισαγωγή.....	7
A. Θεωρίες Επιστήμης και Τεχνολογίας: Η Συμβολή των Σπουδών Επιστήμης και Τεχνολογίας	11
A.1 Το φιλοσοφικό ρεύμα του Θετικισμού και οι επιρροές του σε Επιστήμη και Τεχνολογία	11
A.2 Θετικισμός στην Επιστήμη και στην Τεχνολογία: Η Δυσϊστική Προοπτική	12
A.3 Λογικός θετικισμός και Κύκλος της Βιέννης.....	13
A.4 Καρλ Πόπερ: Η διάκριση ανάμεσα στην αληθινή και ψευδή επιστήμη	14
A.5. Κοινωνικός Κονστρουκτιβισμός.....	16
<i>A.5.1. Κριτική στον Κονστρουκτιβισμό: η ηγεμονία της κατασκευής.....</i>	<i>18</i>
<i>A.5.2. Κοινωνικός κονστρουκτιβισμός και σπουδές Επιστήμης και Τεχνολογίας : Κατασκευή κοινωνικών και υλικών περιβαλλόντων</i>	<i>20</i>
<i>A.5.3. Η γενικότερη συμβολή της θεωρίας της Κοινωνικής κατασκευής για τις θεωρίες της τεχνοεπιστήμης.....</i>	<i>21</i>
A.6. «Μέρτον και η επιστημονική ηθική».....	22
A.7. Η Κουινιανή Επανάσταση στον χώρο της επιστήμης.....	26
<i>A.7.1. Ασυμμετρία και κριτική του Κουινιανού έργου.....</i>	<i>28</i>
A.8. Ισχυρό Πρόγραμμα της Κοινωνιολογίας της Γνώσης και Κοινωνική Κατασκευή της Τεχνολογίας (SCOT)	29
<i>A.8.1. Ερμηνεία: Η κινητήριος δύναμη της Γνώσης.....</i>	<i>31</i>
A.9. Ισχυρό Πρόγραμμα Κοινωνιολογίας της γνώσης και Επιστήμης της τεχνολογίας : Η συμβολή του ισχυρού προγράμματος στην κονστρουκτιβιστική ανάγνωση της τεχνολογίας.....	33
A.10 Μπρούνο Λατούρ και Στιβ Γόυλγκαρ: Η θεωρία δρώντων δικτύου και οι μελέτες εργαστηρίου.	35
<i>A.10.1 Οι φιλοσοφικές καταβολές της θεωρίας δρώντων δικτύου: Βιταλισμός και Νεοβιταλισμός.....</i>	<i>35</i>
<i>A.10.2. Η ιστορία συγκρότησης της Θεωρίας δρώντων Δικτύου</i>	<i>39</i>
<i>A.10.3 Η απομάκρυνση της θεωρίας δρώντων δικτύου από την κοινωνική κατασκευασιοκρατία</i>	<i>41</i>
B. Κοινωνίες της Τεχνητής Νοημοσύνης	44
<i>B.1 Οι καταβολές της τεχνητής νοημοσύνης και η πορεία της έως σήμερα</i>	<i>46</i>
B.2. Κοινωνιολογία της Τεχνητής Νοημοσύνης	49
<i>B.2.1. Η (Α)πορεία για την Κοινωνιολογία της Τεχνητής Νοημοσύνης.....</i>	<i>49</i>

<i>B.2.2. Η επανάσταση των μεγάλων δεδομένων και τα νέα τεχνοεπιστημονικά παραδείγματα</i>	51
<i>B.2.3. Σύγχρονες μορφές της Κοινωνιολογίας της τεχνητής νοημοσύνης</i>	53
<i>B.2.4. Κριτικές Σπουδές Αλγορίθμων-Critical Algorithm Studies (CAS)</i>	58
B.3. Τεχνητή Νοημοσύνη και Μεροληψίες	60
<i>B.3.1. Η φύση των μεροληψιών</i>	60
<i>B.3.2 Τύποι μεροληπτικών διακρίσεων μέσω της μηχανικής μάθησης</i>	61
<i>B.3.3 Κριτική Προσέγγιση των μεροληψιών της Τεχνητής Νοημοσύνης</i>	63
B.4.Οι διακινδυνεύσεις της Τεχνητής Νοημοσύνης στις Κοινωνίες του Μαύρου Κουτιού	68
<i>B.4.1. Οι καταβολές της έννοιας του μαύρου κουτιού και η αξιοποίηση της από το πεδίο των STS</i>	69
<i>B.4.2. Η συνύπαρξη των Κοινωνιών της Διακινδύνευσης και των Κοινωνιών του Μαύρου Κουτιού: Μια ανάλυση στα πλαίσια των συστημάτων τεχνητής νοημοσύνης</i>	71
<i>B.4.3. Η διάνοιξη των μαύρων κουτιών μέσω των θεωριών ANT και SCOT</i>	75
B.5. Η μεροληπτική διακυβέρνηση μέσω μεγάλων δεδομένων και αλγορίθμων	83
Γ.«Toeslagenaffaire»	85
Γ.1.Το κοινωνικοπολιτικό υπόβαθρο του αναστοχαστικού μοντέλου ταξινόμησης ρίσκου: Υψηλό ρίσκο-υψηλή αμοιβή	86
<i>Γ.1.1 Από την ανίχνευση κινδύνων στην παραγωγή τους</i>	90
Γ.2. Η μεροληπτική σχεδίαση του συστήματος ταξινόμησης ρίσκου	91
Γ.3.Το μαύρο κουτί του συστήματος ταξινόμησης ρίσκου	94
Γ.4. Τα αποτελέσματα του μαύρου κουτιού του αλγοριθμικού συστήματος παιδικής μέριμνας ως συμπεράσματα	97
Επίλογος-Συμπεράσματα	100
Προτάσεις	104
Βιβλιογραφία	107

Εισαγωγή

Οι βιομηχανικές επαναστάσεις πάντα αποτελούσαν περιόδους στις οποίες οι κοινωνίες βίωναν σημαντικές αλλαγές όσον αφορά τον τρόπο συγκρότησης και αναπαραγωγής τους. Κάτι τέτοιο συμβαίνει καθώς οι τεχνολογικές ανακαλύψεις που συνέβαιναν εντός αυτών των περιόδων επέτρεψαν την εξέλιξη των μέσων παραγωγής και την αναδιαμόρφωση των κοινωνικών δομών. Στις περιπτώσεις αυτές, η τεχνολογία δεν ήταν η μόνη η οποία άλλαζε την κοινωνία. Μεγάλο μέρος των τεχνολογικών και επιστημονικών δημιουργημάτων πραγματοποιούνται βασιζόμενα σε ανάγκες και βλέψεις της κοινωνίας σχετικά με την εικόνα που η ίδια επιθυμεί για τον εαυτό της. Πρόκειται κατά κύριο λόγο για μια αμφίδρομη, δυναμική σχέση όπου η τεχνολογία, η επιστήμη που την διέπει αλλά και η ίδια η κοινωνία που τις περιβάλλει, αλληλοεπιδρούν, διαμορφώνοντας η μια την άλλη.

Οι υστερονεωτερικές κοινωνίες της διακινδύνευσης (Beck, 2015) χαρακτηρίζονται από την έντονη εμφάνιση τεχνολογιών τεχνητής νοημοσύνης οι οποίες έχουν επηρεάσει σε σημαντικό βαθμό πληθώρα κοινωνικών, οικονομικών, πολιτικών διαδικασιών συγκροτώντας μια πρωτόγνωρη, πολλές φορές υβριδική εικόνα. Οι τεχνολογίες τεχνητής νοημοσύνης, η επανάσταση των μεγάλων δεδομένων, το Ίντερνετ των Πραγμάτων, αποτελούν σημαντικό τεχνοεπιστημονικό τμήμα της τέταρτης βιομηχανικής επανάστασης η οποία λαμβάνει παγκόσμιες διαστάσεις (Ross & Maynard, 2021; Schwab, 2023). Η διευρυμένη χρήση των τεχνολογιών τεχνητής νοημοσύνης μέσω της χρήσης μεγάλων δεδομένων έχει συμβάλλει καθοριστικά στην αλλαγή των τρόπων με τους οποίους πληθώρα ενεργειών πραγματοποιούνται στις μεταμοντέρνες κοινωνίες. Εξαιτίας λοιπόν όλων αυτών των κρίσιμων αλλαγών σε όλο τον κόσμο, πολλοί διανοητές, επιστήμονες και τεχνολόγοι θεωρούν τις σύγχρονες κοινωνίες, «Κοινωνίες της Τεχνητής Νοημοσύνης».

Η παρούσα εργασία αποτελεί μια κοινωνιολογική πραγμάτευση των Κοινωνιών της Τεχνητής Νοημοσύνης, η οποία εστιάζει κυρίως στις κοινωνικές προκαταλήψεις που αναπαράγονται από αυτά τα συστήματα και σε πολλές περιπτώσεις στα συστήματα λήψης αποφάσεων που αξιοποιούνται όχι μόνο στον ιδιωτικό αλλά και στον δημόσιο τομέα και τη διακυβέρνηση. (Liu, 2021). Όπως όλα τα κοινωνικά φαινόμενα, έτσι και οι μεροληψίες εντοπίζονται σε όλες τις τεχνολογικές κατασκευές. Η τεχνολογική αναβάθμιση που χαρακτηρίζει όλες τις τεχνολογίες εξαιτίας της ανάπτυξης της τεχνητής νοημοσύνης, δεν σταματάει την αναπαραγωγή των διακρίσεων και των

στερεοτύπων , βάσει φύλων, θρησκείας, φυλής, εθνικότητας και κοινωνικοοικονομικής θέσης. Αντιθέτως, η αξιοποίηση αυτοματοποιημένων συστημάτων τεχνητής νοημοσύνης μέχρι στιγμής έχει αναδείξει πληθώρα διακρίσεων και μεροληψιών κατά πολλών κοινωνικών ομάδων και ιδίως κατά των αδύναμων και κοινωνικά υποεκπροσωπούμενων (O'neil, 2016; Eubanks, 2018; Noble, 2018; Benjamin, 2019; Gebru, 2020; Crawford, 2021)

Οι κοινωνικές μεροληψίες μεταλαμπαδεύονται από τα κοινωνικά περιβάλλοντα εντός των συστημάτων τεχνητής νοημοσύνης με την μορφή δεδομένων, τα οποία αποτελούν αντικείμενο επεξεργασίας από ανθρώπινα σχεδιασμένους αλγόριθμους. Οι αλγόριθμοι αυτοί, όπως και οι μέθοδοι με τις οποίες τα δεδομένα αυτά συλλέγονται, είναι επηρεασμένα από τις αξίες και τις ανθρώπινες οπτικές σχετικά με τον κόσμο και την πραγματικότητα. Ο σχεδιασμός των αλγορίθμων δεν αποτελεί μια διαδικασία η οποία διέπεται από επιστημονική αντικειμενικότητα και ουδετερότητα. Ο κώδικας ο οποίος χρησιμοποιείται από τους προγραμματιστές για να συγκροτήσουν τους αλγόριθμους αυτούς αλλά και για να ορίσουν τον τρόπο με τον οποίο αυτοί θα λειτουργούν αποτελεί πέρα από τεχνολογική μια έντονα κοινωνική διαδικασία. Στην αμφίδρομη σχέση αλληλοδιαμόρφωσης τεχνοεπιστημονικών κατασκευών και κοινωνίας (Sismondo, 2016) ιδιαίτερη σημασία έχει ο τρόπος με τον οποίο διακρίσεις και μεροληψίες εισχωρούν στον κώδικα με τον οποίο τα συστήματα αυτά προγραμματίζονται (Benjamin ,2019; Airoidi, 2021).

Ένα χαρακτηριστικό που εμφανίζουν πολλές φορές τα αλγοριθμικά συστήματα είναι η μετατροπή τους σε μαύρα κουτιά. Ο όρος αυτός χρησιμοποιείται από πολλούς επιστήμονες προκειμένου να αναδειχθεί η άγνοια και η αορατότητα που διέπει τα αυτοματοποιημένα συστήματα λήψης αποφάσεων καθώς και τα συστήματα ταξινόμησης ρίσκου. Οι λειτουργίες των αλγορίθμων που χρησιμοποιούνται στα συστήματα τεχνητής νοημοσύνης πολλές φορές πραγματοποιούν διαδικασίες, οι οποίες είναι άγνωστες στους πολίτες για τους οποίους οι τεχνολογίες αυτές προορίζονται (Pasquale, 2016). Οι διαδικασίες αυτές, είτε εξαιτίας των δεδομένων που αξιοποιούνται από τον αλγόριθμο (Suresh & Guttag, 2021) ή εξαιτίας του σχεδιασμού τους, πολλές φορές συνεπάγονται την παραγωγή μεροληπτικών αποτελεσμάτων και αποφάσεων. Η έλλειψη γνώσης των κοινωνικών υποκειμένων σχετικά με την λειτουργία τέτοιων τεχνοεπιστημονικών κατασκευών καθιστά τις Κοινωνίες Τεχνητής Νοημοσύνης, «Κοινωνίες του Μαύρου Κουτιού» (Pasquale, 2016). Υπό αυτό το πρίσμα τα τεχνολογικά κουτιά της τεχνητής νοημοσύνης, των αλγορίθμων και της

αυτοματοποιημένης λήψης αποφάσεων αποτελούν νέες, παγκόσμιες μορφές διακινδύνευσης (Beck, 2015), οι οποίες λαμβάνουν μέρος σε πολυδιάστατα κατακερματισμένα ψηφιακά περιβάλλοντα.

Ένα χαρακτηριστικό παράδειγμα μεροληπτικής εφαρμογής τεχνολογίας τεχνητής νοημοσύνης είναι η περίπτωση των ολλανδικών παροχών παιδικής μέριμνας, γνωστή ως «Toeslagenaffaire». Οι παροχές παιδικής μέριμνας αφορούν γονείς οι φροντιστές από αδύναμες κοινωνικοοικονομικές ομάδες οι οποίες προκειμένου να ανταπεξέλθουν στις οικονομικές τους υποχρεώσεις λαμβάνουν αυτή την μορφή κρατικής ενίσχυσης. Το ολλανδικό κράτος έπειτα από περιπτώσεις οικονομικής εκμετάλλευσης που βίωσε (Dutch Data Protection Office, 2020), αυτοματοποίησε πολλές διαδικασίες οι οποίες προηγουμένως πραγματοποιούνταν από κρατικούς υπαλλήλους. Έτσι, η διαδικασία ελέγχου εγκυρότητας των αιτήσεων αντικαταστάθηκε από ένα αυτοματοποιημένο σύστημα ταξινόμησης ρίσκου το οποίο είχε βαθμολογήσει τις αιτήσεις με υψηλό κίνδυνο εξαπάτησης του προγράμματος (Amnesty International, 2021).

Το σύστημα αυτό αποδείχτηκε ότι βαθμολογούσε με υψηλό βαθμό ρίσκου τις αιτήσεις ολλανδών πολιτών με διπλή ιθαγένεια (Amnesty International, 2021). Σύμφωνα με την νομοθεσία του προγράμματος, η ιθαγένεια των πολιτών όχι απλώς δεν αποτελούσε παράγοντα που επηρεάζει το αποτέλεσμα της βαθμολόγησης της αιτήσεως του πολίτη, αλλά αυτός ο μεροληπτικός τρόπος επεξεργασίας αυτών των ευαίσθητων προσωπικών δεδομένων δεν κοινοποιήθηκε ποτέ στους πολίτες. Η διασφράγιση του μαύρου κουτιού του συστήματος ρίσκου σε συνάρτηση με την μεροληπτική αντιμετώπιση πολλών Ολλανδών πολιτών, είχε ως αποτέλεσμα την διακοπή της χρηματοδότηση που προηγουμένως τους προσφερόταν, την επιστροφή μεγάλων χρηματικών ποσών, τα οποία είχαν όμως χρησιμοποιηθεί, και εν τέλει την απομάκρυνση πολλών παιδιών από την οικογενειακή τους εστία (NI.Times, 2021; Pascoe, 2022).

Όπως και άλλα μεροληπτικά συστήματα που αναφέρονται εν συνεχεία έτσι και το σύστημα ταξινόμησης ρίσκου, αξιοποιήθηκε στα πλαίσια της αντικειμενικότητας και της ουδετερότητας. Οι αξιώσεις αυτές πορεύονται κατά κύριο λόγο από την πεποίθηση ότι τα συστήματα τεχνητής νοημοσύνης εξαιτίας της συγκρότησής τους από επιστήμονες και τεχνολόγους δεν πραγματοποιούν σφάλματα και δεν μεροληπτούν. Η Κοινωνιολογία της Τεχνητής Νοημοσύνης, πέρα από τους τρόπους με τους οποίους τα συστήματα αυτά επηρεάζουν την καθημερινή ζωή των πολιτών εξετάζει την

διαμόρφωση των συστημάτων αυτών ως μια κοινωνική διαδικασία, στην οποία οι κοινωνικοί δρώντες διαμορφώνουν με τις αξίες και τις ενέργειές τους ο ένας τον άλλο. Η δυνατότητα της Κοινωνιολογίας να λαμβάνει μέρος σε αυτές τις πραγματεύσεις, εκπορεύεται σε μεγάλο βαθμό από τις Σπουδές Επιστήμης και Τεχνολογίας ή αλλιώς STS.

Οι Σπουδές Επιστήμης και Τεχνολογίας προσφέρουν ένα ιδιαίτερα γόνιμο πεδίο για την ανάλυση των σύγχρονων κοινωνικοτεχνικών συστημάτων. Μέσω αυτών αξιοποιούνται διάφορες κοινωνικές θεωρίες, οι οποίες εξετάζουν την ζητήματα όπως η ουδετερότητα της επιστημονικής διαδικασίας ή η κοινωνικότητα των τεχνολογικών κατασκευασμάτων. Ειδικότερα, μελετούν τους τρόπους όπου η επιστήμη και τεχνολογία μετατρέπονται από δύο ξεχωριστά μεταξύ τους πεδία δημιουργίας και εφαρμογής της γνώσης, σε ένα ενιαίο πεδίο το οποίο από διάφορους διανοητές και επιστήμονες του χώρου χαρακτηρίζεται «τεχνοεπιστήμη» (Γεωργοπούλου, 2010; Sismondo, 2016). Με τον όρο αυτό αναδεικνύεται η ανάμειξη τόσο επιστημόνων όσο και τεχνολόγων στην διαδικασία παραγωγής και αξιοποίησης συστημάτων όπως αυτό της τεχνητής νοημοσύνης, καθώς και η αδυναμία ξεκάθαρα προσδιορισμού των ρόλων που αυτοί διατελούν στις προαναφερθείσες διαδικασίες.

Στην συγκεκριμένη εργασία θα αναλυθούν αυτές οι θεωρίες που μελετούν την κοινωνική διαμόρφωση των επιστημών και τεχνολογιών υπό το πρίσμα των Σπουδών Επιστήμης και Τεχνολογίας . Στο πρώτη ενότητα του κυρίου θέματος της εργασίας θα εξετασθούν οι θεωρίες σχετικά με την επιστήμη και την τεχνολογία και εν τέλει την συγκρότηση του τεχνοεπιστημονικού. Η ανάλυση αυτή θα αποτελέσει στέρεο έδαφος για την καλύτερη κατανόηση της δυναμικής των Κοινωνιών της Τεχνητής Νοημοσύνης. Για την ανάλυση των κοινωνιών αυτών θα αξιοποιηθούν αναλύσεις προερχόμενες από τις κριτικές σπουδές των αλγορίθμων (Critical Algorithm Studies), ένα πεδίο με άμεση σύνδεση και επιρροές από τις STS. Εντός αυτών των πλαισίων θα συζητηθεί ο τρόπος με τον οποίο τα τεχνοεπιστημονικά συστήματα μετατρέπονται σε κοινωνικοτεχνικά συστήματα, τα οποία όχι μόνο διαμορφώνονται κοινωνικά, αλλά ταυτόχρονα διαμορφώνουν την κοινωνία. Στην τρίτη και τελευταία ενότητα θα αναλυθεί η περίπτωση του ολλανδικού δημοσίου τομέα μέσω της μεροληπτική εφαρμογής του μαύρου κουτιού του συστήματος ταξινόμησης ρίσκου στα πλαίσια της αυτοματοποιημένης διακυβέρνησης.

Α.Θεωρίες Επιστήμης και Τεχνολογίας: Η Συμβολή των Σπουδών Επιστήμης και Τεχνολογίας

Α.1 Το φιλοσοφικό ρεύμα του Θετικισμού και οι επιρροές του σε Επιστήμη και Τεχνολογία

Οι κοινωνίες της τέταρτης βιομηχανικής επανάστασης χαρακτηρίζονται σε μεγάλο βαθμό από την ραγδαία και ταχεία τεχνολογική ανάπτυξη, την αυτοματοποίηση, την επανάσταση των δεδομένων, των αλγορίθμων και της τεχνητής νοημοσύνης. Οι κλάδοι αυτοί, ακόμα και αν έχουν διεισδύσει σημαντικά στην καθημερινότητα των ανθρώπων, αντιλαμβάνονται την πραγματικότητα με όρους θετικούς. Η έννοια θετικός σε αυτή την περίπτωση δεν χρησιμοποιείται με σκοπό να αναδείξει την επωφελή επιρροή των σύγχρονων τεχνοεπιστημονικών ευρημάτων και καινοτομιών. Χρησιμοποιείται για να τονίσει την ισχυρή επιρροή της θετικιστικής προοπτικής, την επιρροή του θετικισμού σε όλες τις εκφάνσεις της κοινωνικής ζωής και πραγματικότητας.

Ενώ λοιπόν κάποιος θα πίστευε ότι ο Θετικισμός αφορά κατά κύριο λόγο τον χώρο των θετικών επιστημών, όπως την επιστήμη της φυσικής, των μαθηματικών, της χημείας, της βιολογίας ή πιο σύγχρονα διαμορφωμένους επιστημονικούς κλάδους όπως την επιστήμη των δεδομένων ή αυτόν της τεχνητής νοημοσύνης, οφείλει να γνωρίζει ότι έχει ασκήσει σημαντική επιρροή επιπλέον, στις ανθρωπιστικές και κοινωνικές επιστήμες. Για πολλά χρόνια οι μεθοδολογικές πρακτικές του θετικισμού ταύτιζαν τον θετικισμό με την επιστήμη κατ' επέκταση και την αλήθεια.

Αν και πολλά στοιχεία τα οποία δομούν την έννοια του θετικισμού είχαν εμφανιστεί ήδη από τα χρόνια των αρχαίων Ελλήνων φιλοσόφων, ο θετικισμός, ως φιλοσοφικό ρεύμα εκφράστηκε με στέρεο τρόπο τον 19ο αιώνα, με σημαντικό εκφραστή του τον Αύγουστο Κόντ (Auguste Comte). Η θετικιστική φιλοσοφία του Κοντ έχει τις καταβολές της στην ιδέα ότι η γνώση πρέπει να προκύπτει σαν αποτέλεσμα της παρατηρήσιμης, εμπειρικής πραγματικότητας και θα πρέπει να υποβάλλονται σε αυστηρό επιστημονικό έλεγχο (Coser, 1977). Σύμφωνα με τον Κοντ, η ανθρώπινη νόηση και σκέψη υπάγεται στον «Νόμο των Τριών Σταδίων» βάσει του οποίου, οι παραπάνω έννοιες, περνούν από τρία στάδια: το Θεολογικό, το Μεταφυσικό και το Θετικιστικό. Το τελευταίο χαρακτηρίζεται από την εμπειρική παρατήρηση σε

συνάρτηση επιστημονική τεκμηρίωση και θεωρείται η ύψιστη μορφή ανθρώπινης διανόησης (Comte, 1853; The Editors of Encyclopaedia Britannica, 2020) .

Η σημασία της έκφρασης του Θετικισμού από τον Κοντ είναι μεγάλη καθώς η θεώρηση αυτή θα εκφράσει τα χαρακτηριστικά με τα οποία θα πορευτεί η έννοια της επιστήμης, ακόμα και στον χώρο των κοινωνικών επιστημών, μέχρι την αντιουσιακρατική έλευση του Κοινωνικού Κονστρουκτιβισμού η οποία με την σειρά της θα επηρεάσει σημαντικά την κατανόηση της έννοιας του επιστημονικού (Γεωργοπούλου, 2010, σσ.27-41).

A.2 Θετικισμός στην Επιστήμη και στην Τεχνολογία: Η Δυϊστική Προοπτική

Η θετικιστική επιρροή δεν περιορίστηκε μονάχα στον τομέα της επιστήμης. Η δυνατότητα πρόβλεψης, η αναγωγιστική λογική, η κλειστότητα και περιπλοκότητα των τεχνολογικών συστημάτων συγκροτούν ένα γόνιμο πεδίο για την εφαρμογή της θετικιστικής λογικής σε τεχνολογικές κατασκευές και δημιουργίες. Κάτι τέτοιο έγινε ορατό και την περίοδο της βιομηχανικής επανάστασης με την δημιουργία της μηχανής του ατμού από τον Τζέιμς Βατ (James Watt) (1736-1819). Η κατασκευή της βασίστηκε σε πειραματισμούς και εμπειρικές μελέτες καταφέροντας να αλλάξει την ζωή των πολιτών και την ροή του εμπορίου εκείνη την χρονική περίοδο. Παράλληλα, η θετικιστική προσέγγιση που εφαρμόστηκε κατά την δημιουργία και ανάπτυξη αυτής της τεχνολογικής κατασκευής, επηρέασε με την σειρά της τις παραγωγικές διαδικασίες, καθώς και την επιστήμη της μηχανικής (Hills, 2002).

Ενώ λοιπόν ο θετικισμός επηρεάζει σημαντικά την επιστήμη και την τεχνολογία, λειτουργεί αντιλαμβανόμενος αυτά τα δύο πεδία ως ξεχωριστά ή πολλές φορές θεωρεί το πρώτο προέκταση του δεύτερου (Sismondo, 2016, σσ.13-30). Αυτό συμβαίνει διότι η θετικιστική λογική αποτελεί μια λογική του δυισμού. Επιχειρεί να διακρίνει τον άνθρωπο από την φύση, καθώς η ανθρώπινη φύση διαφέρει σημαντικά από τον υπόλοιπο φυσικό κόσμο ο οποίος υπάγεται σε νομοτέλειες. Με λίγα λόγια, ο άνθρωπος δεν προέχει της φύσης ούτε της λογικής που την διέπει. Αποτελεί ένα κομμάτι της του οποίου η πολυπλοκότητα διαφέρει της φυσικής περιπλοκότητας.

Η αιτιοκρατική λογική του θετικισμού και η ανάγκη του για προβλέψεις θα συνεχίσουν και τις επόμενες δεκαετίες να επηρεάζουν φιλοσοφικά τόσο την επιστήμη όσο και την τεχνολογία. Τα χαρακτηριστικά του θα επηρεάσουν και άλλα φιλοσοφικά ρεύματα που θα μετατραπούν σε παρακλάδια της θετικιστικής αντίληψης, όπως ο

Λογικός Θετικισμός, ο οποίος θα επηρεάσει σημαντικά την εικόνα για την φύση της επιστήμης ενώ παράλληλα θα κινητοποιήσει και θα προβληματίσει διανοούμενους και επιστήμονες σχετικά με την φύση του επιστημονικού.

A.3 Λογικός θετικισμός και Κύκλος της Βιέννης

Σε πολλές συζητήσεις σχετικά με την επιστήμη και τις φιλοσοφικές θεωρίες είναι δύσκολο να μην πραγματοποιηθούν αναφορές, πόσο μάλλον συνδέσεις μεταξύ του Κύκλου της Βιέννης και του Λογικού Θετικισμού. Ο Κύκλος της Βιέννης αποτελεί ομάδα φιλοσόφων που δραστηριοποιήθηκαν όσον αφορά το θέμα της επιστήμης στις αρχές του 1930. Στόχος του ήταν να επεκτείνει την φιλοσοφική επιστημονική κοσμοθεωρία μέσω της συγκρότησης ενός θετικιστικού, εμπειρικού φιλοσοφικού πλαισίου. Ξεκίνησε ως ένα γενικό πρόγραμμα στο οποίο συμμετείχαν αρκετοί επιστήμονες και διανοούμενοι όπως ο R.Carnap, ο M.Schlick κ.α, επηρεασμένοι από τον κλάδο των θετικιστικών επιστημών (Puntel, 1999).

Σημαντικό σημείο στην εστίαση της οπτικής των θετικιστών διανοούμενων ήταν εννοιολογικά ζητήματα που αφορούσαν, όχι μόνο τους δικά τους επιστημονικά πεδία μελέτης, αλλά την ίδια την επιστήμη και την φιλοσοφία γύρω από την οποία συγκροτείται. Ο λογικός θετικισμός βασίζει την επιστημονική οπτική του στον Θετικισμό (Ναγόπουλος, 2015, σσ.83-85) σε βαθμό που πολλές φορές, κυρίως από τους θετικιστές θεωρείται ως συνέχεια του ενώ από άλλους ερευνητές θεωρείται ως ένα παρακλάδι του.

Κεντρική ιδέα, η οποία τροφοδότησε την οπτική του λογικού θετικισμού όσον αφορά την επιστήμη, ήταν η Αρχή της Επαλήθευσης (Kazemier, 1946). Σύμφωνα με την Λογική της Επαλήθευσης, το επιστημονικό νόημα μιας θέσης έγκειται στην ικανότητα να επαληθευτεί εμπειρικά μέσω της παρατήρησης ή της εμπειρίας. Η εμπειρική επαλήθευση των επιστημονικών θέσεων πέρα από την χρησιμότητα την εξέταση των επιστημονικών τεκμηρίων, πραγματοποιεί ταυτοχρόνως έναν διαχωρισμό ανάμεσα σε θέσεις με επιστημονικό νόημα και απλούς ισχυρισμούς. Έτσι, στην περίπτωση του λογικού θετικισμού, το νόημα και η εμπειρική επαλήθευση συνιστούν το πραγματικά επιστημονικό.

Η διαχωριστική ικανότητα της αρχής αυτής διέπεται έντονα από την δυϊστική λογική του Θετικισμού. Η επιστημονικότητα ή μη, το νόημα ή όχι, η επαλήθευση ή διάψευση μιας θέσης σχετίζονται με το αν αυτός που την εκφράζει, δύναται να την

επιβεβαιώσει μέσω της εμπειρίας. Με αυτόν τον τρόπο, θωρακίζεται ο όρος επιστήμη και λαμβάνει τυπικό χαρακτήρα αφού όσο περισσότερα εμπειρικά στοιχεία εντοπίζονται στον φυσικό κόσμο, τόσο ορθότερη γίνεται μια θεωρία. Πρόκειται για την απόπειρα συγκρότησης ενός επαγωγικού πλαισίου από τους ειδικούς, η οποία στοχεύει στην συγκρότηση ενός συμπαγούς επιστημονικού πλαισίου που θα μπορεί να εφαρμοστεί καθολικά.

Συνεπώς η ανάπτυξη θεωριών και κατ' επέκταση η ανάπτυξη της επιστημονικής γνώσης για τον Λογικό θετικισμό νοηματοδοτείται έντονα μέσω της συστηματικά επεξεργασμένης πληροφορίας από τους ειδικούς και την αξιολόγηση και επαλήθευσή τους μέσω των δικλίδων της παρατήρησης και της εμπειρίας.

A.4 Καρλ Πόπερ: Η διάκριση ανάμεσα στην αληθινή και ψευδή επιστήμη

Ο Καρλ Πόπερ (Carl Popper) είναι φιλόσοφος ο οποίος ασχολήθηκε έντονα με την επιστήμη της φιλοσοφίας. Έδειξε ενδιαφέρον για την φύση της επιστήμης και τα κριτήρια συγκρότησής της. Δέχθηκε σημαντική επιρροή από τον Λογικό θετικισμό και τον Κύκλο της Βιέννης, συνέβαλε στην ανάπτυξη του «κριτικού ορθολογισμού» (Ναγόπουλος, 2015, σσ.156-159) ενώ ιδιαίτερη σημασία για εκείνον είχε η διαδικασία με την οποία θα ήταν δυνατός ο διαχωρισμός ανάμεσα σε επιστήμες και μη επιστήμες. Σε αυτό το πλαίσιο, ο Πόπερ διατύπωσε την θέση του σχετικά με την Διαψευσιοκρατία.

Η Διαψευσιοκρατία συνιστά λοιπόν μια ιδέα περί του τι είναι επιστημονικά αληθινό και τι ψευδές. Βασίζεται γενικότερα στον δυισμό που εμφανιζόταν πολύ έντονα μέχρι τότε στις νεωτερικές επιστήμες, οι οποίες δέχθηκαν ισχυρή επιρροή από το θετικιστικό τρόπο σκέψης των Νευτώνειων επιστημών (Γεωργοπούλου, 2010, σσ.69-73). Οποιαδήποτε θεωρία, σύμφωνα με τον Πόπερ, μπορεί να θεωρηθεί επιστημονική από την στιγμή που μπορεί να πραγματοποιήσει προβλέψεις σχετικά με το φαινόμενο που πραγματεύεται. Από την στιγμή όμως που η θεωρία αυτή δεν οδηγείται σε πρόβλεψη, πρέπει να αντιμετωπίζεται ως μη αληθής, μη επιστημονική θεωρία γεγονός που επεκτείνεται βέβαια και στην επιστήμη. Μια από τις σημαντικές αρχές της πραγματικής επιστήμης για τον Πόπερ, είναι η ριψοκινδυνότητα της διάψευσης.

Ειδικότερα, κάθε θεωρία που αναπτύσσεται σε οποιοδήποτε επιστημονικό κλάδο λαμβάνει το «ρίσκο» έκθεσής της στην επιστημονική κοινότητα. Η θεωρία αυτή αξιολογείται αυστηρά από την επιστημονική κοινότητα και η εγκυρότητά της έγκειται,

όπως προαναφέρθηκε, στην ικανότητα να προβλέπει. Σε περίπτωση που μια θεωρία αδυνατεί να προβλέψει κρίνεται ως μη επιστημονική. Αντίθετα, αν προβεί σε πρόβλεψη η θεωρία αυτή αντιμετωπίζεται ως επιστημονική με ένα όμως χρονικό όριο.

Η επιστημονική αλήθεια διατηρεί τον χαρακτήρα της έως την διάψυσή της. Ακριβώς επειδή αποτελεί κοινωνική κατασκευή, η διάρκεια της επιστημονικής της ιδιότητας διατηρείται έως ότου μια άλλη επιστημονική θεωρία καταφέρει και οδηγήσει σε καλύτερη πρόβλεψη από την προγενέστερη. Στη διαψευσιοκρατία, η ανθρώπινη διαισθητικότητα είναι ένα κεντρικό στοιχείο πάνω στο οποίο δομείται η επιστημονικότητα. Αφενός οριοθετεί την λογική, μέσω της ανθρωποκεντρικής αντίληψής της, αφετέρου καθιστά την επιστήμη μια κοινωνική διαδικασία διεύρυνσης της αλήθειας μέσω της συνεχούς απόρριψης και προσθήκης αυστηρά αξιολογούμενων επιστημονικών θεωριών.

Στη Διαψευσιοκρατία λοιπόν, γίνεται ξεκάθαρο ότι το ζητούμενο κάθε πραγματικής επιστήμης είναι η πρόβλεψη. Για τον Πόπερ, το ρίσκο της πρόβλεψης συνιστά μια μορφή κοινωνικής αποδοχής, μια κοινή παραδοχή μεταξύ της επιστημονικής και ερευνητικής κοινότητας στον βωμό της προόδου της επιστήμης. Πέρα όμως από αυτή την κοινή παραδοχή, η ίδια η επιστημονική κοινότητα οφείλει να ακολουθεί κάποια κριτήρια τα οποία θα οριοθετούν και ταυτόχρονα θα οχυρώνουν την φύση της πραγματικής επιστήμης από το ψεύδος. Υπό αυτές τις συνθήκες, ο Πόπερ εισαγάγει τον όρο του κριτηρίου της διάψευσης.

Το κριτήριο της διάψευσης συνιστά ένα ιδιαίτερα σημαντικό στοιχείο της Διαψευσιοκρατίας. Η ανάγκη της επιστήμης να αποτελέσει ένα ξεχωριστό πεδίο στην κοινωνική ζωή συνεπάγεται την συγκέντρωση κάποιων χαρακτηριστικών τα οποία θα την ξεχωρίζουν από οποιαδήποτε άλλη μορφή απλής κοινωνικής γνώσης. Ο έντονος αυτός διαχωρισμός, μεταξύ επιστήμης και μη επιστήμης, τονίζει τους έντονους δυϊσμούς ανάμεσα στο αληθινό και το ψευδές, το επιστημονικό και το μη επιστημονικό. Μάλιστα, συνδέεται άμεσα με την Αρχή της Επαλήθευσης, που εκφράστηκε από τους λογικού θετικιστές διαχωρίζοντας όμως όχι με σκοπό απλώς την επιβεβαίωση μιας θεωρίας αλλά με στόχο να αναδείξει εάν αυτή είναι επιστημονική ή όχι. Αυτό το κριτήριο οριοθέτησης της πραγματικής επιστήμης βασίζεται έντονα στον εμπειρικό ορθολογισμό, έτσι όπως αυτός αναδύθηκε από τον Κύκλο της Βιέννης, δηλαδή την αυστηρή εξέταση των ελεγχόμενων υποθέσεων μέσα από εμπειρικούς ελέγχους και δοκιμές (Popper, 1959; Khun, 1962; Sismondo 2016). Τα κριτήρια διάψευσης των επιστημονικών θέσεων, όπως αυτά προαναφέρθηκαν, δεν έχουν κάποια αξία εάν αυτά

δεν εφαρμόζονται . Η θέση του Πόπερ σχετικά με την πραγματική επιστήμη δεν αφορά την επιστήμη καθαυτή. Πρόκειται για μια στάση, η οποία πρέπει να δεχθεί εφαρμογή από την ερευνητική κοινότητα.

Μέσα από την ανάγνωση του έργου του Πόπερ μπορεί κάποιος να κατανοήσει την επιρροή του Θετικισμού στην αντίληψή του σχετικά με την Επιστήμη. Ο Λογικός θετικισμός όπως αυτός εκφράστηκε κατά κύριο λόγο μέσα από τον Κύκλο της Βιέννης τροφοδότησε την εικόνα του Πόπερ σχετικά με την ουσιοκρατική αντίληψη των δεισμών. Η ποπεριανή αντίληψη της επιστήμης υιοθετεί τις στέρεες θέσεις που δόμησαν την θετικιστική αντίληψη σχετικά με την ικανότητα εμπειρικής, λογικής, αιτιακής, τεκμηρίωσης, των οποίων η αδυναμία πραγμάτωσης συνιστά μια ταυτόχρονη διάψευση, έναν αποκλεισμό της εκάστοτε θεωρίας από την αλήθεια και την επιστήμη. Πρέπει όμως να τονιστεί ότι πέρα από την έντονη παρουσία του Θετικισμού, ο Πόπερ αντιλαμβάνεται την επιστήμη και την έρευνα ως (και) μια κοινωνική διαδικασία. Σε αντίθεση με την λογική του Θετικισμού, της φυσικής αιτιοκρατίας της σταθερής νομοτέλειας και του «λογικού», φυσικού κόσμου ο Πόπερ αντιλαμβάνεται την επιστήμη ως μια κοινωνική διαδικασία, ένα κοινωνικό κατασκευάσμα. Η κοινωνική πτυχή της επιστήμης δεν προϋποθέτει την εξάλειψη οποιασδήποτε επιστήμης, η οποία ενέχει την αντικειμενικότητα και την αυστηρότητα, την δυνατότητα και ανάγκη για προβλέψεις. Έτσι η επιστημολογία κάθε επιστήμης, οι μεθοδολογίες που υιοθετούνται, οι επιστήμονες, τα πρόσωπα που τις αναπτύσσουν και τις χρησιμοποιούν, πρέπει να λειτουργούν συστηματικά, προσεγγίζοντας την αλήθεια.

A.5. Κοινωνικός Κονστрукτιβισμός

Η κοινωνική κατασκευή συνιστά έναν όρο ο οποίος εμφανίζεται τον 20ο αιώνα, χρησιμοποιείται έντονα στην Κοινωνιολογία αλλά και στις κοινωνικές επιστήμες γενικότερα, με καταβολές της, τις αναφορές της από τους κλασσικούς, κοινωνικούς θεωρητικούς όπως ο Μαξ Βέμπερ. Σαν πεδίο επιστημονικής μελέτης αναπτύχθηκε αργότερα χρονικά , την δεκαετία του 1960, μέσω της συμβολής του έργου του Πίτερ Μπέργκερ (Peter Berger) και Τόμας Λούκμαν (Thomas Luckmann), «Η Κοινωνική Κατασκευή της Πραγματικότητας» (1967) στο οποίο αναδεικνύεται η κοινωνική κατασκευή της γνώσης και η καθοριστική επίδραση του κοινωνικού παράγοντα στην δημιουργία οποιουδήποτε γνωσιακού κατασκευάσματος.

Ο θεματικός άξονας του Κοινωνικού Κονστρουκτιβισμού αναπτύσσεται βάσει της διαφωνίας με την θέση σχετικά με μια αυθύπαρκτη μορφή γνώσης. Αντιθέτως, η γνώση κατασκευάζεται μέσω της αλληλεπίδρασης ατόμων και ομάδων με το κοινωνικό και ιστορικό τους πλαίσιο. Η προοπτική αυτή υποστηρίζει ότι η κατανόηση της πραγματικότητας διαμορφώνεται από τη γλώσσα, τον πολιτισμό και τους θεσμούς εντός των οποίων λειτουργούμε (Berger & Luckmann, 1967). Η κοινωνία, η γνώση, η επιστήμη και η τεχνολογία, δεν αποτελούν προϋπάρχουσες οντότητες στον χώρο και τον χρόνο αλλά κατασκευές, αντικείμενα που δομούνται πρωτίστως κοινωνικά (Baron & Gomez, 2016).

Η προοπτική του κοινωνικού κονστρουκτιβισμού υποστηρίζει ότι η γνώση δεν είναι ουδέτερη, αλλά αντίθετα επηρεάζεται από το κοινωνικό και ιστορικό πλαίσιο στο οποίο παράγεται (Bijker και άλλοι, 1989). Αυτό σημαίνει ότι η επιστημονική και τεχνολογική γνώση δεν ανακαλύπτεται, αλλά κατασκευάζεται από τους επιστήμονες και τους τεχνολόγους σε αλληλεπίδραση με το κοινωνικό και ιστορικό τους πλαίσιο (Latour & Woolgar, 1986; Law, 1990). Επισημαίνει επίσης το ρόλο της εξουσίας και των συμφερόντων στη διαμόρφωση της επιστημονικής και τεχνολογικής γνώσης, καθώς και τους τρόπους με τους οποίους η γνώση μπορεί να χρησιμοποιηθεί για τη διατήρηση ή την αμφισβήτηση των υφιστάμενων δομών εξουσίας.

Μια από τις βασικές συνεισφορές του κοινωνικού κονστρουκτιβισμού είναι η προσοχή του στους τρόπους με τους οποίους η γνώση παράγεται, διαδίδεται και λαμβάνεται στην κοινωνία. Επισημαίνει τους τρόπους με τους οποίους διαφορετικές κοινωνικές ομάδες έχουν πρόσβαση σε πόρους που τους επιτρέπουν να παράγουν, να διαδίδουν και να λαμβάνουν γνώση και πώς αυτό μπορεί να οδηγήσει σε ανισότητες στην εξουσία και τα προνόμια που απολαμβάνουν ορισμένες κοινωνικές ομάδες .

Συνολικά, ο κοινωνικός κονστρουκτιβισμός προσφέρει μια κριτική προοπτική για τη σχέση μεταξύ γνώσης και κοινωνίας και παρέχει πληροφορίες για τους τρόπους με τους οποίους η κατανόηση του κόσμου διαμορφώνεται από κοινωνικούς παράγοντες. Πρόκειται για μια προοπτική που έχει σημαντικές επιπτώσεις σε διάφορους κοινωνικούς θεσμούς, όπως η εκπαίδευση, η επιστήμη και η τεχνολογία. Η φιλοσοφία του Κοινωνικού Κονστρουκτιβισμού επηρεάζει αποφασιστικά τις σπουδές Επιστήμης και Τεχνολογίας και ειδικότερα την μελέτη αυτών όσον αφορά την τεχνητή νοημοσύνη και ζητήματα που την αφορούν.

A.5.1. Κριτική στον Κονστρουκτιβισμό: η ηγεμονία της κατασκευής

Η δεκαετία του 1960 θεωρείται ένα κομβικό σημείο στην επιστημολογία της κοινωνική θεωρίας. Η καθιέρωση του Κονστρουκτιβισμού ως το πλέον ισχυρό θεωρητικό εργαλείο των κοινωνικών επιστημών ενίσχυσε σημαντικά την θέση του, αναδεικνύοντας το έντονα σχετικιστικό πνεύμα που επικρατεί στον χώρο των κοινωνικών επιστημών εκείνη την χρονική περίοδο, καθώς και τις επόμενες δύο δεκαετίες. Σε αυτό το χρονικό διάστημα ηγεμονεύει η ιδέα της «κατασκευής», η οποία παύει πλέον να έχει φυσικό χαρακτήρα. Η φύση, στερείται της ουσιαστικής σημασίας της, να αποτελεί δηλαδή τον τελικό κριτή στις ανθρώπινες θεωρήσεις και εκτιμήσεις μέσω της εμπειρικής αποδοχής ή απόρριψης. Η ριζοσπαστικότητα της κατασκευασιοκρατίας έγκειται στην ανάδυση μίας κοινωνικής οπτικής όπου τα πάντα ανάγονται σε κοινωνικές διαδικασίες.

Ένα από τα αποτελέσματα της θεωρητικής επικράτησης του Κονστρουκτιβισμού είναι η μετατόπιση από μια πιο θετικιστική κοινωνιολογική θεώρηση της πραγματικότητας, σε μια αντιθεμελιακή κοινωνιολογική επιστήμη (Γεωργοπούλου, 2010). Βέβαια, ακόμα και αν η θετικιστική προοπτική για την Κοινωνιολογία είχε ξεκινήσει να είναι αναχρονική, όπως και ο Κοντιανός όρος «κοινωνική φυσική», ο θετικισμός είχε ξεκινήσει να υστερεί όσον αφορά το περιεχόμενό του, με την εμφάνιση άλλων κοινωνικών τοποθετήσεων όπως αυτή του Βέμπερ σχετικά με την αξιολογική ουδετερότητα. Ο κοινωνικός κονστρουκτιβισμός μετέβαλε σημαντικά το ενδιαφέρον από την φύση στην κοινωνία, αφού μεταφέρει την ουσία και τους δυισμούς που την διέπουν, στον άνθρωπο, ο οποίος πλέον χαρακτηρίζεται από μια μεταφυσική κατάσταση.

Η ιδιότητα του ανθρώπου να κατασκευάζει μέσω της γνώσης του κόσμου που τον περιβάλλει, ανάγεται στην ικανότητα συγκρότησης ουσίας. Το επιστημολογικό ενδιαφέρον σταματά να εστιάζει στην εξωτερική, αμέτοχη παρατήρηση του Νεύτωνα και μεταφέρεται στον άνθρωπο. Με άλλα λόγια, η παρατήρηση της κίνησης του μήλου που πέφτει από την μηλιά στο έδαφος, ο υπολογισμός του βάρους του, της κίνησης της πτώσης του δεν έχει καμία πλέον αξία. Το επιστημονικό βλέμμα πλέον εστιάζει στο υποκείμενο που ερμηνεύει τα γεγονότα αυτά καθώς και το πως αυτά επικοινωνούνται και γίνονται κατανοητά μεταξύ των κοινωνικών υποκειμένων.

Από αυτή την οπτική, το κυρίαρχο εργαλείο κατασκευής της γνώσης πλέον είναι η γλώσσα του υποκειμένου και όχι αυτή της επιστήμης. Η δεύτερη μάλιστα,

αποτελεί προϊόν της πρώτης, μια κοινωνική κατασκευή που γίνεται κατανοητή εντός των δικών της πλαισίων. Μια τέτοια συνθήκη έχει ως αποτέλεσμα την σχετικοποίηση ακόμα και της μέχρι τότε γνωστής επιστημονικής πραγματικότητας. Η πραγματικότητα που έως τότε διεπόταν από γραμμικότητα και αιτιότητα, έχουν πλέον ανοικτό χαρακτήρα όπου όλα «παγιδεύονται στα δεσμά των λόγων και συνιστούν αποκλειστικό αποτέλεσμά τους» (Γεωργοπούλου,2010, σ.38)

Υπό την προϋπόθεση λοιπόν ότι η πραγματικότητα δεν εντοπίζεται στην φύση, η Κονστρουκτιβιστική προοπτική, ενώ δίνει μια σημαντική κατεύθυνση στον χώρο των κοινωνικών, και όχι μόνο, επιστημών συγκροτώντας μια νέα, αντιθεμελιακή, μεταφυσική προοπτική. Δημιουργεί παράλληλα μια νέα, προβληματική συνθήκη για την οποία θα δεχθεί έντονη κριτική, κυρίως από τον χώρο του ρεαλισμού (Sismondo,2016). Με την «ουσιοκαρτικοποίηση του Κονστρουκτιβισμού» (Γεωργοπούλου, 2010) οτιδήποτε εντοπίζεται στην πραγματικότητα, αποτελεί μια κατασκευή άνευ ουσίας, η οποία, σε τελική ανάλυση δεν είναι απαραίτητο να κατηγοριοποιηθεί ή να ξεχωρίσει από κάτι άλλο, αφού δεν αποτελεί παρά ένα κοινωνικό κατασκεύασμα.

Το ζήτημα αυτό επεκτείνεται σημαντικότερα στον βαθμό όπου και η ίδια η κονστρουκτιβιστική θεωρία, αποτελεί μια κοινωνική κατασκευή. Κάτι τέτοιο συγκροτεί επιπλέον προβληματικές όπως αν η κονστρουκτιβιστική Κοινωνιολογία χρησιμοποιεί τις θεωρίες κατασκευής ώστε να οργανώσει με αυτό τον τρόπο την πραγματικότητα ή αν απλώς τις χρησιμοποιεί για να αναδείξει την αταξία και ρευστότητα που την διέπουν. Πάντως σε κάθε μια από τις παραπάνω περιπτώσεις η κονστρουκτιβιστική Κοινωνιολογία γίνεται αντιληπτή ως «προνομιούχο πεδίο αναφοράς ή κλειδί ανάλυσης των πάντων» (Γεωργοπούλου,2010,σ.31).

Στις σπουδές της Επιστήμης, Κοινωνίας και Τεχνολογίας η κοινωνική κατασκευασιοκρατία ή αλλιώς ο Κοινωνικός Κονστρουκτιβισμός έχει επηρεάσει σημαντικά την διαμόρφωση και την ανάπτυξη του ίδιου του επιστημονικού κλάδου , ο οποίος δομείται βασιζόμενος στην φιλοσοφία του Κονστρουκτιβιστικού προτύπου, την ιδέα ότι τεχνολογία και επιστήμη αποτελούν κατά κύριο λόγο κοινωνικές κατασκευές (Sismondo,2016). Ο κονστρουκτιβισμός πέρα από ισχυρό θεωρητικό σημείο αναφοράς στις σπουδές Επιστήμης και Τεχνολογίας διευρύνθηκε σημαντικά μέσω της αξιοποίησής του από πολλούς θεωρητικούς όπως ο Μπρούνο Λατούρ,(Bruno Latour) και ο Στιβ Γούλγκαρ (Steve Woolgar) οι οποίοι θα παρουσιαστούν αναλυτικά εν συνεχεία.

A.5.2. Κοινωνικός κονστρουκτιβισμός και σπουδές Επιστήμης και Τεχνολογίας : Κατασκευή κοινωνικών και υλικών περιβαλλόντων

Η αμφίδρομη σχέση τεχνολογίας κοινωνίας και επιστήμης τονίζει την κατασκευή από περιβάλλοντα που κατασκευάζουν και κατασκευάζονται μέσω διαρκών αλληλεπιδράσεων. Για παράδειγμα, η ανάπτυξη των μαθηματικών μοντέλων βαθιάς μάθησης (Deep Learning Models), τα οποία συγκροτήθηκαν από ένα περιβάλλον που στελεχώνουν μαθηματικοί οδήγησαν στην συγκρότηση προβλεπτικών μοντέλων, σε κλειστά συστήματα, μέσω τεχνητής νοημοσύνης. Ένα κλειστό, υλικό περιβάλλον, όπως αυτό του παιχνιδιού του σκακιού, αποτέλεσε πρόσφορο έδαφος για την ανάπτυξη του προγράμματος τεχνητής νοημοσύνης «Deep Blue». Το πρόγραμμα αυτό, σε αγώνα που διεξήχθη μεταξύ αυτού και ενός από τους καλύτερους σκακιστές όλων των εποχών και τότε παγκόσμιου πρωταθλητή, Γιούρι Κασπάροφ (Garry Kasparov), λαμβάνοντας ένα αξιομνημόνευτο αποτέλεσμα. Μέσα σε λίγα παιχνίδια το «Deep Blue», πέρα από μια ήττα στο πρώτο του παιχνίδι, κατάφερε να κερδίσει τον Κασπάροφ, ο οποίος σε αντίθεση με το τεχνούργημα, είχε αφιερώσει τεράστιο μέρος από την ζωή του στην εκμάθηση σκακιού.

Το υλικό περιβάλλον του παιχνιδιού του σκακιού, το οποίο είχε συγκροτηθεί κοινωνικά, μέσα από επιστημονικές ομάδες και περιβάλλοντα, ώθησε άλλα τεχνοεπιστημονικά περιβάλλοντα να αναπτύξουν δικά τους λογισμικά τεχνητής εκμάθησης και νοημοσύνης να δραστηριοποιηθούν στον χώρο τόσο της τέχνης του σακακιού όσο και άλλων, παρόμοιων παιχνιδιών. Η πραγματικότητα αυτή βέβαια οφείλεται στην συνθήκη της απρόσμενης νίκης του υπολογιστή της IBM αφού «τέτοιες αναμετρήσεις όχι μόνο θα αύξαναν το προφίλ και θα ενίσχυαν τη φήμη των ερευνητών τεχνητής νοημοσύνης - αλλά θα ενίσχυαν επίσης τη σύνδεση μεταξύ ανθρώπινης και μηχανικής νόησης, τουλάχιστον όπως εφαρμόζεται στο σκάκι» (Esmenger, 2012, σ.20). Η αμφίδρομη κατασκευή περιβαλλόντων, δεν περιορίζεται μόνο μεταξύ της σχέσης επιστήμης και τεχνολογίας καθώς ιδιαίτερα κομβική όσον αφορά την συγκρότηση περιβαλλόντων είναι η πολιτική. Η πολιτική λειτουργεί καθοριστικά στην κατασκευή τεχνοεπιστημονικών περιβαλλόντων και το αντίστροφο. Ο αποκλεισμός της από την κατασκευασιοκρατική προοπτική θα περιόριζε σημαντικά ανάδυση και κατ' επέκταση θέαση υλικοκοινωνικών περιβαλλόντων, καθώς «σχεδόν καμία δράση είτε αφορά τους τομείς της υγείας, της οικονομίας, του περιβάλλοντος ή της άμυνας, δεν μπορεί να

αναληφθεί, εάν δεν συνδεθεί με τον ισχυρισμό ότι υποστηρίζεται από μια μελέτη», «οι κρατικές δράσεις είναι υπόλογες, απέναντι στην επιστημονική τεκμηρίωση» (Sismondo, 2016, σ. 108). Προφανώς, η επιστήμη και οι τεχνολογίες που αξιοποιούνται από αυτήν, δεν κατασκευάζουν μονάχα τα πολιτικά περιβάλλοντα, αλλά διαμορφώνονται και αυτές αποφασιστικά από αυτά.

Σχετικό παράδειγμα αποτελεί η τεχνολογία αναγνώρισης προσώπων, «Facial Recognition Technology» (Κουρούτζας, 2018, σσ.541-545). Η τεχνολογία αυτή, ξεκίνησε την ανάπτυξή της την δεκαετία του 1960 όπου και χρησιμοποιήθηκε για την ανάλυση και επεξεργασία εικόνων μέσω υπολογιστή με την χρήση αλγορίθμων. Αργότερα, από την «κλειστή» στα γραφεία χρήση τους μετατράπηκαν και σε εργαλείο για την επιβολή του νόμου και της κρατικής ασφάλειας στις Ηνωμένες Πολιτείες της Αμερικής από το FBI. Οι τεχνολογίες αυτές έχουν γενικότερα έχουν αποτελέσει αντικείμενο έντονης κριτικής σχετικά με τον τρόπο που διαχειρίζονται τα δεδομένα τα οποία καταγράφουν και συλλέγουν (Κουρούτζας, 2018, σσ.541-545), δημιουργώντας συχνά τον κίνδυνο βιοπολιτικών μορφών επιτήρησης .

Τα παραπάνω παραδείγματα τονίζουν την δυναμική της αμφίδρομης κατασκευής υλικών και κοινωνικών περιβαλλόντων μέσω της συνεισφοράς επιστήμης και της τεχνολογίας καθώς και την σημασία της κοινωνικής κατασκευασιοκρατίας ως εργαλείο κατανόησης της εξέλιξης των κοινωνικών συστημάτων. Ωστόσο οι κλασσικές κονστουκτιβιστικές θεωρίες, ακόμα και αν κατανοούν την κατασκευαστική διαδικασία πρωτογενώς ως διαδικασία παραγωγής γνώσης και δευτερογενώς ως διαδικασία που το κοινωνικό συναντά την ύλη, αντιλαμβάνονται την δεύτερη ως απαθή, στάσιμη, χειραγωγίσιμη (Γεωργοπούλου, 2010).

Ο ανοικτός, πολύπλοκος χαρακτήρας του σύγχρονου κόσμου απαιτούσε πλέον μια νέα, αναθεωρημένη κατεύθυνση που ενώ θα διατηρεί την κατασκευασιοκρατική προοπτική, την σημασία του κοινωνικού παράγοντα στην διαμόρφωση των τεχνοεπιστημονικών διαδικασιών παραγωγής πόρων, παράλληλα θα αντιλαμβανόταν την ύλη ως δυναμική συνιστώσα στην διαμόρφωση της πραγματικότητας. Σε αυτήν την κατεύθυνση σημαντική είναι η συνεισφορά θεωριών των Μελετών Εργαστηρίου και της Θεωρίας Δρώντων Δικτύου που εκφράστηκαν από τους Λατούρ και του Γούλγκαρ.

A.5.3. Η γενικότερη συμβολή της θεωρίας της Κοινωνικής κατασκευής για τις θεωρίες της τεχνοεπιστήμης

Η θεωρία της κοινωνικής κατασκευής εφοδίασε σε σημαντικό βαθμό το πεδίο μελέτης Επιστήμης, Τεχνολογίας και Κοινωνίας μέσω τριών αναπόσπαστων υπομνήσεων σχετικά με την τεχνολογία και τη επιστήμη. Αυτές αφορούν την κοινωνικότητα της επιστήμης και της τεχνολογίας, τον δραστήριο χαρακτήρα τους και τέλος τον μη φυσικό χαρακτήρα τους. Ενώ λοιπόν ο Κοινωνικός Κονστρουκτιβισμός τροφοδοτεί με αυτές τις ιδέες τις σπουδές Επιστήμης και Τεχνολογίας, ταυτόχρονα, ο σχετικιστικός χαρακτήρας της κοινωνικής Κατασκευασιοκρατίας αναδύεται στα πλαίσια της ανοικτής ερμηνείας των παραπάνω αυτών υπομνημάτων. «Παρότι αυτές οι υπομνήσεις έχουν σημαντική ισχύ, δεν επιδέχονται μια μόνο ερμηνεία» (Sismondo, 2016, σ.100).

Το ζήτημα της ερμηνείας ως μέσο θέσπισης και κατανόησης της πραγματικότητας αποτέλεσε έντονο θέμα πραγματείας του Στίβεν Γούλγκαρ στο έργο που δημοσίευσε μαζί με τον Μπρούνο Λατούρ, «Laboratory Life: The Construction of Scientific Facts» (1986). Η λογική του Κονστρουκτιβισμού φαίνεται να έχει επηρεάσει σημαντικά την θεώρηση των δύο αυτών θεωρητικών, χωρίς ωστόσο να ταυτίζονται με αυτή (Latour, 2005; Γεωργοπούλου,2010,σσ.173-174), αφού μελετάνε τον χαρακτήρα της επιστημονικής ανακάλυψης στα πλαίσια της κοινωνικής κατασκευής της επιστημονικής γνώσης.

Ο Γούλγκαρ πραγματευόμενος ζητήματα της επιστήμης υποστηρίζει ότι τα επιστημονικά γεγονότα κατασκευάζονται μέσω μιας διαδικασίας διαπραγμάτευσης μεταξύ των επιστημόνων, κατά την οποία τα γεγονότα επιλέγονται, απορρίπτονται και τροποποιούνται με βάση τα κριτήρια και τα συμφέροντα της επιστημονικής κοινότητας. Τονίζει επίσης τη σημασία της κατανόησης του τρόπου με τον οποίο τα επιστημονικά γεγονότα διαδίδονται και γίνονται αποδεκτά από την κοινωνία και πώς αυτό μπορεί να διαμορφώσει την κατανόηση της επιστήμης από τον κοινωνικό κόσμο. Ωστόσο, για την καλύτερη κατανόηση του τεχνοεπιστημονικού φαινομένου καθίσταται αναγκαία η εξέταση των πρώτων προσεγγίσεων της κοινωνικής κατασκευής της επιστήμης. Διανοούμενοι όπως ο Μέρτον, ο Κούν, ο Μπλερ αποτέλεσαν πυλώνες της εισαγωγής της κονστρουκτιβιστής σκέψης στο πεδίο της επιστήμης, επηρεάζοντας μετέπειτα με τις θέσεις τους την σκέψη περί κοινωνικής διαμόρφωσης της τεχνολογίας.

A.6. «Μέρτον και η επιστημονική ηθική»

Στο έργο του «Sociology of Science: Theoretical and Empirical Investigations» ο Μέρτον πραγματεύεται την επιστήμη ως έναν Θεσμό ο οποίος σαν θεσμικό στόχο έχει την «επέκταση της πιστοποιημένης γνώσης» (Merton, 1973, σ.270). Η επιστήμη ως μια θεσμική οντότητα βρίσκεται σε άμεση εξάρτηση με την διάρθρωση της κοινωνικής δομής. Συνεπώς, η επιστημονική δράση πραγματοποιείται σε συνάρτηση με την κοινωνική δράση, τα πρότυπα και τις συμπεριφορές που την διαμορφώνουν. Έτσι, για την κατανόηση της επιστήμης ο Μέρτον δίνει μεγάλη έμφαση στους κανόνες συμπεριφοράς που διαμορφώνουν μέσω της κοινωνικής τους εφαρμογής, την επιστήμη ως θεσμό.

Ο Μέρτον θεωρεί ότι συνήθως ο όρος επιστήμη χρησιμοποιείται προκειμένου να περιγράψει ένα σύνολο χαρακτηριστικών μεθόδων πιστοποίησης της γνώσης. Το αποτέλεσμα αυτής της μεθοδολογικής επεξεργασίας της γνώσης που παράγει επιπλέον γνώση, ένα σύνολο από πολιτισμικές αξίες και ήθη τα οποία εντοπίζονται εντός της κοινωνίας και συγκροτούν την επιστημονική ταυτότητα του θεσμού, καθώς και οποιοδήποτε από τους παραπάνω συνδυασμούς. Στην πραγματεία του για την επιστήμη, δεν αγνοείται η αμιγώς τεχνική πτυχή της επιστήμης. Οι κανόνες για τους οποίους, ο επιστημονικός κόσμος διατηρεί την αντικειμενικότητά και αμεροληψία του είναι πέρα από ηθικοί και γνωσιακοί. Όμως, κατά κύριο λόγο, επικεντρώνεται στην κοινωνική πτυχή της Επιστήμης και στον κοινωνικό της χαρακτήρα, αυτόν της κοινωνικής κατασκευής.

Τα μοτίβα επιστημονικής συμπεριφοράς, οι «επιστημονικές νόρμες», αποτελούν αναγνωρισμένες, ενδοθεσμικές πρακτικές που συμβάλλουν στην αναπαραγωγή του θεσμού. Ειδικότερα, πρόκειται για θεσμικές επιταγές (Sismondo, 2016, σσ.47-53.) οι οποίες δεν είναι απαραίτητα κωδικοποιημένες με την μορφή ενός αυστηρού νομικού πλαισίου που εφαρμόζεται στο ακέραιό του, αλλά είναι αυτές που επικυρώνουν τον χαρακτήρα και την ταυτότητα του θεσμού. Παράλληλα, η εφαρμογή ή μη των κανόνων αυτών, συνδυάζεται με ανταμοιβές όπως και κυρώσεις σε περίπτωση που δεν ακολουθούνται. Σύμφωνα με τον Μέρτον, αυτά τα άτυπα «mores» της επιστήμης συγκροτούνται μέσω τεσσάρων στοιχείων: της καθολικότητας, της ακεραιότητας, της κοινοκτημοσύνης και της αμεροληψίας.

Η καθολικότητα (Universalism) σχετίζεται άμεσα με την αξιολόγηση μιας πληροφορίας από τους επιστημονικούς δρώντες, τα άτομα δηλαδή που στελεχώνουν την επιστημονική κοινότητα. Τα υποκείμενα που συμμετέχουν στην διαδικασία αξιολόγησης της γνώσης δεν πρέπει να αφήνουν χαρακτηριστικά όπως η εθνικότητα,

η ταυτότητα, η καταγωγή τους, η κοινωνική τους τάξη, ο δικός τους αξιακός κώδικας να επηρεάζουν την κριτική τους σχετικά με το επιστημονικά ορθό. Η οπτική αυτή διαμορφώνει μια εικόνα όχι απλώς μιας ηθικής στάσης, αλλά ενός ηθικού νόμου, ενός ήθους που πρέπει να εφαρμόζεται καθολικά και οικουμενικά από όλους τους επιστήμονες σε οποιαδήποτε επιστημονική διεργασία. Η διαδικασία αυτή, πέρα από την συγκρότηση ενός αυστηρού μεθοδολογικού πλαισίου αξιολόγησης, προάγεται από τον ίδιο τον θεσμό με άμεσο σκοπό την επιλογή του καταλληλότερου επιστημονικά υποκειμένου για τα ζητήματα που επιστήμη εξετάζει κάθε φορά. Σε αυτή την περίπτωση, «σκοπιμότητα και η ηθική συμπίπτουν» (Merton, 1973, σ.273).

Η κοινοκτημοσύνη (Communism) αφορά τον κοινωνικό χαρακτήρα της γνώσης. Εφόσον η επιστήμη συνιστά μια κοινωνική κατασκευή, έναν θεσμό που εξυπηρετεί την κοινωνία, έτσι και η γνώση, το παράγωγο του θεσμού αυτού οφείλει να προσφέρεται ελεύθερα σε όλα τα κοινωνικά μέλη. Ωστόσο, η διαδικασία αυτή δεν πραγματοποιείται ανεξέλεγκτα. Η αναγνώριση της επιστημονικής γνώσης προϋποθέτει την τήρηση κανόνων οι οποίοι επιτάσσουν την αναγνώριση του δημιουργού της εκάστοτε επιστημονικής γνώσης. Πρόκειται για μια από τις ασφαλιστικές δικλίδες διατήρησης του συλλογικού χαρακτήρα της επιστήμης και την ανάπτυξη ενός πεδίου, μιας ζώνης ελεύθερης διάχυσης και διακίνησης της γνώσης κάτω από συνθήκες συλλογικής αναγνώρισης.

Οποιαδήποτε απόκρυψη της επιστημονικής γνώσης συνάδει με μια αντικοινωνική συμπεριφορά. Ακόμα και αν ο δημιουργός της, έχει συμβάλει στην συγκρότηση μιας μεγάλης επιστημονικής ανακάλυψης, εν τέλει, η πράξη αυτή δεν θα έχει κανένα νόημα αφενός γιατί δεν θα αναδειχθεί ποτέ στην κοινωνία, αφετέρου διότι δεν θα βοηθήσει την επιστήμη να αναπτυχθεί, χρησιμοποιώντας την ως γόνιμο έδαφος για επιστημονική επεξεργασία και αξιολόγηση. Η ύπαρξη ενός κοινού ιδιοκτησιακού ήθους της γνώσης αποτελεί απαραίτητη συνθήκη για την αναπαραγωγή του κύκλου διάδοσης και ανάπτυξης του επιστημονικού αγαθού.

Εφόσον λοιπόν διακινούνται και παράγονται πληροφορίες, ο έλεγχος αυτών, δηλαδή η επιστημονική κριτική που ασκείται στα επιστημονικά δεδομένα, για τον Μέρτον πρέπει να πραγματοποιείται κάτω από συνθήκες αμεροληψίας (Disinterestedness). Οι επιστήμονες πρέπει να λειτουργούν στην επιστήμη αμερόληπτα, δίχως να δέχονται επιρροές από εξωτερικούς παράγοντες ή από τα προσωπικά τους συμφέροντα. Τα αποτελέσματα των διεργασιών τους πρέπει να

δημοσιεύονται και να αξιολογούνται από τα υπόλοιπα μέλη της επιστημονικής κοινότητας. Με αυτόν τον τρόπο αποκλείεται από τον επιστημονικό θεσμό η διαφθορά. Η αποτροπή της απάτης συνιστά «έναν από τους κανόνες που πιθανώς εφαρμόζεται σε πολύ ικανοποιητικό βαθμό στην επιστήμη αφού τα ποσοστά εμφάνισής της στον θεσμό της επιστήμης σε σχέση με άλλους θεσμούς είναι ιδιαίτερα μικρά» (Merton, 1976, σ.276). Ένας από τους παράγοντες διαμόρφωσης της χαμηλής συχνότητας εμφάνισης της απάτης από την επιστήμη είναι το ίδιο της το αντικείμενο. Η ίδια η διάρθρωση του θεσμού, η αυστηρή αξιολόγηση και τα υψηλά κριτήρια αποδοχής των επιστημονικών θέσεων από τα ίδια τα μέλη της επιστημονικής κοινότητας, καθιστούν ιδιαίτερα δύσκολη την εξαπάτηση του θεσμού.

Πέρα από την αμεροληψία των αποτελεσμάτων, η επιστήμη πρέπει να αντιμετωπίζεται με αμφισβήτηση. Η διαφορά της επιστημονικής κοινότητας σε σχέση με άλλες κοινωνικές ομάδες έγκειται στην σκεπτικιστική της στάση, έναν «οργανωμένο σκεπτικισμό (Organised Skepticism)», απέναντι στον νέο ισχυρισμό. Μάλιστα, η έντονη αμφισβητική διάθεση της επιστήμης πολλές φορές συνιστά σημείο προστριβών και συγκρούσεων σχετικά με ζητήματα που ερμηνεύονται με διαφορετικό τρόπο από άλλους θεσμούς όσον αφορά όμως τα ίδια τα δεδομένα. Ο επιστήμονας πολλές φορές πρέπει να λαμβάνει μια επιστημονικά ηθική, αλλά για άλλους κοινωνικούς θεσμούς μια μη απαραίτητα ηθική, στάση ουδετερότητας, κριτικής και αμφισβήτησης, ανάμεσα στο ιερό και το βέβηλο του υπόλοιπου κοινωνικού κόσμου, ώστε να βρίσκεται σε θέση να διατηρεί την αντικειμενικότητά της.

Οι θεσμικές επιταγές ηθικής διαμορφώνουν κάτι παραπάνω από έναν απλό κώδικα δεοντολογίας. Κατασκευάζουν συμπεριφοριστικές νόρμες, νοοτροπίες οι οποίες αναπαράγονται διαρκώς εντός του θεσμού διαμορφώνοντας την ίδια στιγμή τον ίδιο τον θεσμό. Η διαδικασία αναπαραγωγής ωστόσο δεν υφίσταται δίχως την παροχή κινήτρων. Για την υιοθέτηση, πόσο μάλλον για την διατήρηση μιας συμπεριφοράς, η ύπαρξη κινήτρων μέσω ανταμοιβών αλλά και η ύπαρξη κυρώσεων συνιστούν ισχυρό μηχανισμό δράσης εντός των θεσμών. Οι ανταμοιβές και οι κυρώσεις στο επιστημονικό πεδίο συνδέονται με το στοιχείο της αναγνώρισης. Η αναγνώριση ενός ατόμου αφενός επικροτεί το έργο του, αφετέρου επιβραβεύει το ήθος του επιστήμονα για την προσφορά του στην κοινωνία. Επιπλέον, παρουσιάζονται διαβαθμίσεις όσον αφορά τις επιβραβεύσεις. Η πιο βασική μορφή ανταμοιβής, λόγω συχνότητας εμφάνισής της, είναι η παράθεση του έργου ενός επιστήμονα στα έργα ενός άλλου ενώ, μέγιστη μορφή επιβράβευσης συνιστά η ονοματοθεσία ενός γνωστικού αντικειμένου ή επιστημονικού

κλάδου με το όνομα του δημιουργού του. Η τελευταία μάλιστα, συνιστά την τιμητική μορφή αναγνώρισης ενός υποκειμένου στον επιστημονικό κλάδο.

Ο Μέρτον, όπως και ο Πόπερ, «υποστηρίζουν ένα ατομικιστικό, δημοκρατικό ιδεώδες της επιστήμης» (Sismondo, 2016, σ.24), αφού και οι δύο αντιλαμβάνονται την επιστημονική ανάπτυξη σαν μια κοινωνική, συλλογική διαδικασία η οποία θα πρέπει να διέπεται από κάποιες τεχνικές απαιτήσεις και ιδεώδη, που να εφαρμόζονται παράλληλα από τα υποκείμενα τα οποία στελεχώνουν αυτές τις κοινότητες. Παρά λοιπόν τα θετικιστικά στοιχεία που χαρακτηρίζουν την επιστημονική δομή και την επιστήμη ως «ένα σύνολο συστηματικών και εμπειρικά επαληθεύσιμων εννοιών, προτάσεων, υποθέσεων και νόμων» (Merton, 1973, σ.2) , η Μερτόνια θέση για την επιστήμη αντιλαμβάνεται την επιστήμη ως ένα κοινωνικό κατασκεύασμα το οποίο αναπαράγεται και αναπτύσσεται εντός των συμβάσεων που πραγματοποιούνται εντός αυτού.

στη συνέχεια

A.7. Η Κουινιανή Επανάσταση στον χώρο της επιστήμης

Κάτω από την πλέον ισχυρή προοπτική του Κοινωνικού Κονστρουκτιβισμού, οι θεωρήσεις σχετικά με την επιστημονική γνώση αρχίζουν να μεταβάλλονται. Το στοιχείο το οποίο αρχίζει να δέχεται έντονη κριτική από πολλούς επιστήμονες της εποχής είναι η διαδικασία εξέλιξης της επιστήμης. Σε αυτά τα πλαίσια, καθοριστική σημασία για την εικόνα της επιστήμης, η οποία πλέον γίνεται αντιληπτή κάτω από ένα κατασκευασιοκρατικό πλαίσιο, έχει το έργο του Τόμας Κουν (Thomas Khun), «The Structure of Scientific Revolutions». Στο δοκίμιο αυτό ο Κουν, μελετώντας ιστορικά την πορεία της επιστήμης ασκεί κριτική στην νομοτελειακή εξέλιξή της, θεωρώντας πως οι επιστήμες δεν αναπτύσσονται απαραίτητα σταθερά με άξονα την πρόοδο προς την αλήθεια. Αντιθέτως, επικεντρώνοντας το βλέμμα του στις ομάδες οι οποίες ασχολούνται με την επιστήμη, δηλαδή τους επιστήμονες και τις επιστημονικές κοινότητες ,αντιλαμβάνεται ότι η επιστήμη υπακούει περισσότερο σε Παραδείγματα παρά σε μια γραμμική, «φυσική» εξελικτική νομοτέλεια.

Ο Κουν εισήγαγε την έννοια των «παραδειγμάτων» για να περιγράψει το κοινό σύνολο θεωριών, μεθοδολογιών και παραδοχών που διέπουν την πρακτική μιας επιστημονικής κοινότητας σε μια δεδομένη χρονική στιγμή. Ένα παράδειγμα, όπως υποστήριξε ο Κουν, παρέχει ένα πλαίσιο εντός του οποίου εργάζονται οι επιστήμονες,

περιγράφοντας τα προβλήματα που πρέπει να επιλυθούν, τις μεθόδους που πρέπει να χρησιμοποιηθούν και τα κριτήρια για την αξιολόγηση των αποτελεσμάτων (Khun, 1962). Η κοινωνική πτυχή των παραδειγμάτων έγκειται στο ρόλο τους ως κοινές συμβάσεις που επιτρέπουν την επιστημονική επικοινωνία και συνεργασία, καθώς και στο γεγονός ότι διαμορφώνονται και διατηρούνται από τις συλλογικές δράσεις των επιστημόνων.

Ωστόσο, ακόμα και αν παρουσιάζεται έντονο το κοινωνικό στοιχείο εντός των παραδειγμάτων, πρέπει να διευκρινιστεί ότι οι κοινές συμβάσεις που παρουσιάστηκαν προηγουμένως, δεν περιορίζονται μονάχα σε συμβάσεις εντός μιας συγκεκριμένης ηθικής τάξης πραγμάτων, όπως αυτή που περιέγραψε ο Μέρτον. Τα Κουινιανά Παραδείγματα δεν διαμορφώνουν μονάχα την συμπεριφορά των επιστημόνων όσον αφορά τις διαπροσωπικές τους σχέσεις, αλλά επηρεάζουν κατά κύριο λόγο τις επιστημονικές πρακτικές, μεθοδολογίες και ερμηνείες, τονίζοντας με αυτόν τον τρόπο τον έντονο εννοιολογικό τους χαρακτήρα καθώς και την πρακτική μορφή που έχουν. Με αυτόν τον τρόπο, «οι επιστημονικές κοινότητες οργανώνονται κυρίως γύρω από ιδέες και πρακτικές και όχι γύρω από ιδεώδη συμπεριφοράς» (Sismondo, 2016, σ. 45).

Εφόσον λοιπόν η επιστήμη δεν αναπτύσσεται γραμμικά προς μια μονάχα κατεύθυνση, παρουσιάζει και αστάθειες. Οι αστάθειες αυτές αποτελούν αναπόσπαστο κομμάτι της λογικής που διέπει την επιστήμη για τον Κουν, δηλαδή την λογική της επίλυσης πρακτικών προβλημάτων, καθώς η παρουσία ανωμαλιών αποτελεί χαρακτηριστικό ζωτικής σημασίας για την επιστήμη. Δεδομένου ότι τα παραδείγματα παρουσιάζουν άμεση σύνδεση με τις πρακτικές επίλυσης των επιστημονικών προβλημάτων, στις περιπτώσεις που οι πρακτικές αυτές δεν επαρκούν, δημιουργούνται αστάθειες εντός της κοινότητας που χαρακτηρίζονται από τον Κουν ως «περίοδοι κρίσης» (1962).

Όταν ένα Παράδειγμα βρίσκεται σε μια περίοδο αδυναμίας επίλυσης των επιστημονικών προβλημάτων, συγκροτείται ένα κλίμα επιστημονικής επανάστασης. Οι περίοδοι των επαναστάσεων συνιστούν περιόδους έντονης αμφισβήτησης του κυρίαρχου επιστημονικού παραδείγματος. Η μέχρι τότε γνωστή επιστημονική πραγματικότητα, οι μέθοδοι επίλυσης των επιστημονικών γρίφων καθώς και η ευρύτερη στάση των επιστημόνων μεταβάλλεται. Τα άτομα τα οποία συμβάλλουν κατά κύριο λόγο σε αυτή τη μορφή εξέλιξης είναι οι νέοι επιστήμονες. Κάτι τέτοιο συμβαίνει

καθώς οι νέοι επιστήμονες δεν έχουν την ίδια διάδραση με το παλαιότερο κυρίαρχο παράδειγμα. Η συνθήκη αυτή, σε συνδυασμό με την περίοδο της επιστημονικής κρίσης, συγκροτούν ένα κλίμα όπου οι νέοι επιστήμονες πρέπει να υιοθετήσουν νέες μεθόδους, νέα εννοιολογικά πρότυπα, νέες παραδοχές στον επιστημονικό τους τομέα.

Όταν οι αλλαγές αυτές πραγματοποιηθούν, τότε πραγματοποιείται ουσιαστικά μια «Αλλαγή Παραδείγματος» (Paradigm Shift). Οι αλλαγές παραδείγματος εισάγουν μια νέα εποχή «κανονικής επιστήμης», όπου οι επιστήμονες εργάζονται για την επίλυση προβλημάτων στο πλαίσιο του νέου παραδείγματος και σημειώνουν σταδιακή πρόοδο. Η φύση λοιπόν των Παραδειγμάτων, κατ' επέκταση και της επιστήμης, δεν είναι στάσιμη. Η επιστημονική εξέλιξη δεν διανύει μια ευθεία προς την εξέλιξή της, αλλά μια διαρκή εσωτερική αλλαγή από ένα παλιό παράδειγμα σε ένα νέο, όπου «Το νέο παράδειγμα θα πραγματοποιήσει ξανά τον ίδιο κύκλο προκλήσεων και αποκρίσεων έως ότου το ίδιο θα αντικατασταθεί από κάποιο άλλο. Σε αυτά τα πλαίσια, θεωρεί ο Κουν, η επιστήμη εξελίσσεται και βελτιώνεται» (Weidhorn, M. 2001, σ.8). Συνεπώς, για τον Κουν η επιστημονική ανάπτυξη δεν είναι γραμμική καθώς δεν «δεν συσσωρεύει με ευθύ και άμεσο τρόπο γνώση, αλλά αντ' αυτού κινείται από ένα λιγότερο ή περισσότερο επαρκές παράδειγμα σε ένα άλλο» (Sismondo, 2016, σ.37).

A.7.1. Ασυμμετρία και κριτική του Κουνιανού έργου

Κάθε επιστημονικό παράδειγμα διαμορφώνει μια μορφή επιστημονικής πραγματικότητας η οποία, όπως προαναφέρθηκε, δεν διαπερνά την επιστημονική κοινότητα μονάχα στο πλαίσιο κανόνων ηθικής, αλλά διαμορφώνει στην ουσία τις έννοιες και τις πρακτικές που χρησιμοποιούνται εντός του Παραδείγματος. Κάθε νέο επιστημονικό παράδειγμα λοιπόν, δημιουργεί νέες επιστημονικές πρακτικές οι οποίες υιοθετούνται από τους επιστήμονες αποτελώντας τα εργαλεία τους για την εύρεση της αλήθειας. Για τον Κουν, ακόμα και αν ένας επιστημονικός κλάδος έχει ξεπεράσει την περίοδο της κρίσης που βίωνε, αυτός δεν θα μπορούσε να συγκρίνει τα επιστημονικά του αποτελέσματα με τα αποτελέσματα του ίδιου κλάδου παλαιότερα. Μια τέτοια σύγκριση θα ήταν ασύμμετρη.

Η ασυμμετρία αποτελεί μια σημαντική έννοια όσον αφορά την δομή των επιστημονικών επαναστάσεων. Αναδεικνύει την μοναδικότητα του εκάστοτε

επιστημονικού παραδείγματος καθώς αναγνωρίζει πως οι συνθήκες που διαμορφώνουν τα επιστημονικά αποτελέσματα διαφέρουν σημαντικά από παράδειγμα σε παράδειγμα. Παραδείγματος χάρη, μια σύγκριση των θεωριών που συγκροτήθηκαν σε διαφορετικά παραδείγματα δεν θα είχε σημαντική αξία, αφού η κάθε θεωρία διαμορφώθηκε κάτω από μια διαφορετική πραγματικότητα. Οι θεωρίες αυτές δηλαδή, αναπτύχθηκαν κάτω από διαφορετικούς συλλογισμούς, με έννοιες οι οποίες, ακόμα και αν ήταν κοινές και στα δύο παραδείγματα, μπορεί να είχαν μεγάλη σημασιολογική απόκλιση.

Για την θέση του αυτή, ο Κουν δέχτηκε έντονη κριτική. Η διαφορά παραδειγμάτων, ακόμα και αν προϋποθέτει την διαφορά της επιστημονικής πραγματικότητας, δεν συνεπάγεται απαραίτητα μια εκ διαμέτρου ασύμμετρη σημασιολογία (Sismondo, 2016, σσ.41-44). Οι έννοιες που μπορεί να δομούν μια θεωρία, δεν παρουσιάζουν πάντα τεράστιες διαφορές μεταξύ παραδειγμάτων ώστε να θεωρούνται προϊόντα εξ ολοκλήρου διαφορετικής κατασκευής. Κάτι τέτοιο σημαίνει ότι ακόμα και αν οι θεωρίες διαφέρουν μεταξύ τους, οι έννοιες οι οποίες τις στελεχώνουν, όπως και οι παρατηρήσεις στις οποίες αναφέρονται, δεν είναι απαραίτητα τόσο διαφορετικές σε κάθε παράδειγμα, αφού μπορούν κάλλιστα να παρουσιάζουν μεγάλο αριθμό ομοιοτήτων ή μάλιστα οι ομοιότητες να ξεπερνούν τις διαφορές.

Παρά την κριτική που δέχτηκε η προσέγγιση της ασυμμετρίας, η θεωρία του Κουν σχετικά με τις επιστημονικές επαναστάσεις, η δυναμική των παραδειγμάτων στην διαμόρφωση της επιστημονικής πραγματικότητας, καθώς και οι ασύμμετρες συγκρίσεις μεταξύ διαφορετικών παραδειγμάτων, αναδεικνύουν την ισχυρή προοπτική του κοινωνικού κονστρουκτιβισμού όσον αφορά την κατανόηση της επιστήμης ως ένα κοινωνικό κατασκεύασμα. Η αμφισβήτηση της γραμμικής κατεύθυνσης της επιστήμης προς μια επιστημονική αλήθεια τονίζει την κοινωνικότητα της επιστημονικής διαδικασίας. Ακόμα περισσότερο όμως, θέτει υπό αμφισβήτηση την θετικιστική προσέγγιση της γραμμικής ανάπτυξης της γνώσης αφού μέσω της θέσης του, «έθεσε σε αμφισβήτηση την άποψη του θετικισμού ότι η επιστήμη βασίζεται σε οικουμενικά a priori» (Κάλλας, 2015,σ.66)

A.8. Ισχυρό Πρόγραμμα της Κοινωνιολογίας της Γνώσης και Κοινωνική Κατασκευή της Τεχνολογίας (SCOT)

Το ισχυρό πρόγραμμα της Κοινωνιολογίας της Επιστημονικής Γνώσης συγκροτήθηκε και αναπτύχθηκε κατά κύριο λόγο βασιζόμενο στις αρχές που θέσπισε ο Τόμας Μπλέρ (Tomas Bloor) στην προσπάθειά του να αποσχετικοποιήσει την απόλυτη σχετικότητα του Κοινωνικού Κονστрукτιβισμού. Η πρώτη από αυτές είναι ότι η Κοινωνιολογία της γνώσης οφείλει να είναι αιτιακή. Δηλαδή, για την κατανόηση ενός φαινομένου είναι αναγκαία η προσφυγή στις ίδιες τις συνθήκες και τα μέσα που το παράγουν. Στη συνέχεια η εξήγηση του φαινομένου πρέπει να συγκροτείται κάτω από συνθήκες αμεροληψίας ως προς τις διχοτομίες που εμφανίζονται κατά την διάρκεια μελέτης του. Συγκεκριμένα το ψεύδος και η αλήθεια, η ορθολογικότητα και ανορθολογικότητα, η επιτυχία και η αποτυχία πρέπει να εξετάζονται από κοινού δίχως να δίνεται η έμφαση μονάχα στην μια πλευρά ερμηνείας.

Ο λόγος για τον οποίο ο Bloor τονίζει την ανάγκη της αμερόληπτης εξήγησης οφείλεται στην αρχή της εφαρμογής της συμμετρικής μεθοδολογίας. Από την στιγμή που ένα αποτέλεσμα αποτελεί προϊόν μιας αιτίας και χαρακτηρίζεται ως αληθές η ψευδές, ο χαρακτηρισμός αυτός, αποτελεί ένα προϊόν ερμηνείας της πραγματικότητας. Η ερμηνεία αυτή, αν και αφορά κατά κύριο το φαινόμενο στο σύνολό του, συγκροτείται από την αξιολόγηση της αιτίας που δημιούργησε το φαινόμενο. Επομένως η προσφυγή στους παράγοντες, στα ίδια τα μέσα που δημιουργούν το προς ερμηνεία αποτέλεσμα είναι αυτοί που συμβάλλουν στην εξήγηση ενός φαινομένου και στον χαρακτηρισμό του ως αληθές η ψευδές, ορθολογικό και μη. Οι παράγοντες αυτοί πρέπει να εξετάζονται συμμετρικά καθώς «οι ίδιοι παράγοντες που παίζουν ρόλο κατά την παραγωγή του αληθούς, υπάρχουν και κατά την παραγωγή του ψεύδους» (Sismondo, 2016, σ.87).

Η μεθοδολογική συμμετρία αποτελεί κάτι περισσότερο από μια προσπάθεια κατανόησης ενός φαινομένου. Η εστίαση της προσοχής στους παράγοντες συγκρότησης παρέχει αφενός μια εξηγητική ισχύ στην ερμηνεία του αποτελέσματος αλλά ταυτόχρονα, διευρύνει τους ορίζοντες κατανόησης της τοποθέτησης του φαινομένου στην μια πλευρά ενός ερμηνευτικού διπόλου. Ουσιαστικά πέρα από την κατανόηση του φαινομένου, «φωτίζεται» και το πεδίο ερμηνείας του και από τις δύο πλευρές με ίσες αποστάσεις, συμμετρικά, γεγονός που οφείλεται στην αμερόληπτη εξηγητική θέση του ερευνητή. Όπως τονίζεται από τους Trevor. Pinch και Wiebe E. Bijker «Στο πλαίσιο ενός τέτοιου προγράμματος όλες οι γνώσεις και όλοι οι ισχυρισμοί γνώσης πρέπει να αντιμετωπίζονται ως κοινωνικά δηλαδή, εξηγήσεις για τη γένεση,

την αποδοχή και την απόρριψη των αξιώσεων γνώσης αναζητούνται στον τομέα της κοινωνικής κόσμο παρά στον φυσικό κόσμο» (Bijker και άλλοι, 1989, σ.6)

Η συμμετρική μεθοδολογία που συνιστά μια αρχή του ισχυρού προγράμματος του Bloor αν και εννοιολογικά ταιριάζει με την Κουνιανή συμμετρία, διαφέρει σημαντικά ως προς το περιεχόμενό της. Η αρχή της συμμετρίας του Bloor υποδηλώνει ότι οι κοινωνικοί παράγοντες και όχι ο ορθολογισμός καθοδηγούν την παραγωγή επιστημονικής γνώσης (Bird, 2000), ενώ το πλαίσιο του Κουν αναγνωρίζει την ύπαρξη ενός ορθολογικού πυρήνα μέσα στα επιστημονικά παραδείγματα. Μια δεύτερη διάκριση μεταξύ των δύο προσεγγίσεων είναι ότι εστιάζουν σε διαφορετικές πτυχές της επιστημονικής διαδικασίας. Το μοντέλο του Κουν ασχολείται κυρίως με τις αλλαγές σε μακρο-επίπεδο που συμβαίνουν κατά τη διάρκεια των αλλαγών παραδείγματος, δίνοντας έμφαση στο ρόλο των επιστημονικών κοινοτήτων στη διαμόρφωση της ανάπτυξης της επιστημονικής γνώσης. Εν τω μεταξύ, η αρχή της συμμετρίας του Bloor τονίζει τη σημασία των ενεργειών και των αποφάσεων των μεμονωμένων επιστημόνων που συμβάλλουν στην αποδοχή ή την απόρριψη συγκεκριμένων επιστημονικών ισχυρισμών (Barnes, 1982).

Η τελευταία αρχή του ισχυρού προγράμματος προϋποθέτει ότι η Κοινωνιολογία της επιστημονικής Γνώσης πρέπει να είναι αναστοχαστική. Τα πρότυπα που χρησιμοποιούνται από την Κοινωνιολογία για την κατανόηση της επιστημονικής Γνώσης πρέπει να χρησιμοποιούνται και στην ίδια. Η αρχή αυτή, χρησιμοποιείται από τον Bloor προκειμένου να «απασφαλίσει» την Κοινωνιολογία της Γνώσης από το μονοπώλιο της εξηγητικής ισχύς της, της απόδοσης του τίτλου κοινωνικό κατασκευάσμα σε οποιαδήποτε διαδικασία παραγωγής γνώσης (Bloor 1976). Κάτι τέτοιο, αν και αποτελούσε μια προσπάθεια περιορισμού της σχετικοποίησης της Κοινωνιολογίας της Γνώσης, υπέπεσε σε δομικές αντιφάσεις καθώς η σχετικότητα που διέπει την ανθρώπινη ερμηνεία και γνώση, μεταφέρεται στα ίδια τα παράγωγά της. Έτσι και η Κοινωνιολογία της Γνώσης από την στιγμή που δεν παύει να αποτελεί και εκείνη ένα ανθρώπινο δημιούργημα στερείται αυτομάτως την οποία αντικειμενική αξία έχει ακόμα και αν ο Bloor επιδιώκει να την αναδείξει.

A.8.1. Ερμηνεία: Η κινητήριος δύναμη της Γνώσης

Από την προοπτική του ισχυρού προγράμματος της επιστημονικής γνώσης, όπως προαναφέρθηκε, μέσω της γραμμικής, αμερόληπτης εξέτασης των παραγόντων που

συγκροτούν μια γνωσιακή κατάσταση, αναδύεται η προβληματική της κινητήριας δύναμης, η μάλλον κατευθυντήριας δύναμης της γνώσης εντός της κοινωνίας. Ο λόγος για τον οποίο συγκροτείται η γνώση εξετάζεται συνήθως ανάλογα με το επίπεδο εστίασης της προσέγγισης.

Τα επίπεδα εστίασης αυτά εξετάζονται βάσει δύο ειδών ερμηνείες. Τις ιντερναλιστικές και τις εξτερναλιστικές ερμηνείες. Οι ιντερναλιστικές ερμηνείες κατά κύριο λόγο εστιάζουν στην δυναμική που αναπτύσσεται εντός των κοινωνικών ομάδων. Στην περίπτωση της επιστημονικής γνώσης, η οποία αποτελεί και αντικείμενο εξέτασης της συγκεκριμένης θεματικής ενότητας, οι ιντερναλιστικές ερμηνείες εξετάζουν την γνώση ως αντικείμενο που συγκροτείται μέσα από τις σχέσεις των ατόμων που στελεχώνουν την επιστημονική κοινότητα. Από την άλλη πλευρά, οι εξτερναλιστικές ερμηνείες εστιάζουν στις ευρύτερες κοινωνικές δυνάμεις και ιδεολογίες που επικρατούν γενικότερα στην κοινωνία την εκάστοτε χρονική περίοδο και διαμορφώνουν την συγκρότηση της κοινωνικής γνώσης, συνακόλουθα και της επιστημονικής.

Όπως τονίζει ο Μάνχαϊμ (1997, σ.311), η κοινωνιολογία της γνώσης «θέλει να εκπονήσει μιαν ανταποκρινόμενη στη σημερινή κατάσταση θεωρία γύρω από τη σημασία των εξωθεωρητικών όρων της γνώσης» επιδεικνύοντας την σημασία των εξωτερικών παραγόντων στην διαμόρφωση της γνώσης καθώς και την «κατασκευαστική» ανοικτότητά της. Η σημασία των ευρύτερων κοινωνικών παραγόντων για την διαμόρφωση της επιστήμης είναι μεγάλη και πρέπει να αποτελεί αντικείμενο εξέτασης σχετικά με του τρόπους διαμόρφωσης του επιστημονικού αφού «η γένεση και διαμόρφωση της σκέψης προσδιορίζεται από ποικίλους πραγματικούς (εξωθεωρητικούς) κοινωνικούς παράγοντες, ανάγοντας έτσι την κοινωνική ομάδα καθεαυτήν σε μια οντική βάση και σε ένα προνομιακό ρόλο σε σχέση με το άτομο (Ναγόπουλος, 2015, σ.213)

Αν και πραγματοποιούνται εξετάσεις τόσο σε ιντερναλιστικό όσο και σε εξτερναλιστικό επίπεδο, η Κοινωνιολογία της Γνώσης αλλά και η Κοινωνιολογία γενικότερα συνδέονται πολλές φορές περισσότερο με τις εξτερναλιστικές προσεγγίσεις. Αυτή η ομολογουμένως βεβιασμένη συλλογιστική, οφείλεται στο γόνιμο πεδίο που ανέπτυξε η Μαρξιστική θεωρία σε σχέση με την εξέταση κατασκευών και την αναγνώριση μιας κατευθυντήριας προοπτικής των κατασκευών αυτών εντός της Κοινωνίας.

Σε αυτό το σημείο, είναι σημαντικό να διευκρινιστεί ότι δεν είναι ορθό να πραγματοποιηθεί μια άμεση συσχέτιση της εξτερναλιστικής μορφής εξέτασης της πραγματικότητας με την Μαρξιστική θεωρία ακόμα και αν η δεύτερη συνέβαλε σημαντικά στην ιδέα της κατανομής πόρων , όπως η γνώση στην Κοινωνία. Οι Μαρξιστικές θεωρίες διεύρυναν το πεδίο μελέτης των κοινωνικών φαινομένων , συνεπώς και της εξτερναλιστικής ερμηνείας, παρέχοντας επιπλέον παράγοντες εξέτασης, δείχνοντας την προοπτική και σημασία του κινήτρου και συγκεκριμένα του συμφέροντος στην παραγωγή της γνώσης.

Οι δύο ερμηνείες που παρουσιάστηκαν, παρόλο που έγιναν κατανοητές με γνώμονα τις διαφορές τους, δεν διακρίνονται από το ισχυρό πρόγραμμα της γνώσης. Αξιοποιούνται και οι δύο με την ίδια συμμετρική, αμερόληπτη σκεπτικιστή τάση. Η έννοια του συμφέροντος, όσον αφορά την επιστημονική γνώση, απασχολεί σε ιδιαίτερο βαθμό τόσο τις ιντερναλιστικές προσεγγίσεις όσο και τις εξτερναλιστικές παρά το γεγονός ότι στις δεύτερες συνήθως γίνεται ευκολότερα κατανοητή η ύπαρξη τους. Ο Γούλγκαρ , βασιζόμενος στην λογική του ισχυρού προγράμματος του Bloor, ασκεί κριτική στην έννοια του συμφέροντος υπερτονίζοντας την αδυναμία αιτιακής συσχέτισης ενός υποκαθορισμένου συμφέροντος με μια συγκεκριμένη πράξη, καθώς η συγκεκριμένη πράξη μπορεί να συνδέεται με περαιτέρω συμφέροντα τα οποία να ενδέχεται να μην λαμβάνονται υπόψιν κατά την διάρκεια αξιολόγησης της αιτίας της. Βάσει των προαναφερθέντων, η θέση του Γούλγκαρ «δεν αφορά τόσο την έννοια του συμφέροντος αλλά αποτελεί ένα ευρύτερο σχόλιο για την ερμηνεία» (Sismondo, 2016, σ.94).

A.9. Ισχυρό Πρόγραμμα Κοινωνιολογίας της γνώσης και Επιστήμες της τεχνολογίας : Η συμβολή του ισχυρού προγράμματος στην κονστρουκτιβιστική ανάγνωση της τεχνολογίας

Το ισχυρό πρόγραμμα της κοινωνιολογίας της επιστημονικής γνώσης συνέβαλε σημαντικά στην αλλαγή θεώρησης σχετικά με την επιστημονική γνώση και τον τρόπο που αυτή παράγεται. Η ιδέα σχετικά με το ότι οποιαδήποτε κοινωνική κατασκευή πρέπει να μην αντιμετωπίζεται ως δεδομένη περιόρισε σημαντικά τον σχετικιστικό χαρακτήρα που δίεπε την κοινωνική κατασκευασιοκρατία δημιουργώντας ταυτόχρονα νέες προοπτικές σχετικά με την ανάγνωση άλλων, κοινωνικών κατασκευών όπως αυτή της τεχνολογίας. Η εξέταση της κοινωνικότητας της τεχνολογίας και των τεχνολογικών

συστημάτων είχε απασχολήσει τις κοινωνικές επιστήμες αλλά κατά κύριο λόγο μονάχα σε ιστορικό επίπεδο. Η ανάδειξη της κοινωνικότητας της τεχνολογίας ως απαραίτητη προϋπόθεση για την δημιουργία της οφείλεται σε πολλούς θεωρητικούς μεταξύ άλλων των Tomas Huges, Michel Callon ,Pinch & Bijker (Bijker και άλλοι, 1989; Law (2008); Bijker,2010).

Η ανάγνωση της τεχνολογίας κάτω από όρους κονστρουκτιβισμού ακόμα και αν πλέον είναι εφικτή, δεν θα πρέπει να θεωρείται δεδομένη. Οι δυαδικότητες που αναπτύχθηκαν ήδη από την ανάπτυξη της θετικιστικής φιλοσοφίας, είχαν ως αποτέλεσμα, η τεχνολογία, όχι μόνο να θεωρείται να θεωρείται μια αμιγώς ξεχωριστή δραστηριότητα από την επιστήμη, αλλά και να γίνεται κατανοητή ως εφαρμοσμένη πτυχή της επιστήμης, δίνοντάς της με αυτόν τον τρόπο δευτερεύοντα χαρακτήρα. Αν και ο Κονστρουκτιβισμός είχε κατορθώσει να διασπάσει την ουσιοκρατική, δύιστική προοπτική διαχωρισμού κοινωνίας και φύσης , δεν κατόρθωσε να πραγματοποιήσει το ίδιο όσον αφορά την τεχνολογία. Διαχωρισμοί όπως «τεχνολογία-επιστήμη», «καθαρή/εφαρμοσμένη», «εσωτερικοί-εξωτερικοί», «τεχνικό-κοινωνικό» εμφανίζονται στον χώρο των Κοινωνικών επιστημών ακόμα και μετά την ανάπτυξη και εδραίωση των κονστρουκτιβιστικών θεωριών.

Η διατύπωση του Ισχυρού Προγράμματος από τον Μπλέρ, τροφοδότησε την μελέτη των τεχνολογικών φαινομένων με μια νέα προοπτική. Σύμφωνα με τον Λέιτον (Layton, 1977 όπως αναφέρεται στους Bijker και άλλοι, 1989) « Οι διαχωρισμοί μεταξύ επιστήμης και τεχνολογίας δεν είναι μεταξύ των αφηρημένων λειτουργιών της γνώσης και της πράξης. Περισσότερο είναι κοινωνικές». Η εξέταση της τεχνολογίας πλέον θα μπορεί να πραγματοποιείται κάτω από όρους παρόμοιους με αυτούς της επιστημονικής γνώσης. Δηλαδή , η εξέταση της τεχνολογίας να πραγματοποιείται ως μια δυναμική διαδικασία ανάγνωσης, χωρίς τίποτα να λαμβάνεται ως δεδομένο, όπως ακριβώς και ένα μαύρο κουτί (Bijker και άλλοι, 1989).

Υπό την προοπτική αυτή, όταν διερευνάται οποιαδήποτε τεχνολογική κατασκευή δεν τίθεται θέμα επιτυχίας και αποτυχίας. Το βλέμμα του κοινωνικού ερευνητή που ασχολείται με ένα τεχνολογικό κατασκεύασμα δεν πρέπει να στρέφεται μόνο σε αυτό το οποίο είναι αναγνωρισμένο ή οικονομικά επιτυχημένο. Η αποτυχία ή επιτυχία ενός τεχνουργήματος πρέπει να εξετάζεται συμμετρικά, και από τις δύο πλευρές ανάλυσης καθώς «για μια κοινωνιολογική θεωρία της τεχνολογίας θα πρέπει να είναι το explanandum, όχι το explanans» (Bijker και άλλοι, 1989, σ.24). Έτσι, για τον Bijker (2010, όπως αναφέρεται στους Baron & Gomez, 2016 , σ.134) «οι

κονστρουκτιβιστικές μελέτες της τεχνολογίας δεν επικεντρώνονται στο ερώτημα «Τι είναι η τεχνολογία;» Αντ' αυτού ανιχνεύουν τη διαδικασία κατασκευής της τεχνολογίας. Οι μελέτες αυτές επεσήμαναν κυρίως τον τρόπο με τον οποίο η τεχνολογία κατασκευάζεται και χρησιμοποιείται, και όχι αυτό που ουσιαστικά είναι».

Στις θεωρίες αυτές αν και η τεχνολογία κατασκευάζεται κοινωνικά, η ύλη από την οποία δημιουργείται είναι πειθήνια στον άνθρωπο ενώ για την ολοκλήρωση οποιαδήποτε τεχνολογικής κατασκευής η ύλη πρέπει να λάβει μορφή ορισμένη από την επιστήμη. Για τον κοινωνικό κονστρουκτιβισμό, η τεχνολογία συνεχίζει να θεωρείται μια μορφή εφαρμοσμένης επιστήμης με ειδοποιό διαφορά από τον θετικισμό και τις νεωτερικές επιστήμες που αυτός επηρέασε, ότι και η επιστημονική γνώση θεωρείται κοινωνική κατασκευή. Μάλιστα, πολλοί επιστήμονες από τον χώρο της Επιστήμης Τεχνολογίας και Κοινωνίας «μεταφέρθηκαν από το να μιλούν για τον κοινωνικό σχηματισμό της επιστημονικής γνώσης σε αυτόν (τον χώρο) που έγινε γνωστός ως Κοινωνική Κατασκευή της Τεχνολογίας (SCOT)» (Law, 2008).

Το ισχυρό πρόγραμμα της επιστημονικής γνώσης, όπως διαμορφώθηκε από τον Τόμας Μπλέρ, αναδεικνύει την συμμετρική εξέταση όλων των γνωσιακών κατασκευών. Η κονστρουκτιβιστική πεποίθηση κάτω από τους όρους του ισχυρού προγράμματος, της αναγνώρισης δηλαδή οποιαδήποτε κατασκευής σε ένα κοινωνικό γνωσιακό κατασκευάσματα παροτρύνει ουσιαστικά τη εξέταση της τεχνολογίας ως μια ακόμα κοινωνική κατασκευή η οποία χρίζει συμμετρικής εξέτασης. Η αντιμετώπιση της τεχνολογίας λοιπόν από το ισχυρό πρόγραμμα της τεχνολογίας με τους ίδιους όρους που αντιμετωπίζεται και η επιστημονική γνώση, επισημαίνει ότι ο διαχωρισμός επιστήμης και τεχνολογίας ως διαφορετικά παιδιά μελέτης μπορεί να πραγματοποιηθεί κάτω από τους ίδιους όρους, αυτούς ενός αντικείμενου, όπου τίποτε δεν πρέπει να λαμβάνεται ως δεδομένο καθώς συνιστά μια κοινωνική κατασκευή.

A.10 Μπρούνο Λατούρ και Στιβ Γούλγκαρ: Η θεωρία δρώντων δικτύου και οι μελέτες εργαστηρίου.

A.10.1 Οι φιλοσοφικές καταβολές της θεωρίας δρώντων δικτύου: Βιταλισμός και Νεοβιταλισμός

Ο βιταλισμός ή αλλιώς ζωτικοκρατία συνιστά ένα φιλοσοφικό ρεύμα που αναπτύχθηκε τον 18ο και 19ο αιώνα ως μια μορφή εναντίωσης απέναντι στην καρτεσιανή φιλοσοφία

της νεωτερικής επιστήμης. Απέναντι στο δυϊστικό πρότυπο της καρτεσιανής φιλοσοφικής σκέψης που αφορά, από την μια πλευρά τους ζωντανούς οργανισμούς και από την άλλη την ύλη η οποία δεν χαρακτηρίζεται από την ζωή, ο βιταλισμός αναπτύσσει μια εντελώς διαφορετική προοπτική όσον αφορά το φαινόμενο ζωή. Αναζητά [όπως και πολλές άλλες φιλοσοφικές σκέψεις και στοχασμοί] το τί είναι ζωή. Στο φιλοσοφικό, υπαρξιακό ερώτημα που θέτει, προκειμένου να προσδιορίσει το τι είναι ζωή, αναζητά την αζωία, την ιδιότητα αυτή που καθιστά ένα αντικείμενο άψυχο. Αυτή η ιδιότητα λοιπόν που καθιστά έναν οργανισμό ζωντανό, σύμφωνα με τον βιταλισμό, ονομάζεται ζωτική δύναμη «vis vitalis» ή αλλιώς «ζωτική ορμή», «elan vital». Αυτή, συνιστά την ουσιαστική δύναμη, η οποία ξεχωρίζει την ζωντανή από την μη ζώσα ύλη. Πρόκειται για μια οπτική που όπως διακρίνεται και από τις προαναφερθείσες ορολογίες που αναπτύχθηκαν σχετικά με αυτήν, δηλώνει την κατανόηση της ύλης ως δυναμική και κινητήρια ιδιότητα η οποία δίνει ζωή στους οργανισμούς σε αντίθεση με την υπόσταση των άψυχων σταθερών αντικειμένων. Αυτή η ιδιότητα συνιστά μια εσωτερική δυναμική, μια τάση της ζώσας ύλης στο να εξελίσσεται στο πέρασ του χρόνου, «μια μορφή αιτιότητας που δεν είναι εξωτερική αλλά εσωτερική» (Γεωργοπούλου, 2010, σ.161, Κουρούτζας, 2018, σσ.61-79). Στην πραγματικότητα ο βιταλισμός θεωρεί ότι οι οργανισμοί που είναι ζωντανοί παρουσιάζουν μια αυτοδημιούργητη τάση να εξελίσσονται, ενώ την ίδια στιγμή η ικανότητά τους να εξελίσσονται διαφυλάττει την ζωτική τους υπόσταση.

Σύμφωνα με τα παραπάνω, γίνεται κατανοητό ότι ο βιταλισμός ασκεί σημαντική επιρροή στον χώρο όχι μόνο της φιλοσοφικής σκέψης αλλά και στο πεδίο των κοινωνικών επιστημών. Διατηρώντας τους δυισμούς της ουσιοκρατίας σχετικά με την ζώσα και μη, την οργανική και ανόργανη, τον έμβιο με τον άβιο κόσμο, τον άνθρωπο και την μηχανή τοποθετεί κάτω από υλιστικούς όρους την προοπτική της υλικής δυναμικότητας και της εξελικτικής προδιάθεσής της στον χρόνο. Στο πλαίσιο αυτό, όπως και ο κονστρουκτιβισμός, διατηρεί τον άνθρωπο στην πρωτοκαθεδρία της ζωντανής προοπτικής. (Γεωργοπούλου, 2010).

Ωστόσο, η αδυναμία της βιταλιστικής φιλοσοφίας να καταφέρει αιτιωδώς να αποδείξει τις θέσεις τις οποίες πραγματευόταν, περιόρισε σημαντικά την ισχύ της στις θεωρητικές συζητήσεις στις αρχές του 20ου αιώνα. Η αναγωγή της ζωής, ως μια αυτόνομη δυναμική ιδιότητα, η οποία ενυπάρχει στον άνθρωπο, αποτελώντας την ίδια στιγμή προϋπόθεση οντολογικού διαχωρισμού ανάμεσα σε ζωντανούς και μη οργανισμούς, θεωρήθηκε προβληματική καθώς η βιταλιστική θεωρία δεν παρείχε τις

κατάλληλες αποδείξεις σχετικά με την θέση της για την ζωή. Αυτή η προσέγγιση, την κατέστησε μια θεωρία που διέπεται από μυστικισμό, τοποθετώντας την στον χώρο της μεταφυσικής και περιορίζοντας σημαντικά τις συζητήσεις περί βιταλισμού για τουλάχιστον το μισό του 20ου αιώνα.

A.10.1.1 Απο τον Βιταλισμό στον Νεοβιταλισμό: Το έργο του Deleuze

Ο Βιταλισμός, όπως προαναφέρθηκε, σταμάτησε να εμφανίζεται στις φιλοσοφικές συζητήσεις του 20ου αιώνα με την ίδια συχνότητα που εμφανίστηκε τους προηγούμενους δύο αιώνες. Μολαταύτα, η βιταλιστική θεώρηση των πραγμάτων επανήλθε στο πεδίο των θεωρητικών προβληματισμών, σχετικά με την ζωή, στα τέλη του 20ου αιώνα, χάρη στην καθοριστική συμβολή του έργου του Deleuze (Γεωργοπούλου, 2010, σσ.157-168). Ο Deleuze, βασιζόμενος στις ιδέες του βιταλισμού ανασυνθέτει την εικόνα σχετικά με το τι πραγματικά είναι η ζωή. Η βασική του θέση δεν πηγάζει από την διαφοροποίηση ζωής και αζωίας. Στην πραγμάτευσή του εκμεταλλεύεται την εξελικτική και δυναμική θεώρηση, που γινόταν ορατή στους έμβιους οργανισμούς, και την θέτει ως αξίωμα σχετικά με την ουσία της ζωής. Σε αυτή την περίπτωση, η ζωή δεν προσδιορίζεται μέσω από το δίπολο ζωής και αζωίας. Αντιθέτως, ο Deleuze, αρκείται σε μια μονιστική προοπτική, όπου η ζωή συνάδει με κίνηση και εξέλιξη.

Εστιάζοντας στην οπτική αυτή, γίνεται αισθητή η μεταστροφή των δεισμών της ουσιοκρατίας, σε μια μονιστική προσέγγιση σχετικά με την προέλευση της ουσίας, μάλλον καλύτερα στον επαναπροσδιορισμό προέλευσης της. Οι ζωντανοί οργανισμοί δεν προσδιορίζονται σε αντίθεση με τους μη ζωντανούς οργανισμούς. Η ζωή αναγνωρίζεται σε ό,τι υπάρχει και εξελίσσεται στα πλαίσια του χωροχρόνου. Κάτω από την οπτική, αυτή δεν αξιολογείται τμηματικά ούτε διακριτικά αναλόγως με την οντολογία του κάθε οργανισμού, ζωντανού και μη. Στα πλαίσια του χρόνου, όλοι οργανισμοί βρίσκονται κάτω από μία διαδικασία ενός συνεχούς γίνεσθαι, (continuum), μια ατέρμονη αλλαγή. Αυτή οπτική μετατοπίζει το ενδιαφέρον των θεωρήσεων για την ζωή και προτάσσει την μεταβολή του ενός πάγιου πλαισίου εξέτασης της ύλης όπου «οι λογικές συσχετίσεις ομοιοτήτων και διαφορών πρέπει να ενισχυθούν από τις εμπειρικές συσχετίσεις της γέννησης και της ιστορικής προέλευσης... Ο κόσμος ως σύστημα αντικειμένων μπορεί μονάχα να περιγραφεί^ο Για να εξηγηθεί πρέπει να γίνει

αντιληπτός ως ένα σύστημα διαδοχικών διαδικασιών στο χρόνο». (Macdougall, 1913 σ.106).

Πέρα από την αναγνώριση της ζωής ως μια δυναμική σχεσιακή διαδικασία καθώς και την ιστορική προοπτική που εισαγάγει, η νεοβιταλιστική σκέψη αναδιαμορφώνει την οντολογική οπτική που κληρονόμησε ο βιταλισμός επηρεασμένος από την θετικιστική ουσιοκρατία. Σε αντίθεση με τον διαχωρισμό ζωντανών και μη οργανισμών, μέσω από την αναγνώριση του ποια ακριβώς οντότητα έχει το αποκλειστικό προνόμιο στην ζωή, η νεοβιταλιστική προσέγγιση δεν πραγματοποιεί κάποια σαφή οντολογική διάκριση, πιο συγκεκριμένα, «δεν υπάρχει ουσιώδης οντολογική διαφορά ανάμεσα στα ανθρώπινα και τα μη ανθρώπινα» (Γεωργοπούλου, 2010, σσ.157-168).

Η νέα οντολογία που εισαγάγεται μέσω του νεοβιταλισμού αίρει την ουσιοκρατική προσέγγιση που ξεχώριζε την ύλη σε ζώσα και με ζωντανή καθώς και την στατική εικόνα που ο σχετικισμός είχε σχηματίσει για εκείνη. Η εξέλιξη πλέον δεν λαμβάνεται υπόψιν ως μια τελεολογική διαδικασία. Βασίζεται στην σχετικότητα των οργανώσεων των ενδεδειγμένων σε ένα «μονοδιάστατο επίπεδο» (Γεωργοπούλου, 2010). Στο επίπεδο αυτό δεν επικρατεί το καντιανό σχήμα υποκειμένου-αντικειμένου όπου τα δεύτερα συνιστούν πειθήνια όργανα περαιώσης επιθυμιών, αντιθέτως προτάσσεται μια επιστημολογική και μεθοδολογική θεώρηση όπου οι οντότητες πρέπει να θεωρούνται «αυτόνομες, και (πρέπει) να αντιμετωπίζονται ως ενδεδειγμένες» (Macdougall, 1913 σ.106).

Ο νεοβιταλισμός λοιπόν τοποθετεί ξανά στο προσκήνιο των θεωρητικών, επιστημονικών και φιλοσοφικών συζητήσεων παρέχοντας μια νέα προοπτική για τις κοινωνικές επιστήμες. Η διαφωνία της με τις ουσιοκρατικές δυϊστικές προσεγγίσεις όπως άτομο-αντικείμενο, ζωντανή και μη ύλη, περιορίζεται χάρη στη μονιστική προσέγγισή της ύλης η οποία πλέον κατανοείται μέσω της δυναμικής σχεσιακότητάς της θέτοντας μια ενναλακτική βάση κοινωνιολογικής συζήτησης, ένα νέο, γόνιμο πεδίο ανάπτυξης κοινωνιολογικών θεωριών που αναπτύσσονται σε επιστημονικό και τεχνολογικό επίπεδο, δίχως την μονιστική κοινωνικοκονστρουκτιβιστική προοπτική, όπου όλα ανάγονται σε κοινωνικές κατασκευές. Η νεοβιταλιστική επιστημολογική πρόταση, θα αποτελέσει βάσει των προαναφερθέντων μια υλικοαναστοχαστική κοινωνιολογική συζήτηση στην οποία θα συμμετάσχουν με αρκετά αποφασιστικό τρόπο, σύγχρονοι κοινωνικοί επιστήμονες και κοινωνιολόγοι, οι οποίοι μέσω νέων

θεωριών, όπως αυτή της θεωρίας Δρώντων Δικτύου, «Actor.Network.Theory.» θα προσεγγίσουν τεχνολογιστημονικά ζητήματα με ρηξικέλευθες προσεγγίσεις.

A.10.2. Η ιστορία συγκρότησης της Θεωρίας δρώντων Δικτύου

Η δημιουργία της Θεωρίας Δρώντων Δικτύου (ANT) μπορεί να κατανοηθεί ως απάντηση στους περιορισμούς των παραδοσιακών κοινωνιολογικών προσεγγίσεων που κυριαρχούσαν στις δεκαετίες του 1970 και 1980. Ειδικότερα, η εστίαση στις κοινωνικές δομές και τα συστήματα είχε την τάση να αγνοεί τον ρόλο των μη ανθρώπινων δρώντων στη διαμόρφωση του κοινωνικού κόσμου. Η Θεωρία Δρώντων Δικτύου εν συντομία, «ANT» εμφανίστηκε τη δεκαετία του 1980 ως ένας τρόπος αντιμετώπισης αυτού του περιορισμού και οι ρίζες της εντοπίζονται στο έργο των Michel Callon, Bruno Latour και John Law, οι οποίοι ήταν μέλη του Centre de Sociologie de l'Innovation στην École des Mines στο Παρίσι.

Ένας από τους βασικούς παράγοντες που οδήγησαν στη δημιουργία της ANT ήταν η μεταβαλλόμενη φύση της τεχνολογίας και η σχέση της με την κοινωνία. Στις δεκαετίες του 1970 και του 1980, άρχισε να αναγνωρίζεται όλο και περισσότερο ότι η τεχνολογία δεν ήταν απλώς ένα ουδέτερο εργαλείο, αλλά είχε σημαντικό αντίκτυπο στις κοινωνικές σχέσεις και πρακτικές. Η αναγνώριση αυτή οδήγησε σε ένα αυξανόμενο ενδιαφέρον για τη μελέτη της επιστήμης και της τεχνολογίας, και το Κέντρο Κοινωνιολογίας της Καινοτομίας βρισκόταν στην πρώτη γραμμή αυτής της εξέλιξης.

Η βασική έννοια που προέκυψε από το έργο τους ήταν η ιδέα του δικτύου δρώντων. Το δίκτυο δρώντων συνιστά έναν τρόπο κατανόησης του κοινωνικού κόσμου ως ένα δίκτυο σχέσεων μεταξύ φορέων, οι οποίοι μπορεί να περιλαμβάνουν ανθρώπους, θεσμούς, τεχνολογίες και αντικείμενα. Η συγκεκριμένη προσέγγιση είναι υλιστική. Δεν αντιλαμβάνεται ως δρώντες μονάχα τους ανθρώπους. Οι «actors» δεν μπορούν να κατανοηθούν εκτός δικτύου παρά μόνο εντός αυτού, καθώς αυτού υποστασιοποιούνται μέσω των συνδέσεών τους στο δίκτυο (Sismondo, 2016, σσ. 137-146). Η υποστασιοποίηση των δραστών μέσω της ανάπτυξης δυναμικών σχέσεων εντός της διάταξης του δικτύου που αυτοί συγκροτούν ονομάζεται «σχεσιακή υλικότητα» (Γεωργοπούλου, 2010, σσ.174-178). Οι σχέσεις αυτές δεν είναι σταθερές ή προκαθορισμένες, αλλά αναδιαμορφώνονται συνεχώς καθώς οι φορείς αλληλοεπιδρούν μεταξύ τους. Αυτή η έννοια επέτρεψε στους Callon, Latour και Law

να αναπτύξουν μια πιο διαφοροποιημένη θέση σχετικά κατανόηση του ρόλου της τεχνολογίας στη διαμόρφωση του κοινωνικού κόσμου.

Ένας άλλος παράγοντας που οδήγησε στη δημιουργία της ANT ήταν η επιρροή της φιλοσοφικών θεωριών όπως ο νεοβιταλισμός αλλά και οι κοινωνιολογικές, φιλοσοφικές θεωρήσεις σχετικά με την επιστήμη. Οι Callon, Latour και Law επηρεάστηκαν από το έργο φιλοσόφων όπως ο Thomas Kuhn και ο Ludwig Wittgenstein, οι οποίοι αμφισβήτησαν την παραδοσιακή άποψη της επιστημονικής γνώσης ως αντικειμενικής και ουδέτερης. Αντίθετα, έβλεπαν την επιστημονική γνώση να διαμορφώνεται από το κοινωνικό και ιστορικό πλαίσιο στο οποίο παρήχθη. Η επιρροή αυτή μπορεί να φανεί στον τρόπο με τον οποίο η ANT αμφισβητεί την ιδέα της σαφούς διάκρισης μεταξύ αντικειμενικών γεγονότων και υποκειμενικών ερμηνειών. Αντίθετα, η ANT θεωρεί ότι η γνώση διαμορφώνεται πάντα από το κοινωνικό πλαίσιο στο οποίο παράγεται και αναγνωρίζει το ρόλο των μη ανθρώπινων δρώντων στη διαμόρφωση αυτού του πλαισίου (Latour, 1986, 2005; Law, 2008)

Στα πλαίσια αυτά αναδεικνύεται η σημαντική επιρροή του κοινωνικού Κονστρουκτιβισμού στην συγκρότηση μιας θεωρητικής κατασκευής η οποία ενστερνίζεται την τεχνολογία και την επιστήμη, μάλλον τεχνοεπιστήμη σύμφωνα με τον Λατούρ ως μια διαδικασία κοινωνικής παραγωγής και όχι απλώς ένα αποτέλεσμα των κοινωνικών διαδικασιών. Στην θεωρία δρώντων δικτύου, με σαφή την επιρροή της κατασκευής του Deleuze οι οντότητες σχηματίζουν και σχηματίζονται μέσω μιας συνεχούς διαδικασίας αλληλεπίδρασης τονίζοντας έτσι τη χρονική συνέχεια των δικτύων, καθώς και τη σχεσιακή προοπτική που τις διέπει.

Η εμφάνιση της θεωρίας δρώντων δικτύου μπορεί επίσης να κατανοηθεί ως μέρος μιας ευρύτερης τάσης προς τη μελέτη των κοινωνικών και πολιτισμικών διαστάσεων της τεχνολογίας. Στη δεκαετία του 1980, άρχισε να αναγνωρίζεται όλο και περισσότερο ότι η τεχνολογία δεν ήταν απλώς ένα εργαλείο ή ένα μέσο για την επίτευξη ενός σκοπού, αλλά ήταν ενσωματωμένη σε κοινωνικές και πολιτιστικές πρακτικές. Η αναγνώριση αυτή οδήγησε στην ανάπτυξη νέων διεπιστημονικών πεδίων, όπως οι Επιστημονικές και Τεχνολογικές Σπουδές, οι οποίες προσπάθησαν να κατανοήσουν τις πολύπλοκες σχέσεις μεταξύ τεχνολογίας, κοινωνίας και πολιτισμού.

Η θεωρία δρώντων δικτύου αναπτύχθηκε λοιπόν βασιζόμενη στην προοπτική της κοινωνικής κατασκευασιοκρατίας, της ανάδειξης της επιστήμης και στην συνέχεια της τεχνολογίας ως αποτέλεσμα κοινωνικών διεργασιών και όχι ενός αυθύπαρκτου πλάνου εξέλιξης ή ηθικών επιταγών και νόμων που διασφαλίζουν την εξελικτική τους

πορεία. Παρά την έντονη επιρροή που δέχτηκε από τον Κοινωνικό Κονστρουκτιβισμό, διατηρεί έντονα χαρακτηριστικά διαφοροποίησης από αυτόν, σχηματίζοντας μια, όχι τόσο για τον Τζόν Λο, θεωρία η οποία συνιστά μια απάντηση στους περιορισμούς των παραδοσιακών κοινωνιολογικών θεωριών και της στάσιμης οπτικής τους σχετικά με τον χειραγωγήσιμο και παθητικό ρόλο της ύλης στην διαμόρφωση του κοινωνικού. (Baron & Gomez, 2016).

A.10.3 Η απομάκρυνση της θεωρίας δρώντων δικτύου από την κοινωνική κατασκευασιοκρατία

Όπως προαναφέρθηκε, η θεωρία Δρώντων Δικτύου δέχθηκε σημαντική επιρροή από τον Κοινωνικό Κονστρουκτιβισμό, ο οποίος ενίσχυσε σημαντικά τις συζητήσεις των κοινωνικών επιστημών διαμορφώνοντας ένα νέο πεδίο συζητήσεων σχετικά με τον κοινωνικό χαρακτήρα της επιστήμης και της τεχνολογίας. Η αποφασιστική συνδρομή του Ισχυρού Προγράμματος της Κοινωνιολογίας της Επιστημονικής Γνώσης του Bloor, την δεκαετία του 1970, λειτούργησε ανασταλτικά απέναντι στην ηγεμονία της Κοινωνικής κατασκευής, απομακρύνοντας την κοινωνιολογία από το απυρόβλητο, το «μονοπάλιο της ουσίας» όπου τα πάντα ανάγονται σε κοινωνικογλωσσικές κατασκευές.

Σύμφωνα με τα παραπάνω, η κοινωνιολογία της επιστημονικής γνώσης έπρεπε πλέον να είναι αναστοχαστική, τα εξηγητικά πρότυπα που χρησιμοποιεί δηλαδή, θα έπρεπε να εφαρμόζονται στον ίδιο της τον εαυτό ώστε να εξηγούν τις ίδιες της ερμηνείες που προκύπτουν μέσω της ανάγνωσης των φαινομένων από τις κοινωνιολογικές θεωρίες. Βάσει του Προγράμματος, οι γνωσιακές κατασκευές παύουν πλέον να ερμηνεύονται μονομερώς. Οι παράγοντες που καθιστούν μια γνωσιακή κατασκευή αληθινή είναι ίδιοι που την διαφοροποιούν από μια ψευδή, μη αληθινή μορφή γνώσης.

Η αρχή της συμμετρίας όπως αυτή συγκροτήθηκε από την σχολή του Εδιμβούργου δεν περιορίστηκε μόνο στην επιστημονική γνώση αλλά εντάχθηκε, όπως αναλύθηκε σε προηγούμενη ενότητα, στα πλαίσια της κοινωνιολογικής εξέτασης της τεχνολογίας. Οι τεχνολογικές κατασκευές δεν εξετάζονται μόνο στα πλαίσια των επιτυχιών τους. Οι τεχνολογικές κατασκευές εξετάζονται τόσο για την επιτυχία τους όσο και για την αποτυχία τους, με την διερεύνηση της πορείας μιας τεχνολογικής

κατασκευής να αποτελεί την εξήγηση της θέσης αυτή της κατασκευής στην κοινωνία και όχι το αντίθετο.

Με αυτόν τον τρόπο, οριοθετείτε η σχετικιστική απολυτότητα του Κοινωνικού κονστρουκτιβισμού αφού κάθε κατασκευή δεν λαμβάνεται ως δεδομένη αλλά αντιθέτως, γίνεται αντιληπτή ως ένα «μαύρο κουτί» το οποίο επιτάσσει συμμετρικής εξετάσεως και προσεκτικής εξέτασης. Υπό τις συνθήκες αυτές, ακολούθησαν πολλές αλλαγές στον τρόπο με τον οποίο η επιστήμη και η τεχνολογία αναλύονταν στον χώρο των κοινωνικών επιστημών. Αποφασιστική συμβολή, μέσω των ριζοσπαστικών εθνογραφικών μελετών της επιστήμης, ήταν αυτή του Μπρούνο Λατούρ και το Στιβ Γούλγκαρ. Οι Λατούρ και Γούλγκαρ, μέσω του έργου τους «Laboratory life _The social construction of scientific facts» (Latour & Woolgar, 1986) επιχειρούν να ανοίξουν το μαύρο κουτί της τεχνοεπιστήμης και της συγκρότησης της επιστημονικής γνώσης, αναπτύσσοντας παράλληλα τις αρχές της Θεωρίας Δρώντων Δικτύου.

Μια αρχή που υιοθετήθηκε από την Θεωρία Δρώντων Δικτύου είναι αυτή της γενικευμένης συμμετρίας. Η αρχή αυτή αναπτύχθηκε απέναντι στην αρχή του «Ισχυρού προγράμματος» του Τόμας Μπλέρ. Η γενικευμένη συμμετρία συνιστά ένα «αίτημα της διάσωσης της υλικότητας του φυσικού και κοινωνικού κόσμου και ενός ρεαλισμού απέναντι σε έναν διευρυμένο κοινωνιολογισμό» (Γεωργοπούλου, 2010, σσ.171-172) όπου τα πάντα ανάγονται σε κοινωνικές κατασκευές. Ο διαχωρισμός που προάγει το ισχυρό πρόγραμμα ανάμεσα στους «νικητές και ηττημένους» της επιστήμης και της τεχνολογίας, προάγει τον έντονο μονισμό της κονστρουκτιβιστής οπτικής, της σχετικότητας και ρευστότητας του κοινωνικού και της ερμηνείας του. Απέναντι σε μια συμμετρία που εξετάζει ισοδύναμα την αλήθεια και το ψέμα, τους νικητές και του ηττημένους, οι δύο κοινωνικοί επιστήμονες προτάσσουν μια περαιτέρω , ισοδύναμη εξηγητική ανάλυση που αφορά τις ανθρώπινες και τις μη ανθρώπινες οντότητες.

Στα πλαίσια αυτά αναδύεται η νέα οντολογία που προτείνει η Θεωρία Δρώντων Δικτύου σχετικά με τους ανθρώπινους και μη ανθρώπινους δρώντες. Για τον Λατούρ, ένας διαχωρισμός ανάμεσα σε ανθρώπινους και μη δρώντες, δεν πραγματοποιεί τίποτε άλλο, πάρα να ενισχύει την κονστρουκτιβιστική οπτική, στην οποία ο φυσικό-τεχνικός κόσμος απορροφάτε πλήρως από τον κοινωνικό κόσμο (Latour όπως αναφέρεται Γεωργοπούλου, 2010).

Σε αυτό το σημείο, θα πρέπει να διευκρινιστεί ότι η θεωρία των Δρώντων Δικτύου συνιστά κατά βάση μια υλιστική θεωρία. Αντιτάσσεται στην ουσιοκρατία του θετικισμού αλλά και στην αντιουσιοκρατική θεώρηση του κονστρουκτιβισμού. Και οι

δύο αυτές προσεγγίσεις διαχωρίζουν την ύλη από τον άνθρωπο, η κάθε μια με τον δικό της τρόπο, ωστόσο θεωρούν την δεύτερη υποτελή στην πρώτη. Η ύλη τόσο στην θετικιστική όσο και στην κονστρουκτιβιστική φιλοσοφία διέπεται από παθητικότητα απέναντι στον άνθρωπο. Στο κέντρο των θεωρήσεων αυτών η ύλη μεταβάλλεται, δεν μεταβάλλει. Η κατασκευασιοκρατία αναγάγει τα πάντα σε κοινωνικές κατασκευές οι οποίες κατασκευάζονται διυποκειμενικά, αγνοώντας τη συμμετοχή των μη ανθρώπινων οντοτήτων στην διαμόρφωση της κοινωνικής πραγματικότητας.

Ο διαχωρισμός λοιπόν ανάμεσα σε άνθρωπο και ύλη θα ενίσχυε σημαντικά την ασύμμετρη κονστρουκτιβιστική θεώρηση, αναδεικνύοντας πάλι, οτιδήποτε σε μια κοινωνική κατασκευή δίχως να δίνεται σημασία σε οποιαδήποτε συμβολή της ύλης στην διαμορφωτική διαδικασία της πραγματικότητας. Απέναντι λοιπόν σε αυτή την ασύμμετρη εξέταση της πραγματικότητας όπου μόνο οι ανθρώπινες οντότητες συνιστούν ενεργητικούς δρώντες, ο Λατούρ με την θεωρία δρώντων δικτύου, αντιτάσσεται προάγοντας μια γενικευμένη μορφή συμμετρίας (Γεωργοπούλου, 2008, 2010), μια «υπερσυμμετρία» (Sismondo, 2016, σσ.137-155) όπου όλες οι οντότητες εξετάζονται ισοδύναμα όσον αφορά την παραγωγή της τεχνοεπιστήμης.

Συμπερασματικά, η Θεωρία Δρώντων Δικτύου αποτελεί μια από τις σημαντικότερες κοινωνικές θεωρίες που πραγματεύονται ζητήματα τεχνολογίας και επιστήμης. Η υλιστική της προσέγγιση μέσω των υβριδικών δρώντων που συμμετέχουν ενεργά στην συγκρότηση του τεχνοεπιστημονικού μέσω των δικτύων που αυτοί δημιουργούν προσφέρει μια νέα οντολογική διάσταση στην κατανόηση της τεχνολογίας και της επιστήμης. Η ANT καθώς και οι «Μελέτες Εργαστηρίου», μέσω της προσέγγισης των διαδικασιών παραγωγής ως ενέργειες με κοινωνικό χαρακτήρα, διασπούν όλους τους προαναφερθέντες δυϊσμούς ανάμεσα σε τεχνολογία και επιστήμη. Επηρεασμένος από αυτές, ο ένας εκ των δύο διαμορφωτών τους, ο Γούλγκαρ, αντιλαμβανόμενος την ανάπτυξη του επιστημονικού (για εκείνη την εποχή) χώρου των «expert systems», έθεσε τους πρώτους κοινωνιολογικούς προβληματισμούς σχετικά με την ύπαρξη «μηχανών», οι οποίες μπορούν από μόνες τους να δημιουργούν προβλήματα που οι ίδιες λύνουν. Ο κοινωνιολογικός προβληματισμός του Γούλγκαρ αποτελεί την αφετηρία της κοινωνιολογικής ενασχόλησης με τα γνωστά την σημερινή εποχή συστήματα τεχνητής νοημοσύνης.

B. Κοινωνίες της Τεχνητής Νοημοσύνης

Τη δεκαετία του 1980, ο Στίβ Γούλγκαρ, αφουγκραζόμενος την ανάπτυξη ενός ανερχόμενου κλάδου των expert systems, πραγματεύτηκε το τότε νέο στον χώρο των Κοινωνικών Επιστημών ζήτημα της Τεχνητής Νοημοσύνης. Στο κείμενο αυτό χρησιμοποίησε έναν τίτλο, οποίος εξέφρασε ταυτόχρονα την απορία του για τον τομέα των συστημάτων των ειδικών «*Why not a Sociology of Machines*» (1985). Η πρόταση-ερώτηση του Γούλγκαρ, πόσο μάλλον το περιεχόμενο του έργου του, αποτέλεσε μια συνειδητοποίηση ενός ανερχόμενου τεχνοεπιστημονικού πεδίου το οποίο θα επηρέαζε σημαντικά την δομή των κοινωνιών αλλά ακόμα και την ίδια την Κοινωνιολογία (Airoldi, 2021). Από την στιγμή συγκρότησης της τεχνολογίας των αλγορίθμων και της τεχνητής νοημοσύνης, έως και σήμερα, την εποχή των Μεγάλων Δεδομένων (Kitchin, 2021), όλες οι τεχνοεπιστημονικές ανακαλύψεις διαμορφώνουν όλο και περισσότερο τα σύγχρονα οικονομικά, κοινωνικά, πολιτικά και για πολλούς υβριδικά (Latour,2005;Γεωργοπούλου,2010 ;Blok & Jensen, 2011a; Beck,2015) περιβάλλοντα.

Οι αλλαγές αυτές έχουν προσεγγιστεί από πολλούς κοινωνικούς επιστήμονες, καθώς το πεδίο της τεχνητής νοημοσύνης έχει επηρεάσει τόσο τον Ιδιωτικό όσο και τον δημόσιο τομέα. Η παρούσα εργασία θα εστιάσει στον δεύτερο τομέα, αυτόν επιρροής των αλγορίθμων και της τεχνητής νοημοσύνης στην κοινωνία, η οποία χάρη σε αυτές τις συνθήκες γίνεται όλο και πιο πολύ περίπλοκη. Οι σύγχρονες κοινωνίες αξιολογούν και διαχειρίζονται την πληροφορία με πολύ διαφορετικό τρόπο από τις κοινωνίες πριν μια δεκαετία. Η επανάσταση των μεγάλων δεδομένων δεν μετέβαλε απλώς τις μεθόδους διακυβέρνησης και τον τρόπο που οι πολίτες αλληλοεπιδρούν μεταξύ τους. Δημιούργησε παράλληλα νέα επιστημονικά πεδία, όπως την επιστήμη των δεδομένων, νέες θέσεις εργασίας, όπως αναλυτές, μηχανικούς δεδομένων ενώ ακόμα αλλάζει και τον τρόπο που οι κυβερνήσεις λαμβάνουν επιλογές, αντιδρούν σε κινδύνους και ασκούν τις πολιτικές τους (Crawford,2021;Kitchin,2021Benjamin 2019).

Ωστόσο, παρά τις ιδιαίτερα αποδοτικές εφαρμογές της τεχνητής νοημοσύνης στον δημόσιο τομέα, στην δημόσια διακυβέρνηση αλλά και γενικότερα στην Κοινωνία, η τεχνητή νοημοσύνη δεν εφαρμόζεται ισότιμα σε όλες τις κοινωνικές ομάδες, διευρύνοντας έτσι τις κοινωνικές ανισότητες. Συγκεκριμένα, οι αλγόριθμοι που πλέον χρησιμοποιούνται ευρέως στον δημόσιο τομέα με σκοπό την περάτωση πολλών διαδικασιών, οι οποίες παλαιότερα διενεργούνταν από υπαλλήλους, μεροληπτούν με

βάσει το φύλο, την εθνικότητα, την φυλή, το χρώμα, την ονομασία και τα εξωτερικά, φυσιογνωμικά χαρακτηριστικά (O’Neil, 2016; Eubanks, 2018; Benjamin, 2019; Crawford, 2021).

Η άνιση αυτή αντιμετώπιση μέσω των μεροληπτικών αλγορίθμων που χρησιμοποιούνται σε συστήματα τεχνητής νοημοσύνης στον δημόσιο τομέα, στην Δημόσια Διακυβέρνηση και Διοίκηση, δεν αποτελούν απλώς συστήματα άνισης αντιμετώπισης πολλαπλών κοινωνικών πληθυσμών εντός των κοινωνιών. Λειτουργούν πολλές φορές χρησιμοποιώντας λογισμικά που εκπορεύονται από ιδιωτικούς κολοσσούς, οι οποίοι ασχολούνται με την συλλογή δεδομένων και πληροφοριών, παράγοντας λογισμικά ανάλυσης και προβλέψεων. Ακόμα και αν αυτά συγκροτούνται κάτω από νομικά πλαίσια προστασίας των προσωπικών δεδομένων και των πλαισίων προστασίας των ανθρωπίνων δικαιωμάτων, δεν λειτουργούν ουδέτερα για όλους τους πληθυσμούς (O’Neil, 2016; Eubanks, 2018; Benjamin, 2019).

Η λειτουργία των αλγορίθμων αυτών, είναι πολλές φορές δύσκολο να κατανοηθεί και κατ' επέκταση να γίνει γνωστή στο ευρύ κοινό. Βρίσκεται διαρκώς σε μια συνθήκη αορατότητας και ακατανοησίας, πολλές φορές ακόμα και από τους ίδιους τους δημιουργούς της. Η έντονη διεπιστημονικότητα η οποία διέπει την συγκρότηση και ανάπτυξη τέτοιων τεχνοεπιστημονικών δημιουργημάτων, σε συνάρτηση με τις οικονομίες και τις πολιτικές που αναπτύσσονται γύρω από την τεχνητή νοημοσύνη, αλλά και τις «μάχες» και τους «αγώνες» γύρω από αυτήν (Chow & Perrigo, 2023), καθιστούν πολλές πτυχές της τεχνητής νοημοσύνης μαύρα κουτιά. Μια τέτοια συνθήκη διαταράσσει τα πλαίσια διαφάνειας και κατανόησης των τεχνολογιών, οι οποίες εφαρμόζονται παγκοσμίως, παγκοσμιοποιώντας ταυτόχρονα διακινδυνεύσεις και αβεβαιότητες μέσα από ψηφιακές αορατότητες. Οι κοινωνίες της Διακινδύνευσης μετατοπίζονται από το φυσικό περιβάλλον της γης σε κατακερματισμένα, πολυσύνθετα, ψηφιακά περιβάλλοντα, όπου μαύρα αλγοριθμικά κουτιά επηρεάζουν πολλές πτυχές της ζωής των πολιτών σε έναν πολύπλοκο, υβριδικό κύκλο. Με λίγα λόγια, οι Κοινωνίες της Διακινδύνευσης μετατρέπονται σε Κοινωνίες των Μαύρων Κουτιών (Paquale, 2016) δίχως να χάνουν τα δομικά χαρακτηριστικά που αναγνώρισε σε αυτές ο Μπέκ (Beck, 2016).

Η πολυπλοκότητα που προαναφέρθηκε σχετικά με το πεδίο της τεχνητής νοημοσύνης, τις αλγοριθμικές μεροληψίες, τις εφαρμογές τεχνολογιών τεχνητής νοημοσύνης αλλά και των περιστατικών μεροληψίας και άνισης μεταχείρισης θα

αναπτυχθούν εν συνεχεία. Στα πλαίσια αυτά, θα πραγματοποιηθεί μια σύντομη αναζήτηση των ιστορικών καταβολών της τεχνητής νοημοσύνης, μια ανάλυση των βασικών χαρακτηριστικών της, καθώς και μια διερεύνηση της σχέσης της με την Κοινωνιολογία και τον επιστημονικό χώρο των STS.

B.1 Οι καταβολές της τεχνητής νοημοσύνης και η πορεία της έως σήμερα

Ο έντονος, πολεμικός ανταγωνισμός του Β παγκοσμίου πολέμου συγκρότησε ένα ευνοϊκό κλίμα για την ανάπτυξη τεχνολογιών, οι οποίες θα έδιναν στρατιωτικό προβάδισμα σε ορισμένα κράτη (De Spiegeleire και άλλοι, 2017). Αμέσως μετά τον πόλεμο, το 1945, ο Άλαν Τιούρινγκ είχε δημιουργήσει ένα «πίνακα οδηγιών», ο οποίος αποτελούσε ουσιαστικά μια μορφή υπολογιστικού προγράμματος. Οι τεχνολογικές υποδομές καθώς και το επιστημονικό δυναμικό, το οποίο πλέον διέθετε την ικανότητα ενασχόλησης με αυτήν την πολύπλοκη τεχνολογία, είχαν ξεκινήσει ήδη να εμφανίζονται.

Τη δεκαετία του 1950, ο Άλαν Τιούρινγκ, αξιολογώντας τις δυνατότητες των «έξυπνων μηχανών», θέτει σαν προβληματισμό το κατά πόσο η νοημοσύνη, η οποία θεωρούντο ένα αποκλειστικά ανθρώπινο χαρακτηριστικό, θα μπορούσε σαν διαδικασία να πραγματοποιείται μέσω της εφαρμογής ισχυρών αλγοριθμικών διαδικασιών (Bieri, 1988). Έτσι, ο Τιούρινγκ, προτείνει ένα τεστ ελέγχου, γνωστό σήμερα ως έλεγχος Τιούρινγκ, βάσει του οποίου σε ένα σενάριο επικοινωνίας ενός δέκτη (ανθρώπου) με έναν άνθρωπο και μια μηχανή, χωρίς να υπάρχει γνώση για την ταυτότητα του κάθε πομπού, ο δέκτης πρέπει να τακτοποιήσει την ταυτότητα του κάθε πομπού. Στην περίπτωση που στο μήνυμα, το οποίο συγκροτήθηκε από κάποιο μηχανήμα, ο Πομπός δεν είναι σε θέση να ταυτοποιήσει την τεχνητή προέλευσή του, το συγκεκριμένο τεχνούργημα ελέγχεται θετικά (Turing, 2004). Η αξιολόγηση του μηχανήματος σαν ένα μηχανήμα του οποίου τα μηνύματα δεν μπορούν να ταυτοποιηθούν όσον αφορά την προέλευσή τους, συνιστά μια από τις πρώτες σοβαρές κοινωνικές προβληματικές, καθώς, όχι μόνο αναδεικνύει την αναγνώριση μιας προοπτικής επικοινωνίας ανθρώπου μηχανής, αλλά ταυτόχρονα αποτελεί έναν ευρύτερο οντολογικό προβληματισμό ο οποίος θα απασχολήσει μελλοντικά πληθώρα επιστημονικών πεδίων.

Σε αυτά τα πλαίσια, η συζήτηση γύρω από μια μορφή υπολογιστικής νοημοσύνης είχε ξεκινήσει ήδη. Μια από τις πρώτες μορφές πραγμάτωσής του, νέα για τα δεδομένα της εποχής εγχειρήματος της τεχνητής νοημοσύνης, έλαβε χώρα μαζί με την πρώτη άνοιξη

της τεχνητής νοημοσύνης. Την δεκαετία του 1950 στο Κολλέγιο του Ντάρμουθ, μια ομάδα επιστημόνων, με υπεύθυνο τον Τζον Μακάρθι (John McCarthy), ανέπτυξε το ερευνητικό της έργο γύρω από την υπόθεση ότι οι ανθρώπινες λογικές διεργασίες θα μπορούσαν να αναπτυχθούν με την μορφή αλγορίθμων (Verdegem, 2021). Οι επιστήμονες που ασχολήθηκαν με το ζήτημα της τεχνητής νοημοσύνης εκείνη την περίοδο, ανέπτυξαν εργαλεία ικανά να διεκπεραιώσουν πολλές ενέργειες, οι οποίες μέχρι τότε είχαν αναπτυχθεί από ανθρώπους, όπως αλγεβρικές και γεωμετρικές αποδείξεις ή και παιχνίδια, εξελίξεις που οδήγησαν και στην χορήγηση «σοβαρών ποσοστών επιστημονικών- και εξαιτίας της αύξουσας πίεσης του ψυχρού πολέμου, στρατιωτικών επίσης, χορηγήσεων» (De Spiegeleire και άλλοι, 2017, σ.31)

Παρά την ισχυρή και ένθερμη διάθεση της ακαδημαϊκής κοινότητας γύρω από το πεδίο της τεχνητής νοημοσύνης και τα ανθηρά αποτελέσματα της πρώτης άνοιξής της, η ανάπτυξη και το ενδιαφέρον αυτό ήρθαν αντιμέτωπα με τον πρώτο χειμώνα του πεδίου. Με τον όρο χειμώνας περιγράφονται «περίοδοι κατά τις οποίες η χρηματοδότηση και το δημόσιο ενδιαφέρον για τον χώρο της τεχνητής νοημοσύνης μειώνονται. Εμφανίζονται όταν οι υποσχέσεις και οι δυνατότητες των εξελίξεων της ΤΝ δεν επαρκούν ή αποτυγχάνουν να αποδώσουν την επένδυση, με αποτέλεσμα να μειώνεται η προσοχή και ο ενθουσιασμός στον κλάδο» (Glover, 2022). Σημαντικό γεγονός, το οποίο οδήγησε στην έλευση του νέου χειμώνα της τεχνητής νοημοσύνης, ήταν η αρνητική κριτική που δέχθηκε από αξιολογήσεις και εκθέσεις, όπως αυτή του Lighthill, που δημοσιεύθηκε στις αρχές της δεκαετίας του 1970. Σε αυτή, ισχυρό σημείο κριτικής δέχθηκε η αδυναμία της κοινότητας των ειδικών να αντιμετωπίσει το ανερχόμενο επιστημονικό πρόβλημα της συνδυαστικής έκρηξης (De Spiegeleire και άλλοι, 2017; Schuchmann, 2021b).

Ο Schuchmann (2021b), πραγματευόμενος τον όρο «συνδυαστική έκρηξη», τον παρουσιάζει ως «ένα γνωστό πρόβλημα σε χώρους αναζήτησης, όπως τα δέντρα, όπου ο αριθμός των κόμβων αυξάνεται εκθετικά όταν προχωράμε προς τα κάτω στο δέντρο». Εξηγώντας τον περεταίρω, φέρνει το παράδειγμα του σακακιού, όπου κάθε κίνηση από τις πρώτες που πραγματοποιεί ο σκακιστής στην αρχή του παιχνιδιού, δεν μειώνει την διάρκεια της παρτίδας, αντιθέτως αυξάνει την συνθετότητά της. Μέσω των δυνητικών κινήσεων που μπορούν να πραγματοποιηθούν, ξεκινώντας από είκοσι (20) με την πρώτη κίνηση και φτάνοντας τις τέσσερα εκατομμύρια οκτακόσιες εξήντα πέντε χιλιάδες και εξακόσιες εννέα (4.865.609) δυνητικές εκβάσεις, μόλις στην πέμπτη κίνηση. Με αυτόν τον τρόπο, η τεχνητή νοημοσύνη σταμάτησε να δέχεται μεγάλο

μέρος των επιχορηγήσεων που λάμβανε την πρώτη δεκαετία, μειώνοντας συνεπώς και το αντίστοιχο ακαδημαϊκό ενδιαφέρον.

Την μειωμένη εξέλιξη του κλάδου της τεχνητής νοημοσύνης της δεκαετίας του 1970, διαδέχθηκε μια καλύτερη δεκαετία. Η ανάπτυξη των ειδικών συστημάτων (expert systems), «οντοτήτων που πραγματοποιούν ενέργειες που θεωρούνται ανθρώπινες, σε συνάρτηση της προσβάσεώς τους και χρήσης βάσεων γνώσης» (Woolgar,1985), αυξήθηκε κατά τη διάρκεια της δεκαετίας του 1980, και η πρόοδος που σημειώθηκε από τους φυσικούς και τους επιστήμονες υπολογιστών, μαζί με την εφεύρεση του αλγορίθμου οπισθοδιάδοσης, έπαιξε ουσιαστικό ρόλο στην αναζωπύρωση του ενδιαφέροντος για την έρευνα στα νευρωνικά δίκτυα. Προκειμένου να δημιουργηθεί μια τεχνητή νοημοσύνη που θα μπορούσε να κάνει φυσικές συνομιλίες και να ερμηνεύει εικόνες με γρήγορη πρόσβαση σε τεράστιες αποθήκες κανόνων, μέσω μαζικών παράλληλων υπολογιστικών συστημάτων, η Ιαπωνία δρομολόγησε το έργο «Fifth Generation Computer Systems Project» το 1981 (De Spiegeleire και άλλοι, 2017).

Η ίδια διαδικασία με την οποία η τεχνητή νοημοσύνη, ως επιστημονικό πεδίο, βιώνει διαδοχικές περιόδους ακμής και παρακμής, ανοίξεις και χειμώνες, συνεχίζεται έως και τη πρώτη δεκαετία του 2000. Μετά το 2010 το πεδίο της τεχνητής νοημοσύνης βιώνει μια ραγδαία ανάπτυξη, χάρη στην επανάσταση των μεγάλων δεδομένων (Runting και άλλοι, 2020). Η επιστημονικότητα του πεδίου διευρύνθηκε σημαντικά εξαιτίας των νέων αναγκών που προέκυψαν από την διογκωμένη, πλέον, ανάπτυξη της τεχνητής νοημοσύνης, σε συνάρτηση με τον όγκο των δεδομένων από πολλαπλές υπηρεσίες και οργανισμούς. Η δημοκρατικοποίηση των δεδομένων και η ευρεία υιοθέτηση τεχνολογιών που αφορούν την διαχείριση και αξιοποίησή τους, ανέπτυξαν ένα τεχνολογικό περιβάλλον συνύπαρξης των δύο αυτών τεχνοεπιστημονικών επιτευγμάτων, όπου πολλές φορές η χρήση του ενός λειτουργεί ευεργετικά για το άλλο (Sinur, 2019)

Ενώ λοιπόν το επιστημονικό πεδίο των αλγορίθμων ήρθε αντιμέτωπο με αρκετούς «χειμώνες», οι περίοδοι ακμής που γνώρισε, σε συνάρτηση με την επανάσταση στον τρόπο διαχείρισης των δεδομένων καθώς και των νέων υποδομών που συγκροτήθηκαν για την επιχειρησιακή εκμετάλλευσή τους, αποτέλεσαν συνθήκες αρκετές ικανές, όσον αφορά την συγκρότηση της επανάστασης των Αλγορίθμων και των Κοινωνιών της Τεχνητής Νοημοσύνης.

B.2. Κοινωνιολογία της Τεχνητής Νοημοσύνης

B.2.1.H (Α)πορεία για την Κοινωνιολογία της Τεχνητής Νοημοσύνης

Η τεχνητή νοημοσύνη συνιστά έναν ευρύ όρο, ο οποίος εμφανίζεται πλέον εντονότερα από ποτέ στις ανθρώπινες κοινωνίες. Μοντέλα μηχανικής μάθησης χρησιμοποιούνται για την προσωποποιημένη πρόταση τραγουδιών σε εφαρμογές μουσικής (Barthle, 2023), για την πρόταση συντρόφου σε εφαρμογή γνωριμιών (Szaniawska-Schiano, 2023), για εξατομικευμένες διαφημίσεις και προτάσεις από διάφορους οργανισμούς (Crawford, 2021), ενώ αλγόριθμοι χρησιμοποιούνται γενικότερα στον δημόσιο Τομέα, στη διακυβέρνηση των πολιτών, στην εφαρμογή μέτρων και πολιτικών όπου, «αυξανόμενη διαθεσιμότητα τεράστιων και πρωτοφανών κλιμάκων δεδομένων και οι πολύπλοκες μαθηματικές και υπολογιστικές εξελίξεις συνδυάζονται με ισχυρούς τρόπους για να ερμηνεύσουν, να ενημερώσουν και να επηρεάσουν τη διακυβέρνηση των ατομικών συμπεριφορών, των οργανωτικών πρακτικών και των επιδόσεων των κοινωνιών»(Liu, 2022, σ.2) . Η ευρεία χρήση τέτοιων τεχνολογιών διαρρηγνύει την φύση της ιδιωτικής και δημόσιας σφαίρας, ενώ η επιρροή των εκροών τέτοιων εφαρμογών, επηρεάζει πλέον άμεσα την ζωή των ανθρώπων.

Οι τεχνολογίες των έξυπνων μηχανών, τα «expert systems», τα οποία εμφανίστηκαν αισθητά την περίοδο της δεύτερης άνοιξης της τεχνητής νοημοσύνης, το 1980-1987 (De Spiegeleire και άλλοι, 2017), προβληματίσαν με τις δυνατότητές τους και το ακαδημαϊκό δυναμικό που ασχολούνταν με αυτές, όπως τον κοινωνιολόγο από τον χώρο των STS, Στίβ Γούλγκαρ .Ο Γούλγκαρ (1985) παρατηρεί ότι η τεχνητή νοημοσύνη, όπως αυτή έχει τεχνοεπιστημονικά διαμορφωθεί, δύναται να συγκροτεί η ίδια συστήματα τα οποία μπορούν να λύσουν προβλήματα με τα οποία παλαιότερα καταπιάνονταν άνθρωποι. Όπως εξετάστηκε προηγουμένως, η μελέτη της κατασκευής του επιστημονικού, αφενός απασχόλησε διανοούμενους από τον χώρο των STS αλλά και τον ίδιο τον Γούλγκαρ, με την «μελέτη εργαστηρίων» (1986), αφετέρου αντιλαμβάνεται το επιστημονικό πεδίο της τεχνητής νοημοσύνης ως ένα ασταθές δίκτυο δρώντων που χρήζει ευέλικτων χειρισμών.

Στο έργο του «Why not a Sociology of Machines» (1985), ο Γούλγκαρ αναφέρεται στην αλλαγή που χαρακτηρίζει την Κοινωνιολογία της Επιστημονικής Γνώσης, έπειτα από την συνεισφορά του Κουν και του έργου του σχετικά με τις επιστημονικές επαναστάσεις. Ωστόσο, η αλλαγή αυτή πρέπει, πέρα από σημείο

αναφοράς, να διατηρήσει την ουσία της όσον αφορά την εξέταση της επιστήμης ως μια κοινωνική διαδικασία. Αναφερόμενος στην τεχνολογία των «expert systems», διαπιστώνει ότι πρόκειται για τεχνοεπιστημονικά επιτεύγματα, τα οποία μπορούν να δημιουργούν από μόνα τους συστήματα επίλυσης προβλημάτων, κάτι το οποίο θεωρούνταν ότι αποτελούσε μια ιδιότητα που κατείχαν αποκλειστικά οι άνθρωποι.

Η ιδιότητα αυτή των «expert systems» απασχόλησε ιδιαίτερα το έργο του Γούλγκαρ, καθώς αποτελούσε μια ανατρεπτική εικόνα για το τι είναι νοημοσύνη, ποια όντα διέπονται από αυτή, αλλά και για την κοινωνικότητα της τεχνητής νοημοσύνης (Mutzner,2020). Η σχεσιακή υλικότητα της θεωρίας δρώντων δικτύου, καθώς και η υβριδικότητα των δρώντων, στην συγκεκριμένη περίπτωση των αλγορίθμων της τεχνητής νοημοσύνης, καθιστούν τα «expert systems» πολύπλοκα στην χρήση τους και ανοιχτά όσον αφορά την ερμηνεία. Οι ρητορικές και τα λογοπλάισια που χρησιμοποιούνται από τους ειδήμονες του πεδίου της τεχνητής νοημοσύνης, αποτελούν διαμορφωτές της μορφής που αναπτύσσεται η τεχνητή νοημοσύνη.

Στα πλαίσια αυτά, ο Γούλγκαρ θεωρεί ότι η Κοινωνιολογία μπορεί πράγματι να εξετάσει την τεχνητή νοημοσύνη, παράγοντας εποικοδομητικά ευρήματα για αυτή. Οι τομείς που η κοινωνιολογία δύναται να διεισδύσει, αφορούν τις επιστημονικές κοινότητες, τις ερμηνείες και τα λογοπλάισια που αυτές χρησιμοποιούν. Η εξέταση της όμως δεν πρέπει να περιοριστεί μονάχα στον εντοπισμό των λογοπλαισίων αλλά, «θα πρέπει να λάβει ως θέμα τις διχοτομίες και τις διακρίσεις που χαρακτηρίζουν και συντηρούν αυτά τα λογοπλάισια» (Woolgar,1985, σ.567). Η δυαδικότητα των διχοτομιών και η μη ξεκάθαρη οντολογική διάκριση που διέπει αυτά τα συστήματα, εγείρει το Κοινωνιολογικό ενδιαφέρον γύρω από τις καθημερινές εφαρμογές της τεχνητής νοημοσύνης, του τρόπου που αυτή γίνεται κατανοητή αλλά και των συνεπειών της χρήσης της.

Η δεκαετία του 1980 αποτελεί μια δεκαετία σημαντικής ανάπτυξης τομέων που θεωρούνται τμήματα του όρου τεχνητή νοημοσύνη. Μέσα από το έργο του Γούλγκαρ αναπτύσσεται η ιδέα συγκρότησης μιας Κοινωνιολογίας, η οποία θα πραγματεύεται διάφορες πτυχές «των μηχανών», με την κοινωνιολογία να έχει καθυστερήσει την ενασχόληση με το συγκεκριμένο πεδίο, συγκριτικά με άλλες επιστήμες. Πραγματευόμενος τις δυνατότητες των «expert systems» της εποχής του και βασιζόμενος στην συνεισφορά του Κούν, στην κοινωνική κατασκευή της επιστήμης, ο Γούλγκαρ, παραθέτει την άποψη σχετικά με το αν η αναπτυσσόμενη αυτή τεχνολογία μπορεί να αποτελέσει ένα νέο επιστημονικό παράδειγμα (1985). Μια τέτοια

τοποθέτηση αναδεικνύει την δυναμική και ταχεία ανάπτυξη των μηχανών δημιουργίας συστημάτων επίλυσης προβλημάτων.

B.2.2..Η επανάσταση των μεγάλων δεδομένων και τα νέα τεχνοεπιστημονικά παραδείγματα

Η ανάπτυξη της τεχνητής νοημοσύνης συνιστά μια σύνθετη πορεία, όπως άλλωστε πολύπλοκος είναι και ο ίδιος ο τεχνοεπιστημονικός κλάδος που την συγκροτεί. Όπως φάνηκε από την πορεία του επιστημονικού πεδίου της τεχνητής νοημοσύνης, ο δρόμος για την συγκρότηση του δεν χαρακτηρίζεται από μια γραμμική ανάπτυξη. Η διεύρυνση του πεδίου των «expert systems» πραγματοποιήθηκε κάτω από συνθήκες ακμής και παρακμής. Οι χειμώνες της τεχνητής νοημοσύνης, οι χαμηλές χρηματοδοτήσεις που έλαβαν τα σχετικά ακαδημαϊκά προγράμματα, καθώς και τα σύνθετα προβλήματα που αντιμετώπισαν οι ειδήμονες, όπως αυτό της συνδυαστικής έκρηξης (Schuchmann, 2021b), αναδεικνύουν ότι οι επιστημονικοί κλάδοι των θετικών επιστημών δεν αναπτύσσονται όπως τα αντικείμενα τα οποία πραγματεύονται, δηλαδή αιτιακά. Δεν υπάρχει κάποια νομοτελειακή συνθήκη ανάπτυξης ενός επιστημονικού κλάδου. Η διαδικασίες συγκρότησής του, η διαμόρφωση των ιδιαίτερων χαρακτηριστικών του, τα λογοπλάσια τα οποία χρησιμοποιούνται κατά την διάρκεια καταγραφής παρατηρήσεων και της ανάλυσης των τεκμηρίων του, διέπονται από στοιχεία κοινωνικότητας.

Η τεχνητή νοημοσύνη, η οποία περιγράφεται στο έργο του Γούλγκαρ (1985), διαφέρει σε σημαντικό βαθμό, όσον αφορά την συνθετότητα, την πολυπλοκότητα, την διεπιστημονικότητα και την ύλη, με τη μορφή που αυτή έχει λάβει σήμερα. Η επανάσταση των μεγάλων δεδομένων, οι πόροι και οι εγκαταστάσεις που χρησιμοποιούνται για την διοχέτευση, αποθήκευση και αξιοποίηση αυτών, έχουν διαμορφώσει συνθήκες, τεχνολογίες και μεθόδους που ευνόησαν παράλληλα διάφορους κλάδους του σύνθετου πεδίου της τεχνητής νοημοσύνης. Πολλαπλά είδη δεδομένων μπορούν να αναλυθούν με ποικίλους δυνατούς τρόπους, μέσω διάφορων αλγοριθμικών διαδικασιών και κατά συνέπεια «εκατοντάδες διαφορετικοί αλγόριθμοι μπορούν να εφαρμοστούν σε ένα σύνολο δεδομένων για να προσδιοριστεί το καλύτερο ή ένα σύνθετο μοντέλο ή εξήγηση» (Siegel, 2013 όπως αναφέρεται στον Kitchin, 2014).

Τα μεγάλα δεδομένα και η ευρεία χρήση τους στις κοινωνίες, αποτελούν μια σημαντική τεχνολογική και επιστημονική τομή, η οποία χαρακτηρίζει την σημερινή εποχή. Η διαδικτυακή ανάπτυξη που γιγαντώθηκε μετά το 2000, συνέβαλε στην συγκρότηση Κοινωνιών, οι οποίες χαρακτηρίζονταν από την διακίνηση και την θεσμική αξιοποίηση των πληροφοριών που παράγονται από την χρήση ηλεκτρονικών υπολογιστών και από την χρήση του διαδικτύου. Αυτές οι μορφές κοινωνιών που θεωρήθηκαν «μοντέρνες» (Κάλλας, 2006) χαρακτηρίστηκαν πολλές φορές αναλόγως το τεχνολογικό επίτευγμα που δυναμικά τις διαμόρφωσε ως Κοινωνίες τις Πληροφορίας (Κάλλας, 2006), Κοινωνίες του Διαδικτύου (Castells, 2002), Κοινωνίες των Αλγορίθμων (Airoldi, 2021), Κοινωνίες των Μεγάλων Δεδομένων (Beaulieu & Leonelli, 2021). Οι τελευταίες μάλιστα, άσκησαν αποφασιστική επιρροή, τόσο στον τρόπο που λαμβάνονται αποφάσεις ή ασκούνται πολιτικές, όσο και στον τρόπο που οι πληροφορίες και τα δεδομένα αποτελούν αντικείμενο επεξεργασίας και ανάλυσης την σημερινή εποχή.

Η αξιοποίηση των μεγάλων δεδομένων συνιστά ένα ιδιαίτερα σημαντικό κομμάτι στην ιστορία της συγκρότησης του πεδίου της τεχνητής νοημοσύνης. Για τον Jim Gray, ο τρόπος διαχείρισης των δεδομένων, σε συνάρτηση με τον όγκο που αυτά παράγονται στις κοινωνίες, αποτελούν μια τομή στον τρόπο με τον οποίο η επιστήμη εξετάζει την πραγματικότητα. Οι Hey και άλλοι (2009), βασιζόμενοι στην θέση του Gray, ξεχωρίζουν τέσσερα «επιστημονικά παραδείγματα». Το πρώτο παράδειγμα, ονομάζεται «πειραματική επιστήμη», αφορά την εμπειρική εξήγηση της πραγματικότητας, όπως αυτή εκφράστηκε επηρεασμένη από τον θετικισμό την περίοδο του Διαφωτισμού. Το επόμενο παράδειγμα, αφορά την «θεωρητική επιστήμη», όπως αυτή αναπτύχθηκε πριν την περίοδο ανάπτυξης των συστημάτων υπολογισμού. Το παράδειγμα αυτό χαρακτηρίζεται από την μοντελοποίηση και την γενικοποίηση των πληροφοριών που επεξεργάζεται. Το τρίτο παράδειγμα, αφορά την «υπολογιστική επιστήμη», όπως αυτή αναπτύχθηκε την εποχή ανάπτυξης των υπολογιστών, πριν την «επανάσταση των μεγάλων δεδομένων», το οποίο επικεντρώνεται στην προσομοίωση σύνθετων φαινομένων.

Η άφιξη των μεγάλων δεδομένων σε συνδυασμό τεχνολογικής και επιστημονικής ανάπτυξης, των ισχυρών υπολογιστών, των κλάδων της μηχανικής μάθησης, αποσταθεροποιούν τους τρόπους βάσει των οποίων η επιστήμη αξιοποιούσε την πληροφορία. Για τον Kitchin, «σε αντίθεση με τις νέες μορφές εμπειρισμού, η επιστήμη με γνώμονα τα δεδομένα επιδιώκει να τηρεί τις αρχές της επιστημονικής

μεθόδου, αλλά είναι πιο ανοικτή στη χρήση ενός υβριδικού συνδυασμού επαγωγικών, απαγωγικών και επαγωγικών προσεγγίσεων για την προώθηση της κατανόησης ενός φαινομένου» (2014, σ.5). Σε αυτά τα πλαίσια, οι υστερονεωτερικές κοινωνίες δεν βιώνουν μονάχα μια απλή τεχνολογική ανάπτυξη. Για κάποιους τίθεται το ζήτημα, κατά πόσο η εισχώρηση τέτοιων όγκων δεδομένων συνιστά μια ικανή συνθήκη, ένα νέο επιστημονικό παράδειγμα, στα πλαίσια της Κουνιανής πραγμάτευσης (Kitchin, 2014).

B.2.3. Σύγχρονες μορφές της Κοινωνιολογίας της τεχνητής νοημοσύνης

Έτσι, ο επιστημονικός κλάδος της τεχνητής νοημοσύνης διευρύνεται και αυτός σημαντικά, χάρη στην ύπαρξη του τεράστιου όγκου δεδομένων. Κάτι τέτοιο συμβαίνει αφού «οι αλγόριθμοι μηχανικής μάθησης γίνονται πιο αποτελεσματικοί όσο αυξάνεται το μέγεθος των συνόλων δεδομένων εκπαίδευσης» (Udacity Team, 2020). Η τεχνολογία της τεχνητής νοημοσύνης αναπτύσσεται και διογκώνει και η ίδια τον αριθμό των μεγάλων δεδομένων, αποτελώντας στην ουσία μια «επαναστατική μορφή καυσίμου» για τις μηχανές προηγμένης νοημοσύνης. Συνεπώς, η επανάσταση των μεγάλων δεδομένων, σε συνδυασμό με τις εξελίξεις του πολυσύνθετου χώρου της τεχνητής νοημοσύνης, δημιουργούν ένα ιδιαίτερα πολύπλοκο, υβριδικό τεχνολογικό, επιστημονικό και τεχνοοικονομικό περιβάλλον, εντός του οποίου, πολλές διεργασίες πραγματοποιούνται σε πολλές πτυχές της ανθρώπινης ζωής, δυσχεραίνοντας τις απόπειρες για ξεκάθαρους διαχωρισμούς της τεχνητής νοημοσύνης και των μεγάλων δεδομένων.

Η ευρεία διάδοση και αξιοποίηση τεχνολογιών τεχνητής νοημοσύνης διευρύνουν και εντείνουν τις συζητήσεις, σχετικά με το αν η κοινωνιολογία πρέπει να συνεχίσει να ασχολείται με αυτή, αν δεχτούμε ότι ασχολείται ήδη, σε ποιο πεδίο και με ποιον τρόπο.

Ωστόσο, όπως έχει αναφερθεί προηγουμένως, ο όρος τεχνητή νοημοσύνη αποτελεί μια γενική ορολογία που πολλές φορές εμπεριέχει άλλες διεργασίες εντός του, όπως την μηχανική μάθηση, τα τεχνητά νευρωνικά δίκτυα και την βαθιά μάθηση. Η πολυπλοκότητα του πεδίου και η διεπιστημονικότητά του δυσκολεύουν ακόμα και επιστήμονες που διαχειρίζονται τέτοιων ειδών τεχνολογίες, οι οποίοι «είναι συχνά τόσο μπερδεμένοι όσο το κοινό» (Jordan, 2019, σ.2). Λαμβάνοντας υπόψη τα δεδομένα αυτά, η συνθετότητα των συστημάτων τεχνητής νοημοσύνης όπως αυτά αναπτύχθηκαν

στον χρόνο, επιτάσσουν την συζήτηση σχετικά με τους τρόπους με τους οποίους η Κοινωνιολογία μπορεί να εμπλακεί σε ένα τόσο δυναμικά εξελισσόμενο πεδίο.

Τα τεχνοεπιστημονικά συστήματα της τεχνητής νοημοσύνης έχουν απασχολήσει έντονα αρκετούς κοινωνικούς επιστήμονες εξαιτίας, όχι μόνο της διευρυμένης και παγκόσμιας πολυεπίπεδης χρήσης τους, αλλά και των αποτελεσμάτων που έχουν αυτά στις κοινωνίες. Πληθώρα πτυχών της κοινωνίας βιώνουν πολλαπλές πρωτόγνωρες συνθήκες σχετικά με την διεπιστημονική, και σε μερικές περιπτώσεις υβριδική, μορφή που λαμβάνουν τα τεχνοεπιστημονικά συστήματα στην (μετα)νεωτερική εποχή. Οι πορείες τέτοιων τεχνολογιών, τα λογισμικά και τα αλγοριθμικά συστήματα, συγκροτούνται από επιστήμονες και τεχνικούς σε εταιρίες συγκέντρωσης μεγάλων δεδομένων και ανάλυσης πληροφοριών και αξιοποιούνται από άλλες ιδιωτικές εταιρίες. Με την χρήση των νέων αυτών κοινωνικοτεχνικών συστημάτων, τα όρια μεταξύ ιδιωτικού και δημόσιου (Giddens, 1991) ξεπερνούν την μικρή σφαίρα του ατομικού.

Τα προσωπικά δεδομένα, τα οποία αν και προφυλάσσονται από νομικά πλαίσια, μεταξύ των οποίων και το GDPR, μετατρέπονται σε υλικό επεξεργασίας από αλγόριθμους μηχανικής μάθησης ιδιωτικών εταιριών και κρατικών οργανισμών (Pasquale, 2016). Η δημιουργία ψηφιακών δεδομένων συνιστά πλέον μια σημαντική καθημερινή κοινωνική πρακτική. Η επαφή των ανθρώπων με συσκευές παραγωγής ψηφιακών πληροφοριών έχει αναπτυχθεί σημαντικά χάρη στην ίδια την αύξηση «έξυπνων συσκευών», οι οποίες αξιοποιούνται πολύτροπα και μάλιστα για αρκετά χρονικά διαστήματα εντός της ημέρας (Dolan, 2023). Μέσω της χρήσης τέτοιων συσκευών, οι αλγόριθμοι έρχονται σε επαφή, επηρεάζουν και επηρεάζονται από τους ανθρώπους. Η διάδραση αυτή, πέρα από τον τεχνικό χαρακτήρα, διέπεται από έντονη κοινωνικότητα, καθιστώντας έτσι εύλογη την θέση - απορία την οποία ο Airoldi εκφράζει δεκαετίες μετά τον Γούλγκαρ, «Γιατί όχι μια κοινωνιολογία των αλγορίθμων;» (2021).

Στο έργο του Airoldi, η τεχνητή νοημοσύνη και πιο συγκεκριμένα, οι αλγόριθμοι, αποτελούν αναπόσπαστο κομμάτι της τεχνοεπιστημονικής καθημερινότητας των ανθρώπων. Η επιρροή τους στην καθημερινότητα των ανθρώπων αναδεικνύει την έντονη κοινωνικότητά τους. Η ανάδειξη μιας κοινωνίας των αλγορίθμων όμως, δεν πραγματοποιήθηκε με χρονική συντομία και ομαλότητα. Για τον Airoldi, δύο μείζονες τεχνολογικοί και κοινωνικοί μετασχηματισμοί που έλαβαν χώρα την τελευταία δεκαετία, έχουν καταστήσει την ανάγκη για μια κοινωνιολογία των

αλγορίθμων ιδιαίτερα επιτακτική. Η πρώτη, ποσοτική μετατόπιση, προέκυψε από την άνευ προηγουμένου διείσδυση των ψηφιακών τεχνολογιών στη ζωή και τις συνήθειες των ανθρώπων και των οργανισμών. Το παγκόσμιο οικονομικό σύστημα βασίζεται σε αλγορίθμους, δεδομένα και δικτυακές υποδομές, με αποτέλεσμα να εγκαθίστανται ταχύτερα συστήματα επικοινωνίας με μικροκύματα ή λέιζερ σε στέγες κοντά στα εμπορικά κέντρα της Νέας Υόρκης.

Η δεύτερη, ποιοτική αλλαγή, αφορά τα είδη των μηχανών και των τεχνολογιών τεχνητής νοημοσύνης που ενσωματώνονται στην ψηφιακή μας κοινωνία. Η ανάπτυξη και η βιομηχανική εφαρμογή αλγορίθμων μηχανικής μάθησης που επιτρέπουν στους υπολογιστές να μαθαίνουν από την εμπειρία, αποτέλεσε σημαντικό σημείο καμπής, καθώς η εμπειρία είναι ουσιαστικά ένα σύνολο δεδομένων ιστορικών γεγονότων και μάθηση σημαίνει «εντοπισμός και εξαγωγή χρήσιμων προτύπων από ένα σύνολο δεδομένων» (2021).

Η πραγμάτευση του Airoldi (2021), αναδεικνύει τις προοπτικές της Κοινωνιολογίας όσον αφορά την εμπλοκή της με την τεχνητή νοημοσύνη. Συγκεκριμένα, η Κοινωνιολογία προσφέρει πολλαπλά θεωρητικά πλαίσια, τα οποία μπορούν να εφαρμοστούν αναλυτικά σε διάφορα πεδία της τεχνητής νοημοσύνης. Ο Liu (2021) συμβάλλει σημαντικά στην κοινωνιολογική συζήτηση γύρω από την τεχνητή νοημοσύνη αφού, αντιλαμβανόμενος τις διαφορετικές πτυχές στις οποίες η αυτή και οι αλγόριθμοι εμπλέκονται και αλληλοδιαμορφώνουν την κοινωνική πραγματικότητα, προτείνει μια τυπολογία, η οποία στόχο έχει την οριοθέτηση του πεδίου της τεχνητής νοημοσύνης. Αυτά είναι η «Επιστημονική τεχνητή νοημοσύνη» (Scientific AI), η «Τεχνική Τεχνητή Νοημοσύνη» (Technical AI) και η «Πολιτισμική Τεχνητή Νοημοσύνη» (Cultural AI). Μάλιστα, στο κάθε ένα από αυτά παρουσιάζονται και τα επιστημονικά έργα, τα οποία σχετίζονται με την υπό συζήτηση κατηγορία τεχνητής νοημοσύνης.

Σημαντική συνεισφορά στην κοινωνιολογική συζήτηση των τύπων τεχνητής νοημοσύνης παρέχει το επιστημονικό πεδίο των STS. Οι σπουδές Επιστήμης και Τεχνολογίας δεν περιορίζονται μονάχα στην εξέταση των χαρακτηριστικών και των δυνατοτήτων της τεχνητής νοημοσύνης, αλλά συνεισφέρουν σημαντικά στην ανάγνωση των αλγοριθμικών συστημάτων και των συστημάτων τεχνητής νοημοσύνης ως κοινωνικότεχνικά συστήματα. Για παράδειγμα, στον χώρο της επιστημονικής γνώσης, εκτός από την θέση του Γούλγκαρ, αρκετό ενδιαφέρον παρουσιάζουν οι θέσεις του Bloomfield σχετικά με την παραγωγή της γνώσης. Σύμφωνα με αυτόν, η παραγωγή

της γνώσης αποτελούσε μια κοινωνική διαδικασία , ακριβώς επειδή οι άνθρωποι μέχρι τότε θεωρούνταν ότι διέθεταν το αποκλειστικό προνόμιο παραγωγής της. Ωστόσο, λόγω της ραγδαίας ανάπτυξης των τεχνολογιών της τεχνητής νοημοσύνης, όπως αυτή ιδίως έχει αναπτυχθεί με τα μοντέλα επεξεργασίας λόγου, όπως το ChatGPT από την Open AI, και το Google Bard από την Google, εντείνονται οι συζητήσεις σχετικά με το κατά πόσο η παραγωγή της γνώσης μπορεί να διεκδικεί πλέον χαρακτηριστικά επιστημονικότητας, αφού πολλά δεδομένα που αξιοποιούνται σε επιστημονικό πλαίσιο έχουν προέλευση από αυτές τις τεχνολογίες, καθώς και με το κατά πόσο αυτή συνιστά πλέον μια αποκλειστικά ανθρώπινη διαδικασία.

Η συζήτηση σχετικά με την «πολιτισμική τεχνητή νοημοσύνη», αναπτύσσει σημαντικό χώρο και για κοινωνιολογικό προβληματισμό και διάλογο γύρω από τα ζητήματα της τεχνητής νοημοσύνης. Όπως επισημαίνει ο Liu, «η «αναλυτική προοπτική της πολιτιστικής τεχνητής νοημοσύνης ανοίγει χώρους για τους κοινωνιολόγους σε άλλους υποτομείς της κοινωνιολογίας για να ερευνήσουν τον αντίκτυπο της τέχνης των νοημάτων στους τομείς τους.» (2021, σ.9). Μια ακριβώς τέτοια πραγμάτευση αποτελεί και αυτή του Airolti (2021) ο οποίος, αξιοποιώντας στην ανάλυσή του την θεωρία του Μπουρντιέ σχετικά με το πολιτισμικό κεφάλαιο και την κεντρική έννοια του «Habitus», εξετάζει την λειτουργία των αλγοριθμικών συστημάτων, την διάδραση τους με τους ανθρώπους και την αξιοποίησή τους από αυτούς. Αναζητώντας την κουλτούρα στον κώδικα (Airolti,2021), πέρα από την ανθρώπινη παρέμβαση στον κώδικα, αναδύεται η απορία σχετικά με το κατά πόσο η τεχνητή νοημοσύνη μπορεί να οδηγήσει σε μεροληπτικές αποφάσεις απέναντι σε συγκεκριμένες κοινωνικές ομάδες.

Οι μεροληψίες, οι διακρίσεις και τα στερεότυπα διαπερνούν οριζόντια όλες τις μορφές της τεχνητής νοημοσύνης, όπως αυτές προτάθηκαν από τον Zheng Liu στο έργο του «Sociological perspectives on artificial intelligence: A typological reading». Η διακινδύνευση της αυτοματοποιημένης παραγωγής ανισοτήτων, μέσω τεχνοπειστημονικών μεροληψιών, έχει αναπτύξει αρκετές κοινωνιολογικές συζητήσεις γύρω από τις μεροληψίες των μεγάλων δεδομένων, της μηχανικής μάθησης (Machine Learning), της βαθιάς μάθησης (Deep Learning), των αλγορίθμων και της τεχνητής νοημοσύνης. Σε αυτές τις πραγματεύσεις κυριαρχεί η επιταγή για κριτικές προσεγγίσεις.

Όπως αναπτύχθηκε και ανωτέρω, οι τεχνολογίες της τεχνητής νοημοσύνης, διαρρηγνύουν σε ένα γενικότερο πλαίσιο τα όρια του ιδιωτικού και το δημόσιου, τόσο

σε θέμα φορέων, όσο και σε προσωπικό επίπεδο. Οι κριτικές θεωρίες, όπως αυτές εκφράστηκαν αρχικώς στην σχολή της Φρανκφούρτης από προσωπικότητες, όπως ο Theodor W. Adorno, ο Max Horkheimer, ο Herbert Marcuse και ο Jurgen Habermas, στρέφουν το ενδιαφέρον τους «σε καταστάσεις και στοιχεία που παρεμποδίζουν την υλοποίηση ... των (δημοκρατικών) ιδανικών, αφετέρου σε τάσεις αλλαγής στο εσωτερικό της κοινωνίας.» (Preglau, 2014, σ.293). Η αντιληπτή αλλαγή των (μετα)νεωτερικών κοινωνιών μέσω των τεχνολογικών αλλαγών μεταβάλλουν την εστίαση της κοινωνιολογικής σκοπιάς, από τους τομείς που αλλάζουν την καθημερινότητα των πολιτών, στον τρόπο που πραγματοποιείται η διαδικασία αυτή και πιο συγκεκριμένα στην εξουσία.

Οι ερωτήσεις των κριτικών θεωριών γύρω από την τεχνητή νοημοσύνη δεν περιορίζονται μονάχα στις πτυχές της αλλαγής της καθημερινότητας των πολιτών, ούτε στο «πώς» πραγματοποιείται η αλλαγή, παρόλο που το τελευταίο, θα αποτελούσε ενδεχομένως ένα ερώτημα που θα εγείρετο στους κόλπους της θεωρίας δρώντων δικτύου. «Τι είναι και τι δεν είναι; Τι είδους τεχνητής νοημοσύνης χρειαζόμαστε, τι προσδοκίες υπάρχουν για αυτήν και ποιοι βρίσκονται πίσω από αυτές... αν θέλουμε να μιλήσουμε για τεχνητή νοημοσύνη...πρέπει να μιλήσουμε για εξουσία» (Verdegem, 2021, p2). Στα πλαίσια αυτά, οι Joyce και άλλοι, (2021) συνοψίζουν τις τοποθετήσεις σχετικά με τις κριτικές προσεγγίσεις της Κοινωνιολογίας της τεχνητής νοημοσύνης γύρω από δύο άξονες. Ο πρώτος άξονας αφορά τις πολιτικές των αλγορίθμων, των δεδομένων και του κώδικα, ενώ ο δεύτερος σχετίζεται με την κοινωνική διαμόρφωση των πρακτικών της τεχνητής νοημοσύνης. Και στους δύο αυτούς άξονες κυριαρχούν οι κριτικές προσεγγίσεις των μεροληψιών, οι οποίες εμφανίζονται σε διαδικασίες από την συλλογή δεδομένων και την κατηγοριοποίησή τους (Leurs & Shepherd, 2017; Benjamin, 2019), μέχρι και τον προγραμματισμό και την αυτοματοποίηση των αλγορίθμων, με σκοπό την εφαρμογή τους σε πολλούς τομείς της κοινωνικής ζωής (O'Neil, 2016; Eubanks, 2018; Benjamin, 2019; Crawford, 2021).

Από τα προαναφερθέντα, γίνεται κατανοητή η δυναμική του επιστημονικού πεδίου της Κοινωνιολογίας, όσον αφορά την εξέταση των τεχνολογιών της τεχνητής νοημοσύνης, των εσωτερικών διεργασιών αυτών των συστημάτων, καθώς και των αποτελεσμάτων τους στην κοινωνική πραγματικότητα. Η σημαντική εισχώρηση των αλγορίθμων στον δημόσιο τομέα εγείρει πολλές απορίες σχετικά με το κατά πόσο ορισμένα φαινόμενα, που εμφανίζονται παράλληλα με την χρήση τεχνολογιών

τεχνητής νοημοσύνης, θα αναπαραχθούν στην δημόσια και κοινωνική ζωή των πολιτών.

B.2.4. Κριτικές Σπουδές Αλγορίθμων-Critical Algorithm Studies (CAS)

Οι σπουδές Τεχνολογίας και Επιστήμης, STS, μέσα από τις πραγματεύσεις και τις θεωρίες αναδιαμόρφωσαν σημαντικά τον τρόπο με τον οποίο πραγματοποιούνταν η ανάγνωση της τεχνολογίας και της επιστήμης. Η κύρια συνεισφορά τους εντοπίζεται στο γεγονός ότι η διαδικασία παραγωγής του τεχνοεπιστημονικού, γίνεται ορατή κάτω από τα πλαίσια της κοινωνικής κατασκευής και της αλληλοδιαμόρφωσης του τεχνοεπιστημονικού και του κοινωνικού. Υπό το πρίσμα αυτό, τα σύγχρονα τεχνοεπιστημονικά συστήματα θεωρούνται παράλληλα κοινωνικοτεχνικά συστήματα (Bijker & Law, 1992). Έτσι λοιπόν, και τα συστήματα τεχνητής νοημοσύνης (AI systems), τα συστήματα μηχανικής μάθησης (ML systems), μελετώνται πλέον από πολλούς επιστήμονες του πεδίου των STS αλλά και της Κοινωνιολογίας.

Η Κοινωνιολογική προσέγγιση των αλγορίθμων, των μεγάλων δεδομένων, της μηχανικής μάθησης και των συστημάτων τεχνητής νοημοσύνης, έχει αναπτυχθεί με την σημαντική ενίσχυση της κριτικής κοινωνιολογική προσέγγισης. Τα νέα υβριδικά, κοινωνικοτεχνικά συστήματα δεν λαμβάνονται πλέον ως δεδομένα, όπως γινόταν πριν την αποφασιστική συμμετοχή των STS στην συζήτηση σχετικά με την επιστήμη και την τεχνολογία. Οι κριτικές κοινωνιολογικές θεωρίες εμπλουτίζουν και διευρύνουν σημαντικά τα προαναφερθέντα τεχνοεπιστημονικά πεδία, συγκροτώντας ένα νέο επιστημονικό χώρο, αυτόν των κριτικών σπουδών των αλγορίθμων, «Critical Algorithm Studies». Η αναγνώριση λοιπόν των αλγοριθμικών συστημάτων ως κοινωνικοτεχνικά συστήματα (Airoidi, 2021), αποτέλεσε πρόσφορο πεδίο για διάφορους μελετητές, ώστε να εξετάσουν κριτικά τον δυναμικά αναπτυσσόμενο και ευρέως πλέον εφαρμοσμένο χώρο της τεχνητής νοημοσύνης.

Πριν πραγματοποιηθεί οποιαδήποτε παρουσίαση διαφόρων θέσεων των μελετητών του πεδίου των κριτικών σπουδών της τεχνολογίας, CAS, θα γίνει λόγος πρώτα για τις καταβολές του επιστημονικού πεδίου και την ανάπτυξή του στον χρόνο. Η διευρυμένη εφαρμογή των αλγοριθμικών συστημάτων σε διάφορους τομείς της κοινωνικής ζωής, όπως τα οικονομικά, η διακυβέρνηση, τα μέσα δικτύωσης και η επιτήρηση, ανέδειξε την ισχυρή δυναμική αυτών εντός των κοινωνικών συνόλων. Ήδη, από τη δεκαετία του 2000 (Beer, 2009 as cited in Airoidi, 2021), διάφοροι μελετητές

από το πεδίο των STS (Benjamin, 2019), της Κοινωνιολογίας και των Social Media (Gillespie, 2018) ξεκίνησαν να μελετούν τις εφαρμογές των τεχνοεπιστημονικών αυτών συστημάτων στις κοινωνίες. Αυτοί προσέγγισαν κριτικά τους τρόπους με τους οποίους τα αλγοριθμικά συστήματα λήψης αποφάσεων (Decision Making Systems) αξιοποιούνταν καθώς και τις αξίες με τις οποίες αυτά είχαν τροφοδοτηθεί.

Μελετώντας λοιπόν, αφενός μεν τις μεροληψίες εντός αυτών των συστημάτων, αφετέρου δε την εσωτερική κατανομή της εξουσίας, τις πολιτικές και τις οικονομίες που αναπτύσσονταν γύρω από αυτές (Pasquale, 2016) καθώς και τις επιπτώσεις των τεχνολογιών αυτών, το πεδίο των CAS ενσωμάτωσε με την πάροδο του χρόνου στην μελέτη του διάφορες προσεγγίσεις. Συγκεκριμένα, φεμινιστικές προσεγγίσεις (D'Ignazio & Klein, 2020; Cirillo και άλλοι., 2022), κριτικές φυλετικές θεωρίες (Benjamin, 2016), και ηθικές προσεγγίσεις (Floridi και άλλοι, 2018) αξιοποιούνται από τους μελετητές των κριτικών σπουδών των αλγορίθμων, προκειμένου αυτοί να μελετήσουν τους τρόπους με τους οποίους τα αλγοριθμικά συστήματα και τα συστήματα τεχνητής νοημοσύνης χρησιμοποιούν την εξουσία (Beer, 2017), αναπαράγουν κοινωνικές δομές και ανισότητες (O'Neil, 2016; Eubanks, 2018; Benjamin, 2019; Crawford, 2021).

Ιδιαίτερα ρηξικέλευθη παρουσιάζεται η κριτική προσέγγιση του Φράνκ Πασκάλ (Frank Pasquale) αναφορικά με τον τρόπο αξιοποίησης των προσωπικών δεδομένων και των αλγορίθμων στις σύγχρονες κοινωνίες. Ο Πασκάλ αντιμετωπίζει κριτικά τις εφαρμογές, τον σχεδιασμό, τις πολιτικές, τις οικονομίες των δεδομένων καθώς και την λανθάνουσα αξιοποίηση αυτών, μέσω ισχυρών αλγοριθμικών συστημάτων, ο τρόπος λειτουργίας των οποίων είναι ιδιαίτερα δύσκολο να κατανοηθεί. Στο έργο «The Black Box Society: The Secret Algorithms That Control Money and Information» (2016), ο Πασκάλ αντιπαραβάλλει τις σύγχρονες κοινωνίες με μαύρα κουτιά. Αντιλαμβανόμενος την ευρεία εφαρμογή των αλγοριθμικών συστημάτων μέσω της αξιοποίησης των μεγάλων δεδομένων (Brevini & Pasquale, 2020), τονίζει την άγνοια γύρω από τον τρόπο λειτουργίας των αλγορίθμων από την κοινωνία. Η συστηματική αξιοποίησή τους από θεσμούς και οργανισμούς, οι οποίοι ξεπερνούν τα όρια του έθνους-κράτους, σε συνάρτηση με την αδιαφάνεια και την κλειστότητά τους, συγκροτούν μια μορφή διακινδύνευσης, η οποία πλέον έχει παγκόσμιο χαρακτήρα.

B.3. Τεχνητή Νοημοσύνη και Μεροληψίες

Στην συγκεκριμένη ενότητα, παρουσιάζονται τα κύρια είδη μεροληψιών που θα εξεταστούν στην εργασία, ήτοι οι τύποι μεροληψιών που προκύπτουν από την χρήση συστημάτων μηχανικής μάθησης, όπως αυτά χρησιμοποιούνται πολλές φορές από συστήματα τεχνητής νοημοσύνης. Στη συνέχεια, θα αναλυθούν οι μεροληψίες της τεχνητής νοημοσύνης στα πλαίσια των κριτικών σπουδών των αλγορίθμων (CAS). Προκειμένου να διερευνηθεί η σχέση των συστημάτων τεχνητής νοημοσύνης με τις κοινωνικές μεροληψίες, θα πρέπει αρχικώς να αναλυθεί ο όρος «μεροληψία» εντός των κοινωνιών, καθώς συνιστά μια κοινωνιογλωσσική κατασκευή, η οποία προεκτείνεται εντός των κοινωνικοτεχνικών συστημάτων (Airoldi,2021).

B.3.1. Η φύση των μεροληψιών

Η μελέτη των μεροληψιών έχει αποτελέσει αντικείμενο μελέτης διαφόρων φιλοσοφικών και επιστημονικών πεδίων, όπως η ψυχολογία, η εθνογραφία, η κοινωνιολογία, η πολιτική και νομική επιστήμη (Ntoutsι και άλλοι, 2020). Η πολυπλοκότητα των μεροληψιών γίνεται κατανοητή μέσω των διαφορετικών εξετάσεων και ορισμών που αυτή λαμβάνει από τα διαφορετικά επιστημονικά πεδία που την μελετούν. Οι Ntoutsι και άλλοι, (2020), λαμβάνοντας υπόψιν την διεπιστημονική μελέτη των μεροληψιών, τις ορίζουν ως «προδιάθεση ή προκατάληψη μιας απόφασης που λαμβάνεται από ένα σύστημα τεχνητής νοημοσύνης που είναι υπέρ ή κατά ενός ατόμου ή μιας ομάδας, ιδίως με τρόπο που θεωρείται άδικος»

Ωστόσο, η ύπαρξη προκαταλήψεων και μεροληψιών εντός των κοινωνιών δεν μπορεί να νοηθεί δίχως την ύπαρξη εννοιών όπως η ισότητα ή η δικαιοσύνη, καθώς αυτές οριοθετούν την κοινωνική πραγματικότητα αποτελώντας σημεία και εννοιολογικούς πόλους εντός των εκάστοτε κοινωνικών συμφραζόμενων. Συγκεκριμένα, οι Selbst και άλλοι, ορμώμενοι από την επιρροή των STS, αναφέρονται στη «μηχανική μάθηση δίκαιης συνείδησης» (Fairness-aware Machine Learning) και παραθέτουν ότι «εντός του αλγοριθμικού πλαισίου, οποιαδήποτε έννοια της “δικαιοσύνης” δεν μπορεί καν να οριστεί.», (2019, σ.60). Εφόσον λοιπόν, όπως αναλύθηκε και προηγουμένως, τα μοντέλα μηχανικής μάθησης και τα αλγοριθμικά συστήματα είναι κοινωνικοτεχνικά συστήματα, οι έννοιες όπως μεροληψίες και

δικαιοσύνη συνιστούν κοινωνικές κατασκευές, οι οποίες πρέπει να ορίζονται στα εκάστοτε κοινωνικά πλαίσια εντός των οποίων αυτές μελετώνται.

Κοινωνιολογικές θεωρίες, όπως η θεωρία της «κοινωνικής κατασκευής της τεχνολογίας» (SCOT) ή όπως οι κριτικές προσεγγίσεις που εκφράζονται και από το προαναφερθέν πεδίο των «κριτικών σπουδών των αλγορίθμων» (CAS), παρέχουν την δυνατότητα ανάλυσης των μεροληπιών υπό όρους κοινωνικής διαμόρφωσης των τεχνοεπιστημονικών συστημάτων. Κάτι τέτοιο συμβαίνει διότι, η θεωρία SCOT αντιλαμβάνεται την διαμόρφωση του τεχνολογικού ως μια δυναμική διαδικασία, όπου διάφορες κοινωνικές ομάδες αποτελούν ενεργό διαμορφωτή της τεχνολογικής δημιουργίας, ακόμα και μέσω της απουσίας τους (Bijker και άλλοι, 1989).

Από την άλλη πλευρά, οι κριτικές προσεγγίσεις των αλγορίθμων μέσω των CAS διεξάγουν διάλογο γύρω από την αμφίδρομη κατασκευή των τεχνοεπιστημονικών και κοινωνικοτεχνικών κατασκευασμάτων. Ασκώντας κριτική στις μεθόδους, οι οποίες χρησιμοποιούνται από θεσμούς και οργανισμούς για την επιβολή εξουσίας, εξετάζουν τα είδη των μεροληπιών, την εξουσία η οποία επιβάλλεται μέσα από αυτές καθώς και τα αποτελέσματα της εφαρμογής της. Η παρούσα εργασία, όσον αφορά τα συστήματα τεχνητής νοημοσύνης, θα αξιοποιήσει κατά κύριο λόγο προσεγγίσεις από το πεδίο των CAS.

B.3.2 Τύποι μεροληπτικών διακρίσεων μέσω της μηχανικής μάθησης

Οι κοινωνικές διακρίσεις και οι μεροληπίες συνιστούν κοινωνικές πρακτικές, οι οποίες δεν μπορούν να θεωρηθούν και να εξεταστούν εκτός των κοινωνικών πλαισίων αναφοράς τους. Οι καταβολές τους είναι όσο παλιές είναι και οι ανθρώπινες κοινωνίες, ενώ παράλληλα προσαρμόζονται και διαμορφώνουν την κοινωνική πραγματικότητα που τις περιβάλλει. Όντας αναπόσπαστο τμήμα της κοινωνικής πραγματικότητας, οι μεροληπίες και οι διακρίσεις μεταφέρονται εντός των κοινωνικοτεχνικών συστημάτων. Συνεπώς, οι τύποι των κοινωνικών διακρίσεων λαμβάνουν πολλές μορφές, ανάλογες της κοινωνικής πραγματικότητας. Η συγκεκριμένη εργασία εστιάζει στις φυλετικές μεροληπίες (Racial Biases), στις μεροληπίες σε σχέση με το φύλο (Gender Biases) και στις εθνοτικές μεροληπίες (Ethnic Biases).

Racial Biases: Οι φυλετικές μεροληπίες ή αλλιώς «racial biases», ασχολούνται με μορφές διακρίσεων που πραγματοποιούνται σχετικά με τα φυλετικά χαρακτηριστικά, το χρώμα των ανθρώπων. Η μελέτη των μεροληπτικών διακρίσεων

αναφορικά με τα φυλετικά χαρακτηριστικά έχει απασχολήσει σε σημαντικό βαθμό επιστήμονες που ασχολούνται αμιγώς με το πεδίο της Μηχανικής Μάθησης (Suresh & Guttag, 2021), αλλά και επιστήμονες από τον χώρο των CAS (Benjamin, 2016, 2019; O’Neil, 2016; Gebru, 2020; Crawford, 2021). Η μεταβλητή του χρώματος φαίνεται να επηρεάζει σε σημαντικό βαθμό τις αποφάσεις των τεχνολογιών τεχνητής νοημοσύνης.

Η Benjamin, στο έργο της «Race After Technology» (2019), αναφέρεται σε πληθώρα λογισμικών τεχνητής νοημοσύνης τα οποία, χρησιμοποιώντας τεχνολογίες αναγνώρισης προσώπου, αδυνατούν να ανιχνεύσουν τις σκουρόχρωμες επιδερμίδες, πολλές φορές εξαιτίας της υπο-εκπροσώπισης μαύρων πληθυσμών στις βάσεις δεδομένων που αξιοποιούνταν (Noble, 2018; Benjamin, 2019; Raji και άλλοι, 2020). Ακόμα, ιδιαίτερο ενδιαφέρον παρουσιάζουν οι αναλύσεις σχετικά με τις φυλετικές μεροληπτικές διακρίσεις στα συστήματα λήψης αποφάσεων (Decision Making Systems), ιδίως στα συστήματα αξιολόγησης ή ανίχνευσης κινδύνου (Risk Assessment/Risk Detection Systems).

Gender Biases: Οι μεροληψίες σε σχέση με το κοινωνικά προσδιορισμένο φύλο θεωρούνται «η τάση να δίνεται προτιμησιακή μεταχείριση σε ένα φύλο έναντι ενός άλλου» (Reiners, 2021). Σε αυτό το σημείο, είναι σημαντικός ο διαχωρισμός ανάμεσα στις έννοιες του κοινωνικού φύλου (gender) αλλά και στο βιολογικό φύλο (sex). Οι δύο αυτές έννοιες συνήθως ταυτίζονται, όχι μόνο στα πλαίσια μεροληψιών, αλλά και ευρύτερα στην κοινωνία, συγκροτώντας ένα κλίμα αστάθειας, μερικές φορές και πόλωσης, στον χώρο των κοινωνικών ανισοτήτων, στον οποίο υπάγονται και οι μεροληψίες.

Η Lindey διαχωρίζει τους δύο όρους, ορίζοντας το βιολογικό φύλο ως «τα βιολογικά χαρακτηριστικά, τα οποία διαχωρίζουν το αρσενικό από το θηλυκό. Ο ορισμός αυτός δίνει έμφαση στα αρσενικά και θηλυκά βιολογικά χαρακτηριστικά των χρωμοσωμάτων, της ανατομίας, των ορμονών του αναπαραγωγικού συστήματος και στην φυσιολογία» (2021, σ.6). Από την άλλη πλευρά, ορίζει το κοινωνικό φύλο ως «τις κοινωνικές, πολιτισμικές και ψυχολογικές διαφορές ανάμεσα στα αρσενικά και θηλυκά, οι οποίες συνδέονται με αρρενωπότητες και θηλυκότητες» (2021, σ.6).

Οι ανισοτικές μεροληψίες λόγω φύλου, που αναπαράγονται μέσω των συστημάτων τεχνητή νοημοσύνης, έχουν αποτελέσει αντικείμενο μελέτης πολλών μελετητών των CAS. Μεταξύ των πολλών προσεγγίσεων, ιδιαίτερο ενδιαφέρον παρουσιάζουν το έργο των Buolamwini και Gebru (2018), «Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification». Σε

αυτό εξετάζουν τους τρόπους με τους οποίους τεχνολογίες αναγνώρισης προσώπου μεροληπτούν υπέρ των λευκών ανδρών, έναντι των μαύρων γυναικών, με έντονο το στοιχείο της εμφάνισης των φυλετικών μεροληπτικών διακρίσεων. Επιπλέον, η Eubanks στο έργο της «Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor» (2018), εξετάζει περιπτώσεις λογισμικών τεχνητής νοημοσύνης, όπως αυτό στην πολιτεία της Ιντιάνα, το οποίο λειτουργούσε μεροληπτικά κατά των γυναικών, ιδίως των γυναικών-μονογονέων.

Ethnic Biases: Οι εθνοτικές μεροληψίες μπορούν να οριστούν ως «η εμπλοκή σε διακριτική συμπεριφορά, η διατήρηση αρνητικών στάσεων απέναντι σε αυτούς ή η ύπαρξη λιγότερο ευνοϊκών αντιδράσεων προς τους ανθρώπους με βάση την εθνικότητά τους» (Blum και άλλοι, 2021, σ.385). Από την βιβλιογραφία προκύπτει ότι οι εθνοτικές μεροληψίες αποτελούν μια υποκατηγορία των φυλετικών μεροληψιών (Silva & Kenney ,2018.). Η O'neil, στο έργο της «Weapons of Math Destruction» (2016), ασχολείται με αλγοριθμικά μοντέλα ανάλυσης λόγου και τους τρόπους με τους οποίους αυτά αξιοποιούνται στα πλαίσια εθνοτικών διακρίσεων, χωρίς ωστόσο να περιορίζει την πραγμάτευσή της μόνο σε εθνοτικές μεροληψίες.

B.3.3 Κριτική Προσέγγιση των μεροληψιών της Τεχνητής Νοημοσύνης

Οι μεροληψίες της τεχνητής νοημοσύνης έχουν απασχολήσει αρκετούς επιστήμονες λόγω της αυξημένης εμφάνισής τους στην καθημερινή ζωή των πολιτών και της αξιοποίησής τους από ιδιωτικούς και δημόσιους οργανισμούς. Η εκτεταμένη χρήση τους οδηγεί στην διόγκωση και την αναπαραγωγή στερεοτύπων και κοινωνικών ανισοτήτων. Η συγκρότηση και αναπαραγωγή μιας μεροληψίας εντός αυτών των συστημάτων συνιστά μια περίπλοκη διαδικασία. Αυτή δεν περιορίζεται μονάχα εντός των κοινωνικοτεχνικών συστημάτων τεχνητής νοημοσύνης, καθώς σημαντικό τμήμα των μεροληψιών συγκροτείται ήδη από την διαδικασία παραγωγής και συλλογής των δεδομένων. Οι κοινωνίες των μεγάλων δεδομένων, δεν μπορούν να απαλλαγθούν της μεροληπτικής φύσης της ίδιας της κοινωνίας. Το ίδιο ισχύει και για τα κοινωνικοτεχνικά συστήματα τα οποία επεξεργάζονται τα δεδομένα αυτά.

Ενώ λοιπόν, η κοινωνική πραγματικότητα διέπεται εκτός των άλλων και από το στοιχείο της μεροληψίας, η αναπαραγωγή στερεοτύπων και οι μεροληπτικές διακρίσεις συνιστούν πρακτικές οι οποίες δεν είναι επιθυμητές από τις σύγχρονες δημοκρατίες

καθώς «στόχος της κριτικής θεωρίας είναι να εξετάσει και να κατανοήσει τα συστήματα κυριαρχίας και καταπίεσης και να αναζητήσει τρόπους αύξησης και απελευθέρωσης της ελευθερίας» (Verdegem,2021,σ.11). Εντός αυτών των πλαισίων, βασίζεται και ο διάλογος που αναπτύσσεται από τις κριτικές κοινωνιολογικές προσεγγίσεις και ιδίως από τον χώρο των κριτικών σπουδών των αλγορίθμων, CAS.

Οι CAS συμμετέχουν ενεργά στην μελέτη των διακινδυνεύσεων και των ευκαιριών που παράγονται από την χρήση συστημάτων τεχνητής νοημοσύνης. Η πληθώρα των κριτικών πραγματεύσεων, παρά την ενδεχόμενη διαφορά στους τρόπους, ή τον τρόπο, εξέτασης των αλγοριθμικών συστημάτων και των συστημάτων τεχνητής νοημοσύνης, επικεντρώνονται σε ένα σημαντικό σημείο. Αυτό είναι η έλλειψη ουδετερότητας των συστημάτων τεχνητής νοημοσύνης (O'neil, 2016; Pasquale 2016; Eubanks, 2018; Benjamin, 2019; Verdegem, 2019; Gebru, 2020; Crawford, 2021)

Τα συστήματα τεχνητής νοημοσύνης, όπως και οι κοινωνικοί δράστες οι οποίοι τα προγραμματίζουν, λαμβάνουν αποφάσεις και δρουν επηρεαζόμενοι από τα περιβάλλοντά τους (Latour & Woolgar, 1986). Ο τρόπος διαμόρφωσης των κοινωνικών περιβαλλόντων επηρεάζει σημαντικά και τον τρόπο ανάπτυξης και δράσης των τεχνοεπιστημονικών συστημάτων. Η ουδετερότητα και αμεροληψία των συστημάτων τεχνητής νοημοσύνης αποτελούν έννοιες, οι οποίες πρέπει να διερευνηθούν υπό το πλαίσιο των συμφερόντων και της ρητορικής (Latour & Woolgar, 1986; Sismondo, 2016).

Ο χαρακτηρισμός και η αξιοποίηση των συστημάτων αυτών, κάτω από το προσωπείο της ουδετερότητας, είναι ένας από τους λόγους που τα συστήματα αυτά διενεργούν μεροληπτικά. Ο Πασκάλ τοποθετείται στο συγκεκριμένο ζήτημα λέγοντας ότι, «παρόλο που οι γίγαντες του διαδικτύου λένε ότι οι αλγόριθμοί τους είναι επιστημονικά και ουδέτερα εργαλεία, είναι πολύ δύσκολο να επαληθευτούν αυτοί οι ισχυρισμοί» (2016, σ.14). Σε αυτή την περίπτωση, σημασία έχει αφενός μεν ο τρόπος, το καθεστώς κάτω από το οποίο τα λογισμικά αυτά παρουσιάζονται (Benjamin, 2019), αφετέρου δε το ποιος και υπό ποιες συνθήκες ορίζει τις δίκαιες και ουδέτερες τιμές των μελετών στήριξης τέτοιων εγχειρημάτων (Beck, 2015, σσ. 121-130).

Η μάχη για την ουδετερότητα συνιστά μια μορφή ανταγωνισμού στα πλαίσια της πολιτικής και οικονομίας της γνώσης. Για την κατανόηση της, ο όρος του συμφέροντός ως μια ορθολογική επιλογή δεν είναι αρκετός. Ιδιαίτερη σημασία έχει ο τρόπος επιβολής της εξουσίας μέσω της εργαλειοποίησης του επιστημονικού και του ουδέτερου. Όπως επισημαίνει ο Verdegem, «Η εξουσία αποφασίζει ποιος θα

επωφεληθεί από τις νέες τεχνολογίες, όπως η τεχνητή νοημοσύνη, αλλά η συγκέντρωση δύναμης θα οδηγήσει πιθανότατα σε αυξανόμενες ανισότητες και άλλα αρνητικά αποτελέσματα», (2021, σ.2). Αυτές οι ανισότητες, μέσω των ουδετεροποιημένων πλατφορμών αναζήτησης, εξετάζονται από την Noble (2018) στα πλαίσια των κυβερνοτόπων (Cybertopias). Η Noble θεωρεί ότι οι συζητήσεις περί αμεροληψίας από τους κολοσσούς διαχείρισης πληροφοριών, συνιστούν «ιδεολογίες που χρησιμεύουν στη σταθεροποίηση και ομαλοποίηση της έννοιας της εμπορικής αναζήτησης, συμπεριλαμβανομένων των ακόμα δημοφιλών και πάντα επίμονων κυρίαρχων αφηγήσεων για την ουδετερότητα και την αντικειμενικότητα του ίδιου του Διαδικτύου» (σ.61).

Η εκροή μεροληπτικών αποτελεσμάτων από τα κοινωνικοτεχνικά συστήματα τεχνητής νοημοσύνης, αλγορίθμων, και μηχανικής μάθησης, ξεκινάει πολλές φορές πριν την πραγματοποίηση κάποιας αλγοριθμικής διαδικασίας. Διεργασίες, όπως η συλλογή των δεδομένων (data collection) και η επεξεργασία τους (data processing), εμπεριέχουν ήδη μεροληπτικές διακρίσεις, οι οποίες με την σειρά τους έχουν ήδη διαμορφωθεί και συλλεχθεί από τον εξωτερικό κόσμο. Μια ακόμα περίπτωση κατά την οποία οι αλγόριθμοι μπορεί να λειτουργήσουν προβληματικά, όσον αφορά τις μεροληψίες, είναι οι μεροληψίες εκπροσώπησης (Representation Biases), (Suresh & Guttag, 2021). Πολλοί πληθυσμοί υποεκπροσωπούνται ή εκπροσωπούνται λανθασμένα εντός των βάσεων δεδομένων.

Χαρακτηριστικό παράδειγμα υποεκπροσώπησης αποτελεί το μοντέλο πρόσληψης της εταιρίας Amazon (Dastin, 2018; Gebru, 2020; Crawford, 2021). Στο συγκεκριμένο μοντέλο, ο αλγόριθμος σχετικά με τις προσλήψεις εκπαιδεύτηκε σε δεδομένα στα οποία εμφανιζόταν υπερεκπροσώπηση από τον ανδρικό πληθυσμό (Dastin, 2018; Gebru, 2020), ενώ παράλληλα το μοντέλο «ευνοούσε τους υποψήφιους που περιέγραφαν τον εαυτό τους χρησιμοποιώντας ρήματα που απαντώνται συχνότερα στα βιογραφικά των ανδρών», πριμοδοτώντας με αυτόν τον τρόπο αρρενωπό λεξιλόγιο (Dastin, 2018). Αποτέλεσμα της αξιοποίησης του αλγορίθμου ήταν η εταιρία να εκπροσωπείται κατά κύριο λόγο από άνδρες σε θέσεις όπως μηχανικοί (Dastin, 2018; Crawford, 2021), οδηγώντας σε έναν κύκλο ανατροφοδότησης μεροληπτικών πληροφοριών (Feedback Loop), που βασίζονται στις μεροληψίες σε σχέση με το φύλο.

Το ζήτημα της υποεκπροσώπησης δεν αποτελεί το μόνο προβληματικό σημείο των μεροληπτικών αλγοριθμικών συστημάτων. Πολλές φορές, τα τεχνοεπιστημονικά αυτά εργαλεία εντείνουν τις διακρίσεις μέσω συστημάτων ταξινόμησης (Classification

Systems). Ο τρόπος με τον οποίο ένα σύστημα μηχανική μάθησης ταξινομεί τα δεδομένα και τις πληροφορίες που του παρέχονται, μπορεί να επηρεάσει σε τεράστιο βαθμό την πρόβλεψη ή την απόφαση ενός συστήματος τεχνητής νοημοσύνης. Ωστόσο, η μεροληψία των κοινωνικοτεχνικών συστημάτων επηρεάζει σαφώς και την διαδικασία ταξινόμησης των δεδομένων. Συνεπώς, το έκβασμα μιας τέτοιας διαδικασίας είναι εκτεθειμένο στις πολιτικές, οικονομικές σχέσεις και καταβολές των συστημάτων τεχνητής νοημοσύνης (Verdegem, 2021).

Τα συστήματα ταξινόμησης αποτελούν ταυτόχρονα συστήματα ταξινόμησης και απόδοσης εξουσίας. Αποτελούν την μέθοδο ελέγχου του μέλλοντος, ελέγχου των αποτελεσμάτων. Οι μεροληψίες εντός αυτών των συστημάτων περιθωριοποιούν ορισμένες κοινωνικές ομάδες και ταυτόχρονα ενισχύουν άλλες, μέσω του περιορισμού των κοινωνικά αδύναμων, των περιθωριοποιημένων και υποεκπροσωπούμενων κοινωνικών ομάδων και υποκειμένων. Κάτι τέτοιο συμβαίνει καθώς, «αυτοί που έχουν τη δύναμη να σχεδιάζουν συστήματα - ταξινόμησης ή τεχνικά συστήματα - έχουν την ικανότητα να δίνουν προτεραιότητα σε ιεραρχικά σχήματα που προκρίνουν ορισμένους τύπους πληροφοριών έναντι των άλλων» (Noble, 2018, σσ.138-139).

Επηρεασμένη από τον χώρο των κριτικών σπουδών και κυρίως τον χώρο των φυλετικών κριτικών σπουδών και των φυλετικών σπουδών του φύλου, η Noble ορίζει τα συστήματα ταξινόμησης ως «μέρος της επιστημονικής προσέγγισης για την κατανόηση των ανθρώπων και των κοινωνιών, και κατέχουν τις προκαταλήψεις εξουσίας εκείνων που είναι σε θέση να προπαγανδίζουν τέτοια συστήματα» (2018,σ.137). Η επιστημονικότητά τους αποτελεί εργαλείο ενίσχυσης της θέσης τους και ταυτόχρονα ισχυρό διαμορφωτή μιας μεροληπτικής κοινωνίας, ανοιχτής στην αναπαραγωγή φυλετικών, εθνικών, κοινωνικοοικονομικών διακρίσεων.

Η ισχύς που λαμβάνει η κωδικοποίηση των προγραμμάτων τεχνητής νοημοσύνης την σημερινή εποχή, έχει απασχολήσει ιδιαίτερα επιστήμονες όπως η Benjamin (2019), η οποία χαρακτηρίζει τις τεχνολογίες προγραμματισμού «τον νέο κώδικα Jim», δηλαδή, ως νέες τεχνολογίες που «αντανακλούν και αναπαράγουν τις υφιστάμενες ανισότητες, οι οποίες όμως προωθούνται και θεωρούνται πιο αντικειμενικές ή προοδευτικές από τα συστήματα διακρίσεων μιας προηγούμενης εποχής». Από την άλλη πλευρά, η O'neil αντιλαμβάνεται τα αλγοριθμικά συστήματα ως «Όπλα Μαθηματικής Καταστροφής» (Weapons of Math Destruction). Χαρακτηριστικό τους αποτελεί το γεγονός ότι εντός αυτών εμπεριέχονται «κωδικοποιημένες ανθρώπινες προκαταλήψεις, παρεξηγήσεις και προκαταλήψεις στα

συστήματα λογισμικού που διαχειρίζονται όλο και περισσότερο τις ζωές μας» (Oneil, 2016).

Τα συστήματα μηχανικής μάθησης, πέρα από την ικανότητα της δύναμης, έχουν και την ικανότητα της κατασκευής. Η ταξινόμηση που προγραμματίζεται εντός αυτών, παράλληλα διαμορφώνει την κοινωνική πραγματικότητα και τις κατασκευές της, μέσω του τρόπου που τις τοποθετεί ιεραρχικά και μέσω των χαρακτηριστικών που τους αναγνωρίζει. Σύμφωνα με την Crawford (2021), οι αλγόριθμοι ταξινόμησης «καθορίζουν τον κόσμο μέσα από τους όρους που έχουν θέσει, και αυτό έχει μακροχρόνιες συνέπειες για τους ανθρώπους που ταξινομούνται». Η κωδικοποίηση των συστημάτων αυτών δεν συνιστά απλώς μια μορφή επικοινωνίας, αλλά εργαλείο οριοθέτησης της πραγματικότητας. Αρκεί κανείς να αναλογιστεί τους τομείς στους οποίους χρησιμοποιούνται τέτοιες τεχνολογίες και την εξάπλωση της τεχνητής νοημοσύνης σήμερα, για να κατανοήσει την σημαντική επιρροή των τεχνολογιών τεχνητής νοημοσύνης στη δημιουργία ταυτοτήτων και στον αποκλεισμό κοινωνικό αποκλεισμό ομάδων.

Ο κώδικας ως μέσο επικοινωνίας στα πληροφοριακά συστήματα και τα συστήματα τεχνητής νοημοσύνης, οδηγεί στην συγκρότηση κύκλων αναπαραγωγής μεροληψιών (FeedBack Loops) (Crawford, 2021). Όπως όλα τα τεχνοεπιστημονικά συστήματα, έτσι και τα συστήματα τεχνητής νοημοσύνης αναπαραγάγουν την κοινωνική πραγματικότητα, η οποία τα περιβάλλει (Latour & Woolgar, 1986). Αν αυτή η πραγματικότητα διέπεται από μεροληψίες, τότε αυτές εκφράζονται τόσο στα δεδομένα αλλά και στα συστήματα ταξινόμησης. Ωστόσο, τα συστήματα τεχνητής νοημοσύνης δεν περιορίζονται μονάχα στην είσοδο και επεξεργασία μεροληπτικών πληροφοριών εντός αυτών. Τα αποτελέσματα, οι έξοδοι, αυτών των συστημάτων δεν αναπαράγουν απλώς τις ανισότητες και τις στερεοτυπικές αντιλήψεις του κώδικά, αλλά τις εντείνουν καθώς «οι προκαταλήψεις και οι πολιτισμικές κλίσεις που προέρχονται από τον κοινωνικό κόσμο εξαπλώνονται τεχνολογικά με αυτοματοποιημένα συστήματα, σε αντίθεση με την ευρέως υποτιθέμενη ουδετερότητά τους» (Airoldi, 2021, σ.35).

Όλες οι παραπάνω τοποθετήσεις έχουν ως αποτέλεσμα την συγκρότηση ρατσιστικών, μεροληπτικών ως προς το φύλο και ξενοφοβικών αλγορίθμων (O'neil, 2016; Noble, 2018; Benjamin, 2019). Η εκτεταμένη χρήση τους στον ιδιωτικό και δημόσιο τομέα, παρά την αποδοτικότητα και την διευκόλυνση πολλών πολιτών στις διάφορες πτυχές της καθημερινότητάς τους, εγείρει προβληματισμούς. Τα δεδομένα

που αξιοποιούνται στα τεχνοεπιστημονικά συστήματα τεχνητής νοημοσύνης, πολλές φορές οδηγούν σε πρακτικές αποκλεισμού και διακρίσεων αφού «μόλις ένα μοντέλο εκπαιδευτεί σε αυτού του είδους τα δεδομένα, επιδεινώνει τα υπάρχοντα κοινωνικά ζητήματα οδηγώντας σε περαιτέρω περιθωριοποίηση» (Gebru, 2020). Υπό το πρίσμα αυτό, η τεχνητή νοημοσύνη συνιστά μια διακινδύνευση (Beck, 2015), ένα μαύρο κουτί (Pasquale, 2016), το οποίο πρέπει να διανοιχθεί κάτω από όρους κριτικής και διαφάνειας.

B.4.Οι διακινδυνεύσεις της Τεχνητής Νοημοσύνης στις Κοινωνίες του Μαύρου Κουτιού.

Οι σύγχρονες κοινωνίες βιώνουν ραγδαίες μορφές τεχνολογικής ανάπτυξης, οι οποίες μεταβάλλουν ριζικά τον τρόπο που αυτές δομούνται στο σύνολό τους. Η επανάσταση των μεγάλων δεδομένων, σε συνάρτηση με την ανάπτυξη υποδομών και τεχνολογιών τεχνητής νοημοσύνης, επιφέρουν σημαντικές αλλαγές στην λειτουργία των σύγχρονων κοινωνιών. Οι μοντέρνες, νεωτερικές κοινωνίες, όπως αυτές διαμορφώθηκαν χάρη στις βιομηχανικές επαναστάσεις, έρχονται αντιμέτωπες με μια νέα συνθήκη. Τα κοινωνικοτεχνικά συστήματα της τεχνητής νοημοσύνης έχουν εισχωρήσει σε όλα τα στάδια της κοινωνικής και προσωπικής ζωής των πολιτών. Οι μεταμοντέρνες κοινωνίες της διακινδύνευσης, (Beck, 2015) με την χρήση των προαναφερθέντων τεχνολογιών, αποτελούν κοινωνίες του μαύρου κουτιού (Pasquale, 2016), όπου τα ανθρώπινα ψηφιακά δεδομένα και οι αλγοριθμικές διαδικασίες μηχανικής μάθησης λειτουργούν μεροληπτικά εν αγνοία συγκεκριμένων ομάδων των πολιτών. Προτού λοιπόν πραγματοποιηθεί η περιγραφή των διακινδυνεύσεων της τεχνητής νοημοσύνης στις κοινωνίες του μαύρου κουτιού, καθίσταται η αναγκαία η απομόνωση και παρουσίαση των καταβολών της έννοιας του «μαύρου κουτιού», η οποία λαμβάνει κεντρικό ρόλο στις πραγματεύσεις των κριτικών σπουδών των αλγορίθμων (CAS).

Ακόμα και αν στην πραγματεία του για τις Κοινωνίες της Διακινδύνευσης (2015) ο Μπεκ δεν αναφέρεται στα συστήματα τεχνητής νοημοσύνης, τοποθετείται σχετικά με τις διακινδυνεύσεις τις οποίες παράγουν τα τεχνοεπιστημονικά συστήματα. Αυτά, δεν αποτελούν εξωτερικές διακινδυνεύσεις όπως αυτές του περιβάλλοντος, αλλά περισσότερο εσωτερικές, διακινδυνεύσεις οι οποίες παράγονται κοινωνικά και επιστημονικά (2015, σσ.269-274). Οι θέσεις του Πασκάλ σχετικά με τις Κοινωνίες του

Μαύρου Κουτιού και τις λανθάνουσες παρενέργειες που έχουν αυτές στους πολίτες αναδεικνύουν το ζήτημα των μη ουδέτερων αποτελεσμάτων συνεπώς και των μη ουδέτερων επιστημονικών τεκμηριώσεων οι οποίες τις συνοδεύουν και νομιμοποιούν την χρήση τους. Σε αυτά τα πλαίσια, «Η κυρίαρχη θεωρητική έννοια της επιστήμης υπονοεί ότι οι επιστήμες...παραδίδουν υποτιθέμενους “ουδέτερους” αριθμούς και πληροφορίες ή εξηγήσεις που λειτουργούν ως η αμερόληπτη βάση για αποφάσεις δε μια ευρύτατη ποικιλία συμφερόντων» (Beck, 2015, σ. 302). Τα συμφέροντα των εταιριών μέσω των τεχνοεπιστημονικών εργαλείων των εταιριών αποτελούν αντικείμενο της πραγματείας του Πασκάλ ενώ οι τρόποι με τους οποίους αυτά παρουσιάζονται κάτω από τον μανδύα της αντικειμενικότητας της επιστήμης αναλύονται στην πραγματεία των διακινδυνεύσεων του Μπέκ (2015).

B.4.1. Οι καταβολές της έννοιας του μαύρου κουτιού και η αξιοποίηση της από το πεδίο των STS

Η προσέγγιση κοινωνικών διαδικασιών με τον όρο του μαύρου κουτιού, αποτελεί κατά κύριο λόγο απόκτημα που λήφθηκε από τις σπουδές της επιστήμης και της τεχνολογίας, στα πλαίσια ανάγνωσης των κοινωνικοτεχνικών συστημάτων. Οι STS ενίσχυσαν σημαντικά τον διάλογο γύρω από τους τρόπους με τους οποίους η επιστημονική γνώση και η τεχνολογία πρέπει να εξεταστούν ως μαύρα κουτιά, δηλαδή σαν κοινωνικές κατασκευές των οποίων οι εσωτερικές διαδικασίες δεν πρέπει να αντιλαμβάνονται ως δεδομένες. Τόσο το Ισχυρό Πρόγραμμα της Κοινωνιολογίας της Επιστημονικής Γνώσης, η θεωρία κοινωνικής κατασκευής της τεχνολογίας (SCOT), όσο και οι Θεωρίες Μελέτης Εργαστηρίου και η Θεωρία Δρώντων Δικτύου, ANT αξιοποιούν στις πραγματείες τους τον όρο του «μαύρου κουτιού» για την περιγραφή του τεχνοεπιστημονικού φαινομένου.

Η έννοια του μαύρου κουτιού είχε χρησιμοποιηθεί πρώτη φορά στα πλαίσια των STS από τον Whitley. Σύμφωνα με τον Pinch, ο Whitley χρησιμοποίησε αυτόν τον όρο ασκώντας κριτική στη Μερτόνια σκέψη σχετικά με την Κοινωνιολογία της Επιστήμης καθώς «οι Μερτονιανοί μελετούσαν τις εισροές και τις εκροές της επιστήμης, αλλά όχι την εσωτερική της λειτουργία - και, κυρίως, τις διαδικασίες με τις οποίες παράγεται η επιστημονική γνώση» (Pinch, 1992, σ.488). Το Ισχυρό Πρόγραμμα της Κοινωνιολογίας της Γνώσης, μέσω της συμμετρικής αρχής για την προσέγγιση της επιστημονικής γνώσης, τροφοδότησε την κοινωνιολογική σκέψη, με τον αγνωστικισμό

σχετικά με τον τρόπο λειτουργίας της επιστήμης. Βάσει αυτής, η επιστήμη δεν συνιστά μια μορφή αυθύπαρκτης γνώσης, αλλά αποτελεί μια κοινωνική διαδικασία σχηματισμού, η οποία δεν πρέπει να θεωρείται δεδομένη. Αντιθέτως, οι εσωτερικές τις διεργασίες θα πρέπει να εξετάζονται σαν να είναι ένα μαύρο κουτί.

Η έννοια του μαύρου κουτιού χρησιμοποιήθηκε και από τον Layton στα πλαίσια της κοινωνικής κατασκευής της τεχνολογίας, καθώς «αυτό που χρειάζεται είναι η κατανόηση της τεχνολογίας από μέσα, τόσο ως σώμα γνώσεων όσο και ως κοινωνικό σύστημα. Αντίθετα, η τεχνολογία αντιμετωπίζεται συχνά ως ένα "μαύρο κουτί" του οποίου το περιεχόμενο και η συμπεριφορά μπορεί να θεωρηθεί ότι είναι κοινή γνώση» (Layton 1977, σ. 198'όπως αναφέρεται στους Bijker και άλλοι, 1989, σ.22). Χάρη λοιπόν στην συμμετρική προσέγγιση, η κοινωνιολογική μελέτη τόσο της επιστημονικής γνώσης, όσο και της κοινωνικής κατασκευής της τεχνολογίας, αντιλαμβανόμενες την συχνή θεώρηση των παραπάνω διαδικασιών κατασκευής ως δεδομένες, ανέπτυξαν την αρχή της συμμετρίας, η οποία επιτάσσει την ενδελεχή, συμμετρική εξέταση αιτιών που συγκροτούν τα κοινωνικά φαινόμενα κάτω από το αγνωστικιστικό και μη δεδομένο πρίσμα των μαύρων κουτιών.

Η προσέγγιση της τεχνολογίας και της επιστήμης σαν μια διαδικασία που κατασκευάζεται κοινωνικά, υιοθετήθηκε κάτω από το μονιστικό πλαίσιο της Θεωρίας Δρώντων Δικτύου και των Σπουδών Εργαστηρίου. Ο μονισμός, τον οποίο έθεσε η Θεωρία Δρώντων Δικτύου του Λατούρ και του Γούλγκαρ (1986), εγκαθίδρυσε μια νέα βάση εξέτασης των μαύρων κουτιών κάτω από ένα υβριδικό σχεσιακό περιβάλλον δραστών οι οποίοι δεν είναι απαραίτητα ούτε άνθρωποι ούτε και τεχνολογικά κατασκευάσματα (Latour, 1996). Υπό το πρίσμα αυτό, τα μαύρα κουτιά της τεχνοεπιστήμης θα έπρεπε να διανοίγονται με διαφορετικό τρόπο, καθώς οι δράστες εντός αυτών φέρουν διαφορετικές ιδιότητες, οντολογίες και χρόνους διάνοιξης.

Το επιστημονικό πεδίο των STS διεύρυνε σημαντικά την κοινωνιολογική οπτική σχετικά με τα σύγχρονα κοινωνικοτεχνικά συστήματα τεχνητής νοημοσύνης. Η συζήτηση γύρω από την φύση των μαύρων κουτιών από την δεκαετία του 1970 μέχρι και τις αρχές της δεύτερης χιλιετίας, όπου και ξεκίνησε η αξιοποίηση των αλγοριθμικών συστημάτων, συντέλεσε στην ανάπτυξη πραγματεύσεων γύρω από τον χώρο των κριτικών προσεγγίσεων των μεγάλων δεδομένων και της τεχνητής νοημοσύνης. Εντός αυτών προσεγγίσεων, αναπτύχθηκε και το έργο του Πασκάλ με θέμα τις Κοινωνίες του Μαύρου Κουτιού (2016). Η κριτική προσέγγιση των νέων, «αόρατων», περιβαλλόντων των δεδομένων και των αλγορίθμων παρουσιάζει αρκετές

ομοιότητες με την προσέγγιση του Μπέκ σχετικά με τις κοινωνίες της διακινδύνευσης και τους κινδύνους που συγκροτούνται εντός αυτών.

B.4.2. Η συνύπαρξη των Κοινωνιών της Διακινδύνευσης και των Κοινωνιών του Μαύρου Κουτιού: Μια ανάλυση στα πλαίσια των συστημάτων τεχνητής νοημοσύνης

Η έλευση των μεγάλων δεδομένων και η ταχύτατη πορεία ανάπτυξης και καθιέρωσης των τεχνολογιών της τεχνητής νοημοσύνης αποτελούν συνθήκες, οι οποίες μετέβαλαν και μεταβάλλουν έως σήμερα την δόμηση και την οργάνωση των σύγχρονων κοινωνιών. Τα κοινωνικά υποκείμενα και οι ενέργειές τους ψηφιοποιούνται και λαμβάνουν την μορφή δεδομένων. Τα δεδομένα πλέον αποτελούν την νέα μορφή πετρελαίου (Yonago, 2015; Bhageshrur, 2019). Ωστόσο, ο χώρος συλλογής, ο τρόπος επεξεργασίας αυτών των δεδομένων καθώς και ο φορέας αξιοποίησης τους παραμένουν άγνωστοι, ένα μαύρο κουτί για τις σύγχρονες κοινωνίες, οι οποίες έρχονται αντιμέτωπες με την άγνοια αξιοποίησης των ψηφιακών αποτυπωμάτων τους, ενώ ταυτόχρονα οι κοινωνικοί δράστες θεωρούν αυτήν την πραγματικότητα δεδομένη.

Το φαινόμενο αυτό δεν περιορίζεται γεωγραφικά στον χώρο στον οποίο θεωρείται συνήθως ότι διατηρεί το μονοπώλιο της καινοτομίας και του τεχνολογικού εκσυγχρονισμού, δηλαδή στον Δυτικό κόσμο. Τα δεδομένα των πολιτών δεν περιορίζονται όπως τα σώματά τους. Αντιθέτως, μεταφέρονται και επεξεργάζονται σε βάσεις δεδομένων και κέντρα, τα οποία απέχουν σημαντικά από την άμεση γεωγραφική θέση των κοινωνικών υποκειμένων από τα οποία αντλήθηκαν. Ο τρόπος με τον οποία τα δεδομένα αυτά φιλτράρονται, κατηγοριοποιούνται, χρησιμοποιούνται για εμπορικούς λόγους αλλά και για την επιβολή πολιτικών, συνιστά ένα μαύρο κουτί, μια διακινδύνευση σχετικά με τις μελλοντικές τους χρήσεις (Paquale, 2016; Eubanks, 2018; Benjamin, 2019; Gebru, 2021). Υπό αυτά τα πλαίσια, τα αλγοριθμικά συστήματα μηχανικής μάθησης, τα οποία αξιοποιούνται για τους παραπάνω σκοπούς, αποτελούν μαύρα κουτιά, τεχνολογίες δηλαδή οι οποίες χρίζουν προσεκτικής διάνοιξης, καθώς ενδέχεται, εξαιτίας της αναστοχαστικότητάς τους, να οδηγήσουν στην εμφάνιση περεταίρω διακινδυνεύσεων, από αυτές που κλήθηκαν να αντιμετωπίσουν.

Ο Μπέκ, στην παρουσίαση των (μετα)μοντέρνων κοινωνιών, των κοινωνιών του εκσυγχρονισμού, υπογραμμίζει την έννοια της αναστοχαστικότητας βάσει της οποίας οι ίδιες αναδεικνύουν «στο ίδιο το θέμα και το πρόβλημά τους» (2015, σ.45). Τα προβλήματα των κοινωνιών αυτών, στην κοινωνιολογική προσέγγιση του Μπέκ,

περιγράφονται ως διακινδυνεύσεις. Αυτές ορίζονται ως ο «συστηματικός τρόπος αντιμετώπισης των κινδύνων και των ανασφαλειών που προκαλούνται και εισάγονται από τον ίδιο τον εκσυγχρονισμό» (2015, σ.48). Οι κοινωνίες της διακινδύνευσης χαρακτηρίζονται από την αναπαραγωγική τάση τους σχετικά με την αναπαραγωγή προβλημάτων και κινδύνων. Αυτές συγκροτούνται εσκεμμένα ή άθελα των εκσυγχρονιστικών δυνάμεων της πολιτικής και της οικονομίας, ή καλύτερα υποπολιτικής. Μια από αυτές τις διακινδυνεύσεις, η οποία χαρακτηρίζει τις (μετα)μοντέρνες κοινωνίες της τέταρτης βιομηχανικής επανάστασης» (Ross & Maynard, 2021), είναι τα φαινόμενα του μαύρου κουτιού.

Οι Κοινωνίες του Μαύρου Κουτιού (2016) όπως και οι Κοινωνίες της Διακινδύνευσης (2015) δομούνται γύρω από διάφορα χαρακτηριστικά, όπως είναι η παγκόσμιά τους διάσταση, οι οικονομίες της πολιτικής και της γνώσης, οι λανθάνουσες παρενέργειες που φέρουν ως αποτέλεσμα καθώς και η αναστοχαστική τους διάθεση για την συγκρότηση περεταίρω μαύρων κουτιών συνεπώς και διακινδυνεύσεων. Μια όμως σημαντική ομοιότητα των δύο αυτών θεωρητικών προσεγγίσεων των μεταμοντέρνων κοινωνιών και των κοινωνιών της τεχνητής νοημοσύνης είναι η αορατότητα. Καμία από τις ενέργειες, οι οποίες πραγματοποιούνται και μπορούν να προκαλέσουν βλάβη, δεν είναι άμεσα ορατή, συνεπώς δεν μπορεί να αποτελέσει κίνδυνο έως ότου τα ίδια της αποτελέσματα να φανερωθούν (Beck, 2015).

Σύμφωνα με τον Πασκάλ, χαρακτηριστικό της Κοινωνίας του μαύρου κουτιού αποτελεί το γεγονός ότι αυτή «δημιουργεί αόρατες δυνάμεις» (2016, σ.193). Τόσο τα δεδομένα των πολιτών, όσο και οι αλγόριθμοι, οι οποίοι τα επεξεργάζονται και πραγματοποιούν προβλέψεις βασιζόμενοι σε αυτά, λειτουργούν κάτω από μη ορατά και κατανοητά πλαίσια. Ο αλγόριθμος, το μοντέλο μηχανικής μάθησης, το σύστημα τεχνητής νοημοσύνης και τα αποτελέσματα αυτών των κοινωνικών τεχνημάτων γίνονται αντιληπτά στον βαθμό που ο κόσμος έχει πραγματική γνώση σχετικά με την λειτουργία τους. Σε αντίθεση μάλιστα με ορισμένες φυσικές καταστροφές, οι οποίες μπορούν να γίνουν αισθητά αντιληπτές, οι διακινδυνεύσεις της τεχνητής νοημοσύνης στις Κοινωνίες του Μαύρου Κουτιού γίνονται ορατές μονάχα μέσω της θέασης των αποτελεσμάτων τους.

Στις κοινωνίες των μεγάλων δεδομένων λοιπόν, η οικονομία και η πολιτική περιβάλλονται κατά κύριο λόγο γύρω από την γνώση. Η οικονομία της γνώσης αποτελεί όρο ο οποίος αναφέρεται «σε μια οικονομία που βασίζεται σε πολύ ανεπτυγμένες μορφές γνώσης» (Sismondo, 2016, σ.309). Για την κατανόηση των

λειτουργιών των συστημάτων τεχνητής νοημοσύνης δεν αρκεί απλώς η γνώση της ύπαρξης ενός αλγορίθμου, ο οποίος πραγματοποιεί εντολές όπως αυτός έχει προγραμματιστεί. Στις κοινωνίες στις οποίες η γνώση αποτελεί μια από τις σημαντικότερες μορφές κεφαλαίου, σημαντικοί διαμορφωτές της πραγματικότητας είναι οι ειδήμονες της επιστήμης και της τεχνολογίας (Sismondo,2016), καθώς αυτοί, σε αντίθεση με τις ανθρώπινες αισθήσεις, έχουν πρόσβαση στα «αισθητήρια όργανα της επιστήμης-θεωρίες πειράματα, όργανα μέτρησης» (Beck, 2015, σ.57), διασφαλίζοντας με αυτόν τον τρόπο πέρα από την προσωπική πρόληψη και τον έλεγχο των διακινδυνεύσεων.

Η εμφάνιση των μαύρων κουτιών τεχνητής νοημοσύνης, των μεροληπτικών και αλγοριθμικών διακρίσεων που εμπεριέχονται σε αυτά, είναι ανάλογη των οικονομιών που σχηματίζονται γύρω από αυτά, δηλαδή παγκόσμια. Τα μαύρα κουτιά και η υποπολιτική που τα διέπει, παγκοσμιοποιούνται όπως άλλωστε και όλες οι διακινδυνεύσεις στην μεταμοντέρνα εποχή. Οι ίδιες εταιρίες συλλογής πληροφοριών και μεγάλων δεδομένων μέσω των μηχανών αναζήτησης, όπως η Microsoft και η Google, αυτή την στιγμή πρωτοστατούν στους αγώνες εγκατάστασης μοντέλων τεχνητής νοημοσύνης στην αναζήτησή τους (Milmo, 2023). Ακόμα και αν αυτή η υβριδική μορφή αναζήτησης αξιοποιεί μοντέλα τα οποία βασίζονται πολλές φορές σε μεροληπτικές πληροφορίες, οι εταιρίες που τα παρέχουν επενδύουν πολλαπλά εκατομμύρια στην ανάπτυξή τους, η οποία μάλιστα έχει ήδη αλλάξει την δομή της αναζήτησης πληροφοριών.

Ωστόσο, παραγωγή των μαύρων κουτιών δεν είναι ενιαία, ούτε περιορίζεται στο φυσικό περιβάλλον της γης. Τα μαύρα κουτιά των λογισμικών επεξεργασίας και αξιοποίησης των μεγάλων δεδομένων, εξαιτίας της ψηφιακής τους φύσης, αποθηκεύονται σε μορφές όπως αποθήκες δεδομένων (Data warehouses), σε λίμνες δεδομένων (Data Leaks) και σε τεχνολογίες Cloud (Juillet, 2022). Μάλιστα, οι τεχνολογίες αυτές συμβάλλουν στην αποκέντρωση των δεδομένων από το υποκείμενο και την αποθήκευσή τους σε πολλά ψηφιακά περιβάλλοντα ταυτόχρονα. Η τοποθέτηση σε κατακερματισμένα περιβάλλοντα ρευστοποιεί την έννοια του εαυτού και της προσωπικότητας καθώς, όχι μόνο παρουσιάζονται προσωπικά δεδομένα και εκδοχές του εαυτού σε πολλαπλούς χώρους συγκέντρωσης δεδομένων, αλλά ταυτόχρονα τα δεδομένα αυτά εκτίθενται στην διακινδύνευση, όπου μια από τις ψηφιακά κατακερματισμένες μορφές του εαυτού να δεχτεί κάποια κυβερνοεπίθεση.

Χαρακτηριστικό παράδειγμα αποτελεί η παραβίαση δεδομένων από το Office of Personnel Management (OPM) των Ηνωμένων Πολιτειών της Αμερικής, όπου ευαίσθητα προσωπικά δεδομένα των πολιτών διέρρευσαν εξαιτίας της κυβερνοεπίθεσης, η οποία όπως δεν είχε καμία γεωγραφική σχέση με τις Ηνωμένες Πολιτείες. Αντιθέτως, τα αποκεντρωμένα προσωπικά δεδομένα των πολιτών δέχτηκαν επίθεση στον χώρο που συγκεντρώνονταν στην Κίνα (Wehbé, 2017). Τα μαύρα κουτιά της συγκέντρωσης των προσωπικών δεδομένων τους βιώνουν την συχνή διακινδύνευση της αόρατης έκθεσης του εαυτού.

Η πραγμάτευση του Πασκάλ αναπτύσσεται ακριβώς γύρω από τις οικονομίες γνώσης του εαυτού. Ωστόσο, η ανάλυσή του, δεν περιορίζεται μονάχα στην αποθήκευση των μεγάλων δεδομένων. Η αξιοποίηση των μεγάλων δεδομένων μέσω αλγορίθμων και συστημάτων τεχνητής νοημοσύνης από εταιρίες και κρατικούς οργανισμούς συνιστά ένα πολύπλοκο μαύρο κουτί, το οποίο εξαιτίας της αναστοχαστικής του φύσης εμπεριέχει και άλλα μαύρα κουτιά. Οι αλγόριθμοι αυτοί σύμφωνα με τον Πασκάλ (2016, σ.18) αποτελούν «Τα μαύρα κουτιά της φήμης, της αναζήτησης και της χρηματοδότησης θέτουν σε κίνδυνο όλους μας. Τα λανθασμένα δεδομένα, οι άκυρες υποθέσεις και τα ελαττωματικά μοντέλα δεν μπορούν να διορθωθούν όταν είναι κρυμμένα»

Ένα σημαντικό στοιχείο το οποίο καθιστά αυτά τα αλγοριθμικά συστήματα «αόρατα» στους πολίτες, είναι η επικοινωνία του μηνύματος της ουδετερότητας. Τα συστήματα αυτά παρουσιάζονται από τους επιστημονικούς ειδήμονες ως ουδέτερα και αντικειμενικά (Beck, 2015; O'neil, 2016; Eubanks, 2018; Noble, 2018; Benjamin, 2019; Crawford, 2021; Gebru, 2021). Σύμφωνα με τον Beck το γεγονός αυτό συμβαίνει καθώς στις (μετα)μοντέρνες κοινωνίες της διακινδύνευσης, «διασπάται το μονοπώλιο της ορθολογικότητας των επιστημών» (2015, σ.61). Σχετικά με τον καθορισμό της αντικειμενικότητας, ο Μπέκ αναπτύσσει την θέση του γύρω από τον ίδιο άξονα σκέψης που μελετούν και οι κριτικές σπουδές των αλγορίθμων, δηλαδή την κατανομή της εξουσίας. Για αυτόν, στις κοινωνίες που η γνώση αποτελεί την ισχυρότερη μορφή πρόληψης από τους κινδύνους του εκσυγχρονισμού, η εξουσία συγκεντρώνεται στους τεχνοεπιστημονικούς ειδήμονες, οι οποίοι ασκούν το μονοπώλιο ορισμού των επιτρεπτών τιμών, των αποδεκτών στατιστικών λαθών.

Αναλύοντας τις προαναφερθείσες συνδέσεις, προκύπτει η ανάδειξη των μαύρων κουτιών των μεγάλων δεδομένων, των αλγοριθμικών λογισμικών και των συστημάτων τεχνητής νοημοσύνης. Το γεγονός μάλιστα ότι στις σύγχρονες κοινωνίες

τα συστήματα τεχνητής νοημοσύνης αξιοποιούνται άκριτα παραβιάζοντας πολλές φορές τα νομικά πλαίσια προστασίας των ανθρώπινων δικαιωμάτων και μεροληπτώντας δίχως κάποια σαφή τεκμηρίωση κατά κοινωνικών μειονοτήτων, αναδεικνύει την παράλληλη συνύπαρξη των Κοινωνιών της Διακινδύνευσης και των Κοινωνιών του Μαύρου Κουτιού. Η συνύπαρξή τους δεν συνεπάγεται της απόλυτης ταύτισης των δύο αυτών μορφών των μεταμοντέρνων κοινωνιών. Μολονότι αποτελούν δύο διαφορετικές προσεγγίσεις, διαθέτουν σημαντικές ομοιότητες όσον αφορά τα δομικά χαρακτηριστικά τους καθώς και την κριτική που ασκούν στον τεχνοεπιστημονικό ντετερμινισμό της πολιτικής και της οικονομίας, η οποία μοιάζει με ένα πεπρωμένο το οποίο πρέπει να αποφευχθεί.

Οι πολύπλοκες λειτουργίες των μαύρων κουτιών σε συνδυασμό με τις διακινδυνεύσεις που αυτές εμπεριέχουν, επιτάσσουν την ανάγκη συγκρότησης συστημάτων «εξηγήσιμης τεχνητής νοημοσύνης», όπως τα XAI και NI4XAI, τα οποία συγκροτούνται σε αυτά τα πλαίσια σε χρηματοδότηση από την Ευρωπαϊκή Ένωση (Cassauwers, 2020). Αφού έχει αναλυθεί η έννοια του μαύρου κουτιού μέσω της πορείας της στο επιστημονικό πεδίο των STS αλλά και ο κεντρικός ρόλος που αυτή έχει στις κριτικές προσεγγίσεις των αλγοριθμικών συστημάτων, θα πραγματοποιηθεί μια συγκριτική ανάλυση των προσεγγίσεων των θεωριών SCOT και ANT όσον αφορά τους τρόπους που αυτά τα σχήματα μπορούν να συμβάλλουν σε μια κοινωνιολογική προσέγγιση διάνοιξης των μαύρων κουτιών της τεχνητής νοημοσύνης.

B.4.3. Η διάνοιξη των μαύρων κουτιών μέσω των θεωριών ANT και SCOT

Το άνοιγμα μαύρων κουτιών μπορεί να πραγματοποιηθεί με πάρα πολλούς τρόπους, καθώς τα σύγχρονα τεχνοεπιστημονικά δημιουργήματα διέπονται από τεράστιο βαθμό συνθετότητας. Αναφερόμενος σε αυτή την συνθετότητα ο Καστελβέτσι, (Castelvecchi, 2016) πραγματεύεται την δυσκολία του να ανοίξει το μαύρο κουτί ενός νευρωνικού δικτύου που είχε φτιάξει σχετικά με την αυτόματη οδήγηση ενός στρατιωτικού αυτοκινήτου εντός του αστικού ιστού, καθώς θεωρεί ότι «η αποκρυπτογράφηση του μαύρου κουτιού έχει γίνει εκθετικά πιο δύσκολη και πιο επείγουσα. Η ίδια η τεχνολογία έχει ραγδαία μετατραπεί σε «πολυπλοκότητα και εφαρμογή». Κάτι τέτοιο γίνεται ακόμα πιο σύνθετο την εποχή των Μεγάλων Δεδομένων, των Αλγοριθμικών Κοινωνιών (Airoldi, 2021) και των Κοινωνιών της Τεχνητής Νοημοσύνης

(Mutzner,2020), όπου εντός ενός μαύρου κουτιού ενδέχεται να βρεθεί ένα ή περισσότερα άλλα μαύρα κουτιά.

Η συνθετότητα του εγχειρήματος του διαφωτισμού, των σκοτεινών πλευρών των σύγχρονων τεχνολογικών, οικονομικών, πολιτικών και επιστημονικών μαύρων κουτιών (Pasquale,2016) μετατοπίζει το βάρος από την απλή διάνοιξή τους στον τρόπο με τον οποίο αυτά θα διανοιχθούν. Όπως λοιπόν προαναφέρθηκε, η αντιμετώπιση κάθε φαινομένου σαν ένα μαύρο κουτί τέθηκε υπό πραγμάτευση από το Ισχυρό Πρόγραμμα της Επιστημονικής Γνώσης του Τόμας Μπλέρ, (Bloor, 1975). Το ισχυρό πρόγραμμα συνέβαλε στην διαμόρφωση της Κοινωνικής Κατασκευής της Τεχνολογίας (SCOT) των Τρέβερ Πίντς και Βίμπε Μπίκερ, ενώ και η Θεωρία Δρώντων Δικτύου του και η σχεσιακή υλικότητα που πραγματεύεται στον τεχνοεπιστημονικό κόσμο εξετάζεται σαν ένα μαύρο κουτί. Ενώ και οι δύο παραπάνω προσεγγίσεις αντιλαμβάνονται την διαδικασία παραγωγής της επιστημονικής γνώσης σαν μια κοινωνική δραστηριότητα (Winner, 1993; Baron και Gomez 2016; Sismondo, 2016), εμφανίζουν διαφορές στον τρόπο διάνοιξης των μαύρων κουτιών παρουσιάζοντας ταυτόχρονα δύο ενδιαφέρουσες μεθοδολογίες διάνοιξης.

Πρώτου όμως να πραγματοποιηθεί κάποια περεταίρω ανάλυση των δύο θεωρητικών αυτών προσεγγίσεων όσον αφορά το άνοιγμα των μαύρων κουτιών στην τεχνητή νοημοσύνη, καθίσταται επιτακτική ανάγκη για διασαφήνιση των λόγων χρήσης αυτών των δύο συγκεκριμένων θεωρητικών σχημάτων, καθώς και της επιλογής να μην χρησιμοποιηθούν άλλες. Αρχικά, οι δύο θεωρητικές κατευθύνσεις αναπτύχθηκαν προηγουμένως στα πλαίσια της κοινωνικής κατασκευής της τεχνοεπιστήμης κάτω από την οπτική των Σπουδών Επιστήμης Τεχνολογίας και Κοινωνίας (STS). Συγκεκριμένα, αναλύθηκε το πώς, η Κοινωνική Κατασκευή της Τεχνολογίας επηρεασμένη, αν όχι διαμορφωμένη πλήρως, από το ισχυρό πρόγραμμα, αντιμετωπίζει τις τεχνολογικές εφευρέσεις ως μια κοινωνική διαδικασία όπου διάφορες κοινωνικές ομάδες αλληλοεπιδρούν και επηρεάζουν την διαδικασία διαμόρφωσής της. Από την άλλη πλευρά, και η Θεωρία Δρώντων Δικτύου ασχολείται με την τεχνοεπιστημονική διαμόρφωση μέσω της δικτυακής, τεχνολογικής και υλικής αλληλεπίδρασης. Και οι δύο θεωρίες διατηρούν μια αξίωση σχετικά με την πολλαπλή ανάγνωση του τεχνοεπιστημονικού ως μια διαδικασία που διαμορφώνεται κάτω από ενέργειες πολλών κοινωνικών δραστών (Bijker και άλλοι, 1989, Latour, 1999).

Στα πλαίσια αυτά και οι δύο θεωρίες μπορούν να αξιοποιηθούν σαν σχήματα που μπορούν να εξυπηρετήσουν στο άνοιγμα μαύρων κουτιών. Ωστόσο, για την

πραγματοποίηση μιας τέτοιας ενέργειας, θεωρητικές κατασκευές όπως το Ισχυρό Πρόγραμμα της κοινωνιολογίας της επιστημονικής γνώσης θα μπορούσαν να παράγουν γόνιμα αποτελέσματα στον διαφωτισμό των σκοτεινών πλευρών της τεχνολογίας των αλγορίθμων, και της τεχνητής νοημοσύνης. Η αποστασιοποιημένη εξέταση της επιστημονικής γνώσης και η συμμετρική απόσταση από την αλήθεια και το ψεύδος των επιστημονικών κατασκευών, αντιμετωπίζουν την επιστημονική γνώση σαν ένα μαύρο κουτί (Collins, 1983).

Στην περίπτωση του Μπλέρ όμως, το ενδιαφέρον των αρχών που θέτει (1975) περιορίζεται στην επιστημονική γνώση χωρίς να αναφέρεται στις τεχνολογικές πρακτικές. Κάτι τέτοιο αναδεικνύει τους δυϊσμούς της σκέψης της κατασκευασιοκρατίας, η οποία αναγνωρίζει την ανθρώπινη κοινωνικότητα ως συγκροτησιακή ουσία. Επιπλέον το ισχυρό πρόγραμμα συγκρότησης της επιστημονικής γνώσης διατηρεί μια διαφορετική σκοπιά από το SCOT. Ενώ λοιπόν το SCOT εξετάζει ένα ευρύ πλαίσιο από δράστες όπως επιστήμονες, μηχανικούς, καταναλωτές, ποικίλες κοινωνικές ομάδες (πολλές φορές και ετερόκλητες μεταξύ τους), το ισχυρό πρόγραμμα επικεντρώνεται στις ιντερναλιστικές και εξωτερικές δυνάμεις που επηρεάζουν συγκεκριμένα την επιστημονική γνώση. Η προοπτική αυτή, δίνει έμφαση στην μη μεροληπτική ανίχνευση της επικρατούσας συνθήκης επιρροής της επιστημονικής γνώσης και όχι στην ανοικτότητα της κοινωνικής διαπραγμάτευσης της μορφής που θα λάβει ένα τεχνούργημα (Barnes και άλλοι, 1996) Βάσει των προαναφερθέντων και δεδομένου ότι τα μαύρα κουτιά συνιστούν κοινωνικά και όχι μόνο κατασκευάσματα, η πραγμάτευση δύο θεωριών, της SCOT και της ANT, οι οποίες εξετάζουν τις διαδικασίες διαμόρφωσης του τεχνοεπιστημονικού στα πλαίσια της τεχνητής νοημοσύνης και των αλγορίθμων κάτω από τον άξονα της κοινωνικής δημιουργίας, αποτελεί μια δόκιμη επιλογή.

Πραγματοποιώντας λοιπόν μια διάνοιξη ενός μαύρου κουτιού, όπως αυτό της τεχνητής νοημοσύνης πρέπει πρώτα να διευκρινιστεί ποια είναι η ουσία μίας τέτοιας πράξης. Με λίγα λόγια πρέπει να αναδειχθούν οι οντολογικές αντιλήψεις που διατηρεί η κάθε θεωρία και που αναπόφευκτα θα επηρεάσουν την διαδικασία διαφωτισμού ενός μαύρου κουτιού. Κοινό σημείο των δύο θεωριών συνιστά η αντίληψη της κοινωνικής κατασκευής της τεχνολογίας και της επιστήμης (τεχνοεπιστήμης), συνεπώς και η κοινωνική διαδικασία παραγωγής της τεχνητής νοημοσύνης. Ωστόσο, ένα άνοιγμα του μαύρου κουτιού βασιζόμενο στην κοινωνική κατασκευή της τεχνολογίας δεν μπορεί να αποκεντρώσει το οντολογικό της ενδιαφέρον από τον άνθρωπο. Για αυτήν, μια μη

μεροληπτική διαδικασία κατανόησης ενός μαύρου κουτιού, μιας δηλαδή κατασκευής η οποία θεωρείται δεδομένη, πραγματοποιείται αντιλαμβανόμενη όλες τις ενέργειες εντός της ως ανθρώπινες και ως εκ τούτου κοινωνικές.

Σε αντίθεση με την θέση αυτή, όπως έχει αναλυθεί, η Θεωρία Δρώντων Δικτύου, δεν περιορίζει το οντολογικό της ενδιαφέρον στους ανθρώπους. Οι δράστες δεν είναι μονάχα άνθρωποι. Σύμφωνα με τον Λατούρ, η Θεωρία Δρώντων Δικτύου «δεν περιορίζεται σε μεμονωμένους ανθρώπινους δράστες, αλλά διευρύνει την λέξη δράστης ή «actant» στο μη ανθρώπινο, μη ανθρώπινες μεμονωμένες οντότητες. Οι υβριδικοί αυτοί δράστες (Γεωργοπούλου, 2010; Blok και Jensen, 2011) λαμβάνουν ουσία, υπάρχουν μόνο εντός των δικτύων που αλληλοεπιδρούν. Συνεπώς, όλες οι οντότητες εντός των μαύρων κουτιών της τεχνοεπιστήμης, αλλά και τα μαύρα τα κουτιά αυτά καθαυτά αποτελούνται από ανθρώπινες και μη οντότητες οι οποίες ορίζονται μέσω της δυναμικής αλληλεπίδρασής τους.

Στην περίπτωση των αλγορίθμων και της τεχνητής νοημοσύνης, το άνοιγμα των μαύρων κουτιών κάτω από την θεωρία της κοινωνικής κατασκευής της τεχνολογίας και της Θεωρίας Δρώντων Δικτύου παρουσιάζει αισθητές οντολογικές διαφορές. Η Θεωρία Δρώντων Δικτύου λαμβάνει υπόψιν όλες τις οντότητες οι οποίες στελεχώνουν ένα σύστημα τεχνητής νοημοσύνης, ξεκινώντας από τους επιστήμονες δεδομένους, τους μηχανικούς δεδομένων, τους νομικούς συμβούλους, έως και τις βάσεις δεδομένων, τους αλγόριθμους και τα νευρωνικά δίκτυα. Κάθε οντότητα αποτελεί έναν δράστη ο οποίος λαμβάνει χαρακτηριστικά μόνο σε αλληλεπίδραση με άλλους δράστες (Law, 2008).

Από την άλλη πλευρά, το πεδίο της Κοινωνικής Κατασκευής της Τεχνολογίας, στην περίπτωση διάνοιξης του μαύρου κουτιού ενός αλγοριθμικού συστήματος, εμβαθύνει στην διαμόρφωση του συστήματος τεχνητής νοημοσύνης εξετάζοντας τις κοινωνικές ομάδες οι οποίες συμμετέχουν στην διαδικασία σχηματισμού του, αυτές που επηρεάζουν αλλά και επηρεάζονται από την χρήση της. Κάτι τέτοιο έχει ως συνέπεια να πραγματοποιείται μια συμμετρική εξέταση των κοινωνικών ομάδων που δημιουργούν μια τέτοια τεχνική δημιουργία (Bijker και άλλοι, 1989) . Δηλαδή, πέρα από τους τεχνικούς, επιστήμονες και πολιτικούς που ευθύνονται για την δημιουργία του συστήματος, υπό εξέταση τίθενται και οι πληθυσμοί που επηρεάζονται από την τεχνητή νοημοσύνη.

Στην συγκεκριμένη περίπτωση, πρέπει να διευκρινιστεί το περιεχόμενο του όρου επιρροή στα πλαίσια της κοινωνικά κατασκευάσιμης τεχνολογίας. Με την χρήση

του δεν πραγματοποιείται αναφορά μονάχα στα άτομα που επηρεάζονται περισσότερο από την χρήση αυτής τεχνολογίας, αλλά και αυτούς που επηρεάζονται σε μικρότερο βαθμό και έκταση. Παραδείγματος χάριν, στην εξέταση ενός μοντέλου εντοπισμού ρίσκου που χρησιμοποιεί τεχνητή νοημοσύνη στα πλαίσια υλοποίησης ενός Δημοσίου προγράμματος, δεν πρέπει να εξεταστούν μονάχα τα άτομα που επωφελήθηκαν από το πρόγραμμα. Η εξέταση του επωφελούμενου πληθυσμού θα εξέθετε στην αφάνεια πληθυσμούς, οι οποίοι για κάποιους λόγους δεν επωφελήθηκαν από το πρόγραμμα. Κάτι τέτοιο, ενδεχομένως θα καθιστούσε αόρατες τυχόν αδικίες και μεροληψίες που χαρακτηρίζουν τα μαύρα κουτιά (O’Neil 2016; Pasquale 2016; Eubanks, 2018; Benjamin, 2019; Crawford, 2021), αλλά παράλληλα θα αποτελούσε πρόσφορο έδαφος για την διατήρηση της αδιαφάνειας που τα χαρακτηρίζει και την αναπαραγωγή των διακινδυνεύσεων αυτών των μεροληψιών (Beck, 1992).

Οι διαφορές σχετικά με την οντολογία που εξετάζει η κάθε θεωρία, οι οποίες επηρεάζουν παράλληλα και το άνοιγμα των μαύρων κουτιών της τεχνοεπιστήμης, δεν είναι το μόνο σημείο που ξεχωρίζουν η ANT και η θεωρία SCOT. Η στιγμή που ένα τεχνοεπιστημονικό εύρημα μετατρέπεται σε μαύρο κουτί, καθορίζεται από διαφορετικούς παράγοντες. Στην Θεωρία Δρώντων Δικτύου, η κλειστότητα ενός μαύρου κουτιού πραγματοποιείται όταν ένα πολύπλοκο δίκτυο από δρώντες απλοποιείται και λαμβάνεται ως δεδομένο. Ο Λατούρ αναφερόμενος στο κλείσιμο ενός μαύρου κουτιού γράφει «από την στιγμή που ένα μαύρο κουτί κλείσει, προσφέρει ένα σταθερό στο σημείο πάνω στο οποίο μπορεί να οικοδομηθεί κάτι περεταίρω. Όταν ανοίξει ξανά, το μαύρο κουτί γίνεται ένα πεδίο μάχης το οποίο προσφέρει μια συναρπαστική οπτική της διαδικασίας διαμόρφωσης της επιστήμης» (1999, σ.186). Έτσι, το κλείσιμο ενός μαύρου κουτιού σηματοδοτεί την σταθεροποίηση του δικτύου δραστών που το συγκροτούν.

Για το SCOT δεν ισχύει το ίδιο. Το κλείσιμο προκύπτει όταν διάφορες κοινωνικές ομάδες αντιλαμβάνονται ένα τεχνούργημα με τον ίδιο, δεδομένο τρόπο. Η διαδικασία αυτή ωστόσο, δεν προκύπτει μόνο κάτω από συνθήκες συμφωνίας και σύμπνοιας. Οι κοινωνικές ομάδες προβαίνουν σε διαπραγματεύσεις και συγκρούσεις, έως ότου καταλήξουν στην παγίωση μιας εικόνας του δεδομένου (Pinch & Bijker, 1984). Ακόμα δηλαδή και αν το «κλείσιμο» είναι στιγμιαίο, η κοινωνική διαδικασία αποδοχής ενός τεχνολογικού κατασκευάσματος ως μαύρο κουτί παρουσιάζει διάρκεια στον χρόνο.

Στο μαύρο κουτί των Αλγορίθμων και της Τεχνητής Νοημοσύνης, το κάθε ένα από τα δύο θεωρητικά σχήματα που παρουσιάστηκαν, μπορούν να αξιοποιηθεί με διαφορετικό τρόπο. Οι διαφορές που παρουσιάζουν τόσο ως θεωρίες, όσο και ως διαφορετικές διαδικασίες διάνοιξης που χρησιμοποιούν, τις καθιστούν, βάσει των όσων αναφέρθηκαν μέχρι στιγμής, ικανές για να φωτίσουν πτυχές των μαύρων κουτιών που παράγονται στα πλαίσια της τεχνητής νοημοσύνης, δίχως όμως να δύνανται να διαφωτίσει τις ίδιες πτυχές. Αυτό οφείλεται στον πολύπλοκο χαρακτήρα της τεχνητής νοημοσύνης.

Κάθε αλγόριθμος και αλγοριθμική λειτουργία μπορεί να χαρακτηριστεί ως μαύρο κουτί από την στιγμή που η εσωτερικές του δράσεις, διαδικασίες δεν είναι διαφανείς. Παράλληλα, σε ένα σύστημα τεχνητής νοημοσύνης ενδέχεται να λειτουργούν για την περαίωση μιας διαδικασίας πάνω από ένας αλγόριθμοι. Η συνθήκη αυτή καθιστά τα παραπάνω θεωρητικά σχήματα ικανά να συνεισφέρουν στον διάλογο σχετικά με την διάνοιξη συγκεκριμένου τύπου μαύρων κουτιών τεχνητής νοημοσύνης. Παραδείγματος χάρη, βάσει της τυπολογίας που προτείνει ο Λιού (Liu, 2021), η Θεωρία Δρώντων Δικτύου μπορεί να επωφεληθεί την διάνοιξη μαύρων κουτιών που αφορούν την Επιστημονική Τεχνητή Νοημοσύνη (Scientific AI) καθώς και στην Τεχνική Τεχνητή Νοημοσύνη (Technical AI).

Η μελέτη των εργαστηρίων αναπτύχθηκε από πολλούς κοινωνικούς επιστήμονες, μεταξύ αυτών ο Λατούρ και ο Γούλγκαρ, οι οποίοι μέσω του έργου τους «Laboratory Life: The Construction of Scientific Facts» (1986) πραγματεύονται την δημιουργία των επιστημονικών δεδομένων και την χειραγώγησή μέσω των υλικών εργαλείων από πολλαπλούς δρώντες. Η επιστημονική διαδικασία δεν θεωρείται μια αμερόληπτη διαδικασία. Πολιτικά, οικονομικά και επιστημονικά συμφέροντα οδηγούν το δίκτυο των δραστών σε πολύπλοκους σχηματισμούς, οι οποίοι πολλές φορές δεν υπάγονται σε κάποια θετικιστική επιστημονικότητα (Knor Cetina, 2013; Sismondo, 2016).

Η επιστήμη των δεδομένων πολλές φορές αποτελεί έναν εργαστήριο παραγωγής μαύρων κουτιών. Οι επιστήμονες που ασχολούνται με την επεξεργασία και ανάλυση δεδομένων, με νευρωνικά δίκτυα, με αλγορίθμους και συστήματα τεχνητής νοημοσύνης, κατά την διάρκεια παραγωγής της τεχνοεπιστημονικής γνώσης, έρχονται σε επαφή με πληθώρα τεχνοεπιστημονικών δρώντων, όπως βάσεις δεδομένων, τεχνολογίες cloud (Cloud Technologies), προγράμματα ψηφιακής επικοινωνίας και άλλους τεχνοεπιστημονικούς δρώντες, οι οποίοι περιπλέκουν την επιστημονική

διαδικασία, δυσχεραίνοντας με αυτόν τον τρόπο την κατανόηση των αποτελεσμάτων αυτών των συστημάτων. Εν τέλει, όλη η προαναφερθείσα διαδικασία καθίσταται ένα μαύρο κουτί. Η Θεωρία Δρώντων Δικτύου ενδείκνυται για την σταθεροποίηση του δικτύου δραστών, καθώς και για την χαρτογράφηση των τεχνοεπιστημονικών διαδικασιών, οι οποίες λαμβάνουν μέρος εντός της επιστημονικής κοινότητας.

Επιπλέον, η θεωρία Δρώντων Δικτύου διαφοροποιείται από άλλες κonstrουκτιβιστικές προσεγγίσεις χάρη στην ικανότητά της να διανοίγει τα μαύρα κουτιά της τεχνοεπιστήμης, μέσω της σχεσιακής υλικότητας. Ένα δίκτυο δρώντων ξεδιπλώνεται μέσω διαφορετικών ερωτήσεων. Αντί της συμμετρικής ερωτήσεως «γιατί», όσον αφορά την επιτυχία ή μη, την συμμετοχή ή απόχη κοινωνικών ομάδων από την τεχνική δημιουργία, που κληροδότησε το ισχυρό πρόγραμμα του Μπλέρ, η Θεωρία Δρώντων Δικτύου ρωτάει τον «πώς» (Law, 2008). Η ερώτηση αυτή αφορά τόσο τους ίδιους τους δράστες, όσο και τα δίκτυα που αυτοί απαρτίζουν. Η ανάλυση του τρόπου με τον οποίο τα δίκτυα δρώντων συγκροτούν συστήματα υβριδικού τύπου (Γεωργοπούλου, 2010; Blok & Jensen, 2011a) μπορεί να είναι ιδιαίτερα ωφέλιμη και στην περίπτωση της Τεχνικής Τεχνητής Νοημοσύνης (Technical AI).

Συγκεκριμένα, η ANT προσφέρει μια «ανιχνευτική» οπτική των τεχνικών χειρισμών και διαδικασιών που συντελούν στην λειτουργία συστημάτων τεχνητής νοημοσύνης. Η προσέγγιση αυτή δεν περιορίζεται μόνο στους ανθρώπινους δράστες. Αυτή η δυναμική εξυπηρετεί στην ανάλυση, τόσο ημιαυτομάτων αλγοριθμικών συστημάτων (semi automated algorithmic systems), όσο και αμιγώς αυτομάτων αλγοριθμικών συστημάτων (automated algorithmic systems). Ακόμα δηλαδή και να μην δρα εντός δικτύου ένας ανθρώπινος δράστης, η συστημική διαδικασία μπορεί να χαρτογραφηθεί σε μεγάλο βαθμό γενίκευσης αλλά και σε μικρό. Παραδείγματος χάρη, μελετώντας ένα αλγοριθμικό σύστημα, μπορούν να ανιχνευθούν όλοι οι κόμβοι που το συγκροτούν, να οπτικοποιηθούν οι σχέσεις μεταξύ των μη απαραίτητα ανθρωπίνων οντοτήτων ενώ παράλληλα, γενικεύοντας το πεδίο εστίασης, παρέχεται και μια συνολική εικόνα του δικτύου δρώντων του μαύρου κουτιού.

Η Θεωρία Δρώντων Δικτύου, ενώ μπορεί να χρησιμεύσει στο άνοιγμα των μαύρων κουτιών της Επιστημονικής Τεχνητής Νοημοσύνης και της Τεχνικής Τεχνητής Νοημοσύνης βάσει την τυπολογία που προτείνει ο Λιού (Liu, 2021), αντιμετωπίζει δυσκολίες στο άνοιγμα του Μαύρου Κουτιού της Πολιτισμικής Τεχνητής Νοημοσύνης (Cultural AI). Το γεγονός αυτό προκύπτει καθώς η ANT είναι μια «πολιτισμικά επίπεδη» θεωρία, αφού αντιλαμβάνεται όλους τις οντότητες που συγκροτούν ως

ορθολογικούς δράστες (Sismondo,2016). Έτσι, μεροληπτικές διακρίσεις και ανισότητες οι οποίες συγκροτούνται από την φύση των μαύρων κουτιών, τα οποία διέπονται από αδιαφάνεια και κλειστότητα, δεν μπορούν να ερμηνευθούν σε πολιτισμικά πλαίσια.

Στη διάνοιξη των μαύρων κουτιών της Πολιτισμικής Τεχνητής Νοημοσύνης και των πολιτισμικών ανισοτήτων που αυτά δημιουργούν, η θεωρία της κοινωνικής κατασκευής της τεχνολογίας (SCOT) επιτρέπει μια πολιτισμικά πληρέστερη οπτική εξέτασης. Συγκεκριμένα, οι Eynon και Young (2021) χρησιμοποιώντας την προσέγγιση της SCOT, εξετάζουν τον τρόπο με τον οποίο η τεχνητή νοημοσύνη χρησιμοποιείται στην εκπαίδευση, με κύριο ενδιαφέρον το πώς αυτή γίνεται κατανοητή από διάφορες κοινωνικές ομάδες.

Στην έρευνα τους εντοπίζουν ότι οι μοχλοί πολιτικής, όπως οι επενδύσεις και οι οικονομικές ρυθμίσεις, δεν αξιοποιούνται επαρκώς για να επηρεάσουν την ανάπτυξη της τεχνητής νοημοσύνης στη δια βίου μάθηση και την εφαρμογή της σε διάφορες ομάδες ενδιαφερομένων. Προκειμένου να διαμορφωθεί θετικά η τεχνητή νοημοσύνη στο πλαίσιο του πεδίου της αέναης μάθησης, είναι απαραίτητο να συμβιβαστούν αυτά τα διαφορετικά μεταξύ τους πλαίσια, όπως τα ζητήματα εμπειρογνωμοσύνης, η ροή δεδομένων, η διαλειτουργικότητα και τα ατομικά δικαιώματα. Επιπλέον, μελετώντας τις αντιλήψεις των ομάδων που εξετάστηκαν, τίθενται ζητήματα όπως η ανάγκη για βελτιωμένη επικοινωνία μεταξύ εκπαιδευτικών και εμπειρογνομόνων της τεχνητής νοημοσύνης, η πρόσβαση σε πόρους, τα δεδομένα και οι αγορές, η συνεργασία μεταξύ ερευνητών και εμπορικών εταιρών, καθώς και ευρύτερα ερωτήματα σχετικά με τους τύπους των ευκαιριών μάθησης με την χρήση τεχνητής νοημοσύνης, που απαιτούνται από τις σύγχρονες κοινωνίες. Τέλος, τονίζεται η σημασία του διαλόγου αλλά και η ενεργή συμμετοχή μεταξύ των εκπαιδευτικών ιδρυμάτων και άλλων φορέων όσον αφορά τα παραπάνω ζητήματα.

Πάρα λοιπόν από το γεγονός ότι υπάρχουν πολλές τεχνικές για την διάνοιξη των μαύρων κουτιών, οι προσεγγίσεις από τον χώρο των STS, αυτή της κοινωνικής κατασκευής της τεχνολογίας, SCOT, όσο και η θεωρία Δρώντων Δικτύου δύνανται να συμβάλουν σημαντικά στην διάνοιξη των μαύρων κουτιών των τεχνολογιών της τεχνητής νοημοσύνης. Ακόμα και αν η κάθε μια διακρίνεται από τα δικά της ιδιαίτερα χαρακτηριστικά, τα οποία την διαφοροποιούν από την άλλη, μπορούν και οι δύο να διαφωτίσουν τις σκοτεινές, ακατανόητες και πολύπλοκες πτυχές της τεχνητής νοημοσύνης μέσω μιας κοινωνιολογικής προσέγγισης.

B.5. Η μεροληπτική διακυβέρνηση μέσω μεγάλων δεδομένων και αλγορίθμων

Η επανάσταση των μεγάλων δεδομένων σε συνδυασμό με την ραγδαία ανάπτυξη των συστημάτων τεχνητής νοημοσύνης, έχουν ωθήσει τα κράτη να πραγματοποιήσουν αλλαγές στον τρόπο με τον οποίο παραδοσιακά λειτουργούσαν. Ωστόσο, παρά τις έντονες εφαρμογές των τεχνολογιών της τεχνητής νοημοσύνης στον ιδιωτικό τομέα, σύμφωνα με τους Sun και Medaglia (2019), τέτοιες τεχνολογίες εμφανίζονται μόνο τα τελευταία χρόνια στον δημόσιο τομέα. Η υιοθέτηση αλγοριθμικών συστημάτων στα πλαίσια μιας «αλγοριθμικής διακυβέρνησης» (Coletta & Kitchin, 2017), διευρύνει σημαντικά τις δυνατότητες των κρατών στην λειτουργία τους και στην άσκηση πολιτικής. Ωστόσο, παρά τις σημαντικές διευκολύνσεις που προσφέρουν τα συστήματα τεχνητής νοημοσύνης στην εποχή των μεγάλων δεδομένων, τίθενται ισχυρά ζητήματα σχετικά με τις μεροληπτικές διακρίσεις των αλγοριθμικών συστημάτων, οι οποίες έχουν έντονα κοινωνικά αντίτυπα.

Στα πλαίσια αυτά, η Κοινωνιολογία της τεχνητής νοημοσύνης και ειδικότερα οι κριτικές θεωρίες που προσεγγίζουν τους τομείς οι οποίοι πολλές φορές συγχέονται, δηλαδή αυτούς των μεγάλων δεδομένων και της τεχνητής νοημοσύνης, μπορούν να συνεισφέρουν σημαντικά στην ανάγνωση των ανισοτήτων στον δημόσιο και τον ιδιωτικό τομέα. Η διακυβέρνηση μέσω αυτών των συστημάτων απαιτεί μια έντονη οργανωτική διερεύνησή τους. Η εξέταση και οι αναλύσεις τους, πραγματοποιούνται αναφορικά με τους ίδιους τους οργανισμούς, δηλαδή εξετάζουν τους τρόπους με τους οποίους οι οργανισμοί μπορούν να αυξήσουν την αποδοτικότητά τους, τα οικοσυστήματα που μπορούν διαμορφώσουν με αυτούς κ.α. Κάτι τέτοιο, «θα μπορούσε να σημάνει τον κίνδυνο μιας οργανοκεντρικής προοπτικής που αφορά πρωτίστως παράγοντες και δυναμικές σχετικά με τα δεδομένα και τον εκσυγχρονισμό των δεδομένων σε οργανωτικό επίπεδο...εις βάρος των εξωτερικών κοινωνικών, πολιτικών και πολιτισμικών παραγόντων σε μακροεπίπεδο που δομούν και (ανα)παράγουν την πολιτική και την πρακτική διακυβέρνησης που σχετίζονται με τα δεδομένα» (J. Liu, 2022).

Η οργανοκετρική προοπτική που διέπει τα τεχνοεπιστημονικά συστήματα, συγκροτεί μια σύγχρονη μορφή διακινδύνευσης. Η διαρκής ενασχόληση με τις τεχνικές πτυχές, τις οικονομίες γύρω από αυτές τις τεχνολογίες, επιδεικνύει μια τεχνοκρατική

προσέγγιση των συστημάτων τεχνητής νοημοσύνης. Τα συστήματα αυτά διαχειρίζονται πληθώρα κοινωνικών δεδομένων, τα οποία ακόμα και αν χρήζουν πολλές φορές εξαιρετικά σημαντικής προστασίας και διαχείρισης, εκτίθενται σε πολλαπλές μορφές κινδύνων, όπως σε έκθεση σε τρίτους ή σε μεροληπτική αξιοποίηση στα πλαίσια λήψης αποφάσεων (Eubanks, 2018). Η διαρκής προσήλωση στην υιοθέτηση τέτοιων ισχυρών τεχνοεπιστημονικών συστημάτων, δίχως την εξέταση των επιπτώσεων που η χρήση τους μπορεί να επιφέρει στις κοινωνίες και στους πολίτες, μαρτυρά την μετατόπιση από μια διακυβέρνηση αξιοποίησης των κοινωνικών δεδομένων για την κοινωνία «Social Data Governance» (Liu, 2022), σε μία διακυβέρνηση ενίσχυσης των ίδιων των συστημάτων.

Η εφαρμογή αλγοριθμικών πολιτικών από τα δημόσια θεσμικά όργανα, πολλές φορές λαμβάνει έντονα χαρακτηριστικά αυτοματοποίησης. Συστήματα λήψης αποφάσεων αξιοποιούν πλέον συστήματα μηχανικής μάθησης και αλγορίθμους, αποσκοπώντας στην εύρεση των καλύτερων δυνατών επιλογών σχετικά με θεσμικές, οικονομικές, πολιτικές ακόμα και νομικές (Platonova, n.d.) επιλογές. Η λήψη αποφάσεων μεταφέρεται σε μια διαδικασία, όπου οι αλγόριθμοι λαμβάνουν μεγαλύτερο ρόλο στην λήψη αποφάσεων από τους ανθρώπους σε όλο το φάσμα του Δημόσιου τομέα. Παράλληλα, αλγοριθμικά συστήματα ανίχνευσης και αξιολόγησης ρίσκου χρησιμοποιούνται για την ανίχνευση διακινδυνεύσεων που σχετίζονται με την εγκληματικότητα και τις υγειονομικές κρίσεις. Η πανδημία του Covid-19 αποτέλεσε μια έντονη περίοδο χρήσης τέτοιων συστημάτων επιτήρησης (Brevini & Pasquale, 2020).

Στις κοινωνίες των μεγάλων δεδομένων, η αξιοποίηση αυτών των αλγοριθμικών μορφών διακυβέρνησης, θέτει σημαντικούς προβληματισμούς γύρω από τους τρόπους με τους οποίους η χρήση συστημάτων τεχνητής νοημοσύνης επηρεάζει την κοινωνική πραγματικότητα. Η αυτοματοποιημένη διακυβέρνηση όπως αναλύθηκε και προηγουμένως μπορεί πολλές φορές να αξιοποιείται στα πλαίσια προστασίας της κοινωνίας αλλά την ίδια στιγμή η αλόγιστη χρήση της μπορεί να αποτελεί πυρήνα διάδοσης διακινδυνεύσεων. Η απουσία ελέγχου σε τέτοια κοινωνικοτεχνικά συστήματα μπορεί αν επιφέρει μεροληπτικά αποτελέσματα μέσω της μετατροπής των συστημάτων αυτών σε μαύρα κουτιά. Χαρακτηριστική περίπτωση αποτελεί το παράδειγμα των ολλανδικών παροχών παιδικής μέριμνας ή αλλιώς «Toeslagenaffaire», στο οποίο ο αλγόριθμος ταξινόμησης ρίσκου που αξιοποιήθηκε στα πλαίσια προστασίας οικονομικής εξαπάτησης, μετατράπηκε σε ένα μαύρο κουτί το οποίο στην

πάροδο του χρόνου δρώντας ανεξέλεγκτα αναπαρήγαγε εξαιτίας της αυτοματοποιημένης του φύσης πληθώρα μεροληπτικών αποφάσεων.

Γ.«Toeslagenaffaire»

Το σκάνδαλο των ολλανδικών παροχών παιδικής μέριμνας, επίσης γνωστό ως η υπόθεση «Toeslagenaffaire», είναι μια εντυπωσιακή απεικόνιση του τρόπου με τον οποίο οι αλγόριθμοι που δημιουργήθηκαν με την πρόθεση να επιτευχθεί γραφειοκρατική αποτελεσματικότητα και δημοσιονομική προσοχή μπορούν να γίνουν ακούσια εργαλεία κοινωνικής ανισότητας και μεροληπτικών πολιτικών. Η συγκεκριμένη περίπτωση αποτελεί χαρακτηριστικό παράδειγμα μεροληπτικών αλγοριθμικών εφαρμογών, οι οποίες αποτελούν ένα μαύρο κουτί για τους πολίτες, καθώς ο αυτόματος αλγόριθμος που χρησιμοποιήθηκε εντός αυτού δρούσε ξενοφοβικά, δίχως να ελέγχεται από κάποια αρχή (Amnesty International, 2021). Οι συνέπειες εφαρμογής αυτή της αυτοματοποιημένης μεροληπτικής αλγοριθμικής πολιτικής διαπερνούν την σφαίρα της οικονομικής ζωής των πολιτών, οι οποίοι λανθασμένα επιλέχθηκαν από τον αλγόριθμο ως πιθανοί απατεώνες (Amnesty International, 2021), ειςχωρώντας στην προσωπική και οικογενειακή τους ζωή, μέσω της μετακίνησης παιδιών από την οικία τους, καθώς και την επιδείνωση της υγείας των επιλεχθέντων.

Η ολλανδική κυβέρνηση υλοποίησε ένα αυτοματοποιημένο σύστημα για τον συντονισμό της κατανομής των παροχών παιδικής μέριμνας στις αρχές του 21ου αιώνα. Ο σχεδιασμός του αλγορίθμου βασίστηκε στην αρχή της αντικειμενικότητας και της ουδετερότητας, με σκοπό την εξάλειψη των ανθρώπινων προκαταλήψεων μέσω βασικών αποφάσεων σχετικά με τα πρότυπα δεδομένων. Από την μελέτη της περίπτωσης προκύπτει ότι ο συγκεκριμένος αλγόριθμος, ακόμα και αν αρχικώς δηλωνόταν από τους υπεύθυνους διαχείρισής του ότι διατηρούσε την ουδετερότητά του και δεν λάμβανε μεροληπτικές αποφάσεις, εν τέλει αυτός λειτουργούσε βασιζόμενος σε μεροληπτικές διακρίσεις.

Το σύστημα σχεδιάστηκε για να εντοπίσει «παρατυπίες», χρησιμοποιώντας μεταβλητές που χρησίμευαν ως μεσάζοντες για πιθανή εξαπάτηση. Είναι κρίσιμο να σημειωθεί, ωστόσο, ότι αυτές οι μεταβλητές δεν ήταν ουδέτερες, αλλά επηρεάστηκαν από κοινωνικές προκαταλήψεις και στερεότυπα, αντανακλώντας τη θεωρία της Noble

για τους «αλγόριθμους της καταπίεσης» (2018). Οι επιλεγμένες προσεγγίσεις ήταν φορτωμένες με υποθέσεις που οδήγησαν σε φυλετικό και κοινωνικοοικονομικό προφίλ, το οποίο εστίαζε στην διπλή ιθαγένεια και την εθνικότητα των πολιτών.

Στην πράξη, αυτό σήμαινε ότι οι οικογένειες από χαμηλό εισόδημα ή μεταναστευτικό υπόβαθρο στοχεύονταν δυσανάλογα σε σχέση με άλλους πληθυσμούς. Παρά την απουσία σαφών αποδεικτικών στοιχείων, αυτές οι οικογένειες χαρακτηρίστηκαν «απατεώνες» με αποτέλεσμα πολλαπλές διακρίσεις και οικονομική καταστροφή για χιλιάδες από αυτές. Σε αυτό το πλαίσιο, η έννοια της «Κοινωνίας των Μαύρων Κουτιών» του Πασκάλ είναι ιδιαίτερα σχετική. Το αλγοριθμικό «μαύρο κουτί» αποκρύπτει τη διαδικασία λήψης αποφάσεων, υπονομεύοντας την λογοδοσία και τη διαφάνεια, με τις οποίες αυτά τα συστήματα είναι νομικά υποχρεωμένα να λειτουργούν (Seldam & Brenninkmeijer, 2021).

Η υπόθεση «Toeslagenaffaire» αποτελεί μια περίπτωση η οποία έχει αποτελέσει αντικείμενο κριτικής από διάφορους ερευνητές αλλά και οργανισμούς διεθνούς κλίμακας, όπως η Διεθνής Αμνηστία. Από την βιβλιογραφική ανασκόπηση για την εξέταση της συγκεκριμένης περίπτωσης, παρά την βιβλιογραφία και σχετικά άρθρα τα οποία παρουσιάζουν το ζήτημα, δεν βρέθηκαν πολλές πηγές οι οποίες εξετάζουν το παρόν περιστατικό κοινωνιολογικά, εκτός της σχετικής έκθεσης της Διεθνούς Αμνηστίας, η οποία ωστόσο δεν αναπτύσσεται αμιγώς σε κοινωνιολογικά πλαίσια. Μέσω λοιπόν της ανάλυσης της Κοινωνιολογίας της Τεχνητής Νοημοσύνης, όπως αυτή παρουσιάστηκε στην προηγούμενη ενότητα, η παρούσα εργασία θα εξετάσει το σκάνδαλο των ολλανδικών παροχών παιδικής φροντίδας, υπό το πρίσμα των θεωριών των κριτικών σπουδών των αλγορίθμων.

Γ.1. Το κοινωνικοπολιτικό υπόβαθρο του αναστοχαστικού μοντέλου ταξινόμησης ρίσκου: Υψηλό ρίσκο-υψηλή αμοιβή

Προκειμένου να πραγματοποιηθεί μια κοινωνιολογική προσέγγιση σχετικά με το σύστημα τεχνητής νοημοσύνης που χρησιμοποιήθηκε στην υπόθεση «Toeslagenaffaire», δεν πρέπει ο αλγόριθμος που αξιοποιήθηκε να αναλυθεί ως ένα αμιγώς τεχνικό σύστημα το οποίο κάτω από ορισμένες συνθήκες πραγματοποιεί σφάλματα. Αντιθέτως, η ανάλυση του αλγορίθμου θα γίνει με δεδομένο ότι αυτός αποτελεί ένα κοινωνικοτεχνικό δημιούργημα. Δηλαδή, θα εξεταστεί ως ένα κοινωνικό κατασκεύασμα, το οποίο συγκροτείται από τις κοινωνικές, πολιτικές, οικονομικές και

τεχνολογικές συνθήκες που επικρατούν αλλά και από τις ιστορικές συνθήκες που έχουν οδηγήσει στην δημιουργία του. Οι προσεγγίσεις και θεωρίες που ανέδειξαν οι STS, συνεισφέρουν σημαντικά στην μελέτη του υποβάθρου συγκρότησης του μοντέλου ταξινόμησης ρίσκου.

Σύμφωνα με την έκθεση της Διεθνούς Αμνηστίας, «Ξενοφοβικές Μηχανές» (2021), η Ολλανδία γνώρισε μια πρωτόγνωρη μορφή εκμετάλλευσης και απάτης όσον αφορά το σύστημα κοινωνικής ασφάλισης, με το ολλανδικό κράτος να βιώνει μια από τις πρώτες μορφές εξαπάτησής του. Η ενέργεια αυτή οδήγησε στην αυστηροποίηση των πλαισίων γύρω από αυτές τις δράσεις, με τις ολλανδικές φορολογικές αρχές να αναθεωρούν το σύστημα διοικητικών κυρώσεων που εφαρμοζόταν έως τότε. Έτσι, το 2005 ξεκίνησαν στην Ολλανδία οι παροχές παιδιών, στις οποίες θα αναμειγνυόταν αποφασιστικά και οι πλέον αυστηρές φορολογικές αρχές.

Το σύστημα παροχών τέκνου συγκροτήθηκε στα πλαίσια ενίσχυσης των οικογενειών με χαμηλό εισόδημα. Συγκεκριμένα, η ολλανδική κυβέρνηση θα κάλυπτε μέρος από τα έξοδα των γονέων με την μορφή αποζημίωσης. Δικαιούχοι του προγράμματος ήταν γονείς ή φροντιστές, οι οποίοι κατοικούσαν νομίμως στην Ολλανδία και πηγαίναν τα παιδιά τους σε ημερήσια κέντρα φροντίδας. Τα χρήματα δίνονταν προκαταβολικά, ενώ σε περίπτωση πλεονάσματος, αυτό θα έπρεπε να επιστραφεί. Το ποσό λήψης που θα λάμβαναν οι φροντιστές ή οι γονείς εξαρτιούνταν από πολλούς παράγοντες με τα ποσά κατά κύριο λόγο να σχετίζονται με το οικονομικό υπόβαθρο των δικαιούχων, με τους χαμηλότερα αμειβόμενους να λαμβάνουν αποζημίωση έως και 96%, ενώ στους υψηλότερα αμειβόμενους θα έφτανε έως και το 33.3%. (Amnesty International, 2021).

Το 2011, η ολλανδική κυβέρνηση προκειμένου να αντιμετωπίσει τον φόβο εκμετάλλευσης του συστήματος κοινωνικής ασφάλισης, όπως στην περίπτωση της «Βουλγαρικής Απάτης» (Dutch Data Protection Office, 2020), υιοθέτησε «την επιβολή αυτοματοποιημένων πολιτικών επιβολής νόμου και πρακτικών, οι οποίες στόχευαν στην καταπολέμηση της απάτης στον τομέα των παροχών παιδικής μέριμνας, αφήνοντας ελάχιστο ή καθόλου περιθώριο για την προσαρμογή των μέτρων στην ατομική κατάσταση των γονέων και των φροντιστών» (Amnesty International, 2021, σ.11). Τέτοιου τύπου αυτοματοποιημένες διαδικασίες όχι μόνο περιορίζουν τον ανθρώπινο παράγοντα από τα να παρεμβαίνει σε αυτές, αλλά όπως περιγράφει η Noble, «τα αυτοματοποιημένα συστήματα λήψης αποφάσεων είναι δυσανάλογα επιβλαβή για τους πιο ευάλωτους και τους λιγότερο ισχυρούς, οι οποίοι έχουν ελάχιστη ικανότητα

να παρεμβαίνουν σε αυτά - από την παραπλανητική εκπροσώπηση έως τη φυλάκιση, την πρόσβαση στην πίστωση και άλλες φόρμουλες που επηρεάζουν τη ζωή» (2018, σ.50).

Το 2013, οι φορολογικές αρχές υιοθέτησαν ένα αλγοριθμικό σύστημα λήψης αποφάσεων για τον εντοπισμό της απάτης στις αιτήσεις και τις ανανεώσεις παροχών παιδικής μέριμνας (Heikkilä, 2022). Αυτό το σύστημα, γνωστό ως μοντέλο ταξινόμησης κινδύνου, ενσωμάτωσε στοιχεία αυτοδιδασκαλίας για να δημιουργήσει προφίλ επικινδυνότητας των αιτούντων για την παροχή παιδικής φροντίδας, οι οποίοι ήταν φαινομενικά πιο πιθανό να υποβάλουν ανακριβείς και δόλιες πληροφορίες και ενδεχομένως να διαπράξουν απάτη. Το άμεσο αποτέλεσμα της επιλογής από το μοντέλο ταξινόμησης κινδύνου ήταν η παύση των παροχών.

Οι φορολογικές αρχές έπρεπε να αποδείξουν την αποτελεσματικότητα των μεθόδων ανίχνευσης απάτης του νέου μοντέλου ταξινόμησης κινδύνου, το οποίο βασιζόταν σε δεδομένα. Για να πληρώσουν για τις νέες μεθόδους που αξιοποιήθηκαν, προέβησαν στην κατάσχεση επαρκών κεφαλαίων από τους φερόμενους απατεώνες. Ανεξάρτητα από την αλήθεια των κατηγοριών για απάτη, αυτή η πρόκληση δημιούργησε ένα διεστραμμένο κίνητρο για την απόκτηση όσο το δυνατόν περισσότερων κεφαλαίων. Μια συνέπεια της υψηλής βαθμολογίας κινδύνου στο μοντέλο ταξινόμησης κινδύνου ήταν ότι οι δημόσιοι υπάλληλοι απαιτούσαν από τους γονείς και τους φροντιστές να παράσχουν πρόσθετα αποδεικτικά στοιχεία για να καθορίσουν την επιλεξιμότητά τους για παροχές.

Ωστόσο, οι γονείς και οι φροντιστές που προσπάθησαν να καθορίσουν ποιες πληροφορίες κρίθηκαν εσφαλμένες ή παραπλανητικές ή ποιες αποδείξεις κρίθηκαν ελλιπείς, ήλθαν αντιμέτωποι συχνά με σιωπή, καθώς οι φορολογικές αρχές αρνήθηκαν να εξηγήσουν τις αποφάσεις τους. Κάτι τέτοιο είχε ως αποτέλεσμα για πολλά χρόνια οι γονείς και οι φροντιστές να μην έχουν την ευκαιρία αφενός να κατανοήσουν τι δεν συμβαίνει σωστά, αφετέρου να αποδείξουν την αθωότητά τους. Επιπλέον, οι γονείς και οι φροντιστές, οι δημοσιογράφοι, οι πολιτικοί, οι εποπτικοί φορείς και τα μέλη της κοινωνίας των πολιτών δεν μπόρεσαν να αποκτήσουν ουσιαστικές πληροφορίες σχετικά με την ύπαρξη και τη λειτουργία του μοντέλου ταξινόμησης κινδύνου. (Amnesty International, 2021)

Η ανάγκη επιβεβαίωσης της ορθής λειτουργίας του συστήματος είχε ως αποτέλεσμα την υιοθέτηση μιας οικονομικής επιλογής, η οποία θα εξέθετε τους πολίτες σε μεγαλύτερο κίνδυνο από το ίδιο το σύστημα αξιολόγησης κινδύνου. Συγκεκριμένα,

το νέο μοντέλο ανίχνευσης ρίσκου θα χρηματοδοτούνταν για τις δράσεις του κατάσχοντας χρήματα μέσω των απορριφθέντων αιτήσεων που ανιχνεύθηκαν από το ίδιο το σύστημα ως «απάτη» (Amnesty International, 2021). Προκειμένου όμως μια αίτηση να χαρακτηριστεί έτσι, έπρεπε πρώτα να λάβει τον ανάλογο βαθμό επικινδυνότητας, δηλαδή τον βαθμό υψηλού ρίσκου (High Risk). Πολλοί δικαιούχοι του προγράμματος, παρά το γεγονός ότι θεωρούσαν ότι δεν υπάρχει κάποιο σφάλμα ή απόπειρα απάτης προς το σύστημα, απευθυνόμενοι στους αρμόδιους του προγράμματος δεν μπορούσαν να λάβουν κάποια βέβαιη απάντηση σχετικά με την αιτιολογία που αυτές επιλέχθηκαν και χαρακτηρίστηκαν με αυτόν τον τρόπο. Το μαύρο κουτί του μοντέλου ταξινόμησης είχε ήδη σφραγιστεί.

Η εκτεταμένη χρήση του μοντέλου ταξινόμησης ρίσκου επέφερε πληθώρα αρνητικών αποτελεσμάτων, προβλήματα τα οποία δεν περιορίστηκαν μονάχα στους αιτούντες του προγράμματος. Από την Διεθνή Αμνηστία (2021) γίνεται γνωστό ότι με την διακοπή της ενίσχυσης εξαιτίας της εσφαλμένης επιλογής από το αλγοριθμικό σύστημα, οι οικογένειες που προηγουμένως είχαν επωφεληθεί από το προνόμιο παιδικής φροντίδας έπρεπε να επιστρέψουν μέρος του χρηματικού ποσού πίσω. Ειδικότερα, οι γονείς και οι φροντιστές που χαρακτηρίστηκαν από τον αλγόριθμο ως «απατεώνες» και είχαν επίσης λάβει περισσότερα από 10.000 ευρώ σε ετήσιες παροχές, έπρεπε να επιστρέψουν περισσότερα από 3.000 ευρώ συνολικά επειδή σημαδεύτηκαν αυτόματα για σκόπιμη πρόθεση ή υπερβολική αμέλεια, χωρίς επαλήθευση. Η προσθήκη της ετικέτας «σκόπιμη πρόθεση ή σοβαρή αμέλεια» απέκλεισε τους γονείς και τους φροντιστές από ένα σχέδιο πληρωμών που θα τους επέτρεπε να επιστρέψουν τα οφέλη που έλαβαν σε δόσεις με την πάροδο του χρόνου. Οι φορολογικές αρχές χρησιμοποίησαν το κατώτατο όριο των 3.000 ευρώ για να επιταχύνουν τις αποφάσεις σχετικά με χιλιάδες αιτήσεις γονέων και κηδεμόνων που ζητούσαν πρόσβαση στα συστήματα πληρωμής (Amnesty International, 2021).

Από τα προαναφερθέντα γίνεται κατανοητό ότι το μοντέλο ταξινόμησης ρίσκου συνιστά μια κοινωνικοτεχνική κατασκευή, η οποία δεν συγκροτήθηκε τυχαία με τον συγκεκριμένο τρόπο. Η δημιουργία του μοντέλου αυτού, όπως και κάθε κοινωνικοτεχνικού κατασκευάσματος, λαμβάνει στοιχεία από το κοινωνικό περιβάλλον στο οποίο και αυτή συγκροτείται. Υπό το πρίσμα αυτό, το μοντέλο ταξινόμησης ρίσκου διαμορφώνεται στα πλαίσια της αναστοχαστικότητας του εκσυγχρονισμού (Beck, 2015). Δηλαδή, το σύστημα αυτό αναπτύχθηκε το 2013 με στόχο την επίλυση προβλημάτων, τα οποία το σύστημα κοινωνικής ασφάλισης

παρουσίαζε ήδη από τις δεκαετίες του 1980 και 1990. Πρόκειται για μια αιτιακή αλυσίδα, όπου για κάθε πρόβλημα που παρουσιάζεται, όπως αυτό της οικονομικής εκμετάλλευσης του ασφαλιστικού συστήματος, αναπτύσσεται μια νέα, σε κάθε περίπτωση εκσυγχρονιστική λύση, όπου και αυτή με την σειρά της αποτελεί πηγή προβλημάτων. Το παράδειγμα αυτό, όπως γράφει η Benjamin «Δεν είναι απλά μια ιστορία του “τότε” εναντίον του “τώρα”. Πρόκειται για το πώς ιστορικές διαδικασίες καθιστούν εφικτό το παρόν» (2019, σ.147).

Γ.1.1 Από την ανίχνευση κινδύνων στην παραγωγή τους

Σε αυτό το σύστημα μεγάλο ενδιαφέρον παρουσιάζει η οικονομική πολιτική που υιοθετήθηκε για την λειτουργία του. Συγκεκριμένα, η οικονομική βιωσιμότητά του εξαρτάται από τον αριθμό των αιτήσεων που απορρίπτονται ως απάτες. Στην πραγματικότητα λοιπόν, όποια αίτηση επισημαινόταν ως ενδεχόμενη απάτη μέσω του υψηλού σκορ που αυτή συγκέντρωνε, απορριπτόταν απευθείας. Όπως συζητήθηκε προηγουμένως, το μαύρο κουτί του αλγορίθμου, από την στιγμή που έκλεισε, κανένας δεν βρισκόταν σε θέση να μπορεί να απαντήσει με σαφήνεια για την αιτιολογία απόρριψης των αιτήσεων στους δικαιούχους. Το κλείσιμο του μαύρου κουτιού του αλγορίθμου, σε συνάρτηση με το οικονομικό σχέδιο που διασφάλιζε την λειτουργία του συγκρότησαν μια συστηματική διακινδύνευση.

Η ντετερμινιστική απόρριψη των αιτήσεων, δημιούργησε μια μορφή πεπρωμένου απόρριψης (Beck, 2015, Pasquale, 2016) για αυτούς που επιλέγονταν από αυτή. Υπό αυτές τις συνθήκες, ο οργανισμός που ήταν υπεύθυνος για την λειτουργία του αλγορίθμου, είχε την δυνατότητα απλώς να προγραμματίσει τον αλγόριθμο να βάζει υψηλότερα σκορ σε ορισμένου τύπου αιτήσεις, οι οποίες θα αναλυθούν εκτενέστερα στην συνέχεια, κάτι το οποίο θα συνεπαγόταν την απόρριψη των αιτήσεων αυτών, άρα και μεγαλύτερο κέρδος. Στις Κοινωνίες του Μαύρου Κουτιού, η αλγοριθμική ποσοτικοποίηση και η συγκρότηση προφίλ ρίσκου (O'neil, 2016; Pasquale, 2016; Eubanks, 2018; Noble, 2018; Benjamin, 2019) στα πλαίσια της άσκησης πολιτικής αποτελεί μια σημαντική μορφή διακινδύνευσης.

Στην περίπτωση του συστήματος εντοπισμού ρίσκου, η υψηλή διακινδύνευση μέσω της βαθμολόγησης με υψηλό σκορ, είχε διαφορετική σημασία από εκείνη που της απέδιδαν οι πολίτες. Για το σύστημα ίσχυε η φράση «υψηλό ρίσκο - υψηλό κέρδος» που συναντάται πολλές φορές σε συζητήσεις με χρηματιστηριακές θεματικές.

Ο χαρακτηρισμός μιας αιτήσεως ως αίτηση με υψηλό ρίσκο, σήμαινε εν τέλει τον ταυτόχρονο αποκλεισμό της και την επιστροφή χρημάτων από τα άτομα που επιχειρούσαν να εξαπατήσουν το σύστημα. Με αυτόν τον τρόπο το σύστημα εξασφάλιζε την χρηματοδότησή του, επομένως την οικονομική βιωσιμότητά του. Αντιθέτως, οι αιτήσεις με χαμηλή βαθμολόγηση κινδύνου, επειδή αξιολογούνταν θετικά περνώντας τον έλεγχο του συστήματος, δεν προσέφεραν κέρδος για την συντήρηση του αλγορίθμου.

Για τους πολίτες ωστόσο, η επισήμανση της αιτήσεώς τους είχε ακριβώς το αντίθετο αποτέλεσμα. Με την απόρριψή τους, αυτοί έπρεπε να επιστρέφουν ένα σημαντικό μέρος των χρημάτων που είχαν λάβει, επιδεινώνοντας με αυτόν τον τρόπο την δεινή οικονομική τους κατάσταση, αυτή δηλαδή που τους είχε ήδη οδηγήσει εξ αρχής να αιτηθούν. Εδώ η αναλογία που παρουσιάστηκε προ λίγου εμφανίζεται αντίστροφα. Για τους πολίτες, μια αίτηση με υψηλό ρίσκο, η οποία συνεπάγεται την ετικετοποίησή τους ως «σημαντικά αμελής» (Amnesty International, 2021), σημαίνει χαμηλή ή καθόλου αμοιβή. Με άλλα λόγια το υψηλό ρίσκο συνεπάγεται πολύπλευρη ζημία.

Αυτή η αντίστροφη αναλογία, η οποία εν τέλει εφαρμόστηκε πολλές φορές αδικώς εξαιτίας της μεροληπτικής διάκρισης του συστήματος, επιβεβαιώνει την αξιοποίηση του αλγορίθμου έναντι των αδύναμων και Υποεκπροσωπούμενων κοινωνικών ομάδων. Η αναστοχαστική τάση βαθμολόγησης κινδύνου και αυτή της συγκρότησης προφίλ ως πανάκεια στην αντιμετώπιση ζητημάτων όπως οι οικονομικές απάτες ή η επιτήρηση πληθυσμών (Zuboff, 2019) απαρτίζει μια μορφή διακινδύνευσης η οποία βιώνεται κατά κύριο λόγο από τους πληθυσμούς στους οποίους αυτή εφαρμόζεται και όχι από αυτούς που την εφαρμόζουν.

Γ.2. Η μεροληπτική σχεδίαση του συστήματος ταξινόμησης ρίσκου

Πέρα από την πολιτική και οικονομία που συγκροτήθηκε γύρω από το αυτοματοποιημένο σύστημα, αυτό ξεχωρίζει για τον τρόπο με τον οποίο είχε προγραμματιστεί. Από την πραγμάτευση της παρούσας εργασίας έχει γίνει κατανοητό ότι η ανάπτυξη αυτοματοποιημένων συστημάτων δεν συνάδει απαραίτητα με ουδέτερα αποτελέσματα. Η παρούσα περίπτωση αποτελεί ένα χαρακτηριστικό παράδειγμα ενός μεροληπτικού συστήματος τεχνητής νοημοσύνης, το οποίο λειτουργεί μεροληπτικά απέναντι σε ομάδες ατόμων με καταγωγή από άλλες χώρες.

Από την έκθεση του Ολλανδικού Γραφείου Προστασίας Δεδομένων δίνονται συγκεκριμένες πληροφορίες σχετικά με το πώς η καταγωγή ενός ατόμου επηρέαζε το αποτέλεσμα βαθμολόγησης από το αυτοματοποιημένο σύστημα ρίσκου. Στα συστήματα αυτά, όπως έχει αναλυθεί γενικότερα στην εργασία, προκειμένου να γίνει κατανοητός ο τρόπος που αυτά πραγματοποιούν διακρίσεις, πρέπει να αναλυθεί ο μεσάζοντας επικοινωνίας των δραστών που τις πραγματοποιούν, δηλαδή ο κώδικας αυτών των συστημάτων (Benjamin, 2019). Η προγραμματιστική τεκμηρίωση μιας ομάδας η οποία αντιμετωπίστηκε μεροληπτικά σε σχέση με τον υπόλοιπο πληθυσμό, εξυπηρετεί κατά κύριο λόγο στις συζητήσεις, όπου μια τέτοια ενέργεια συνιστά «απλώς ένα τεχνικό λάθος». Η εξέταση όμως των κριτηρίων με τα οποία ο αλγόριθμος υπογραμμίζει τις αιτήσεις και τις βαθμολογεί με υψηλό σκορ, αναδεικνύει τις προθέσεις των υποκειμένων που τον διαχειρίζονται και πιο συγκεκριμένα την θεώρησή τους σχετικά με το ποια υποκείμενα θέλουν να το εξαπατήσουν.

Όπως λοιπόν αναλύθηκε προηγουμένως, η αυτοματοποίηση δημόσιων υπηρεσιών από τις δημόσιες αρχές της Ολλανδίας είχε ξεκινήσει ως αποτέλεσμα της προσπάθειας αντιμετώπισης της εξαπάτησης των θεσμών. Τα αλγοριθμικά συστήματα, τα οποία επιλέχθηκαν για την επιτήρηση αυτών των διαδικασιών σύμφωνα με την Διεθνή Αμνηστία (2021), λειτουργούσαν αυτοματοποιημένα βάσει δεδομένων που τους δίνονταν από προηγούμενες αιτήσεις. Μια από τις πρώτες αρμοδιότητες που έλαβαν οι αλγόριθμοι αυτοί ήταν η εκπαίδευσή τους πάνω σε μορφές λανθασμένων αιτήσεων (Dutch Data Protection Office, 2020). Το σύστημα σχεδιάστηκε για να επιλέγει τις αιτήσεις με την υψηλότερη βαθμολογία κινδύνου, δηλαδή το 0,8, οι οποίες στη συνέχεια ελέγχονταν χειροκίνητα από έναν δημόσιο υπάλληλο, ο οποίος εάν χειριζόταν μια αίτηση με την συγκεκριμένη βαθμολογία είχε το δικαίωμα να κηρύξει την αίτηση εσφαλμένη. Ενώ λοιπόν ο υπάλληλος είχε την δυνατότητα χαρακτηρισμού μιας αίτησης ως εσφαλμένης, δεν είχε πληροφορηθεί σχετικά με τα κριτήρια που ο αλγόριθμος την βαθμολογούσε.

Ακόμα και αν το σύστημα θεωρούντο ότι λειτουργούσε στα πλαίσια της ουδετερότητας, η μελέτη που πραγματοποιήθηκε από την Αρχή Προστασίας Δεδομένων αναδεικνύει πως το σύστημα ταξινόμησης ρίσκου αντιλαμβανόταν ως παράγοντα, ο οποίος επιτείνει την πιθανότητα εξαπάτησης του συστήματος, την ιθαγένεια και σε μερικές αιτήσεις την εθνικότητα. Ειδικότερα, ο αλγόριθμος είχε εκπαιδευτεί στην αναζήτηση της μεταβλητή της ιθαγένειας για τους πολίτες που εξέταζε. Ωστόσο, η ιθαγένεια των πολιτών, βάσει της νομοθεσίας των ολλανδικών

παροχών παιδικής φροντίδας, δεν αποτελεί κριτήριο το οποίο σύστημα πρέπει να λαμβάνει υπόψη όσον αφορά την διαδικασία βαθμολόγησης ρίσκου αφού απαραίτητο στοιχείο για την λήψη της παροχής είναι η νόμιμη διαμονή (Dutch Data Protection Office, 2020). Η παραβίαση των κανονισμών του προγράμματος μέσω του αυτοματοποιημένου αλγορίθμου, επιβεβαίωσε τον κεντρικό άξονα πραγμάτευσης των κριτικών αλγοριθμικών σπουδών, την στόχευση δηλαδή των μειονοτήτων και των περιθωριοποιημένων κοινωνικά ομάδων, καθώς ο αλγόριθμος εν τέλει εκπαιδεύτηκε να βαθμολογεί υψηλά και να επισημαίνει ως επικίνδυνες προς απάτη αιτήσεις με διπλή, όχι μόνο ολλανδική ιθαγένεια.

Μάλιστα, αρκετές ήταν οι περιπτώσεις όπου η εθνικότητα των ολλανδών με διπλή ιθαγένεια, επιλέχθηκε ως κριτήριο χειροκίνητου ελέγχου. Ακόμα και αν υπήρξαν αρνήσεις από επίσημους φορείς, εκ των προαναφερθέντων γίνεται κατανοητό ότι, όχι μόνο ο αλγόριθμος βαθμολογούσε υψηλά τους πολίτες δίχως ολλανδική ιθαγένεια αλλά ταυτόχρονα η εθνική τους προέλευση σε περίπτωση που αυτοί δεν ήταν ολλανδοί, μπορούσε να αποτελέσει αιτία χειροκίνητου ελέγχου από τους δημόσιους υπαλλήλους. Χαρακτηριστικό παράδειγμα αποτελεί αυτό των πολιτών με καταγωγή από την Καραϊβική, οι οποίοι επιλέχθηκαν χειροκίνητα από τους υπαλλήλους και επισημάνθηκαν από αυτούς (Amnesty International, 2021).

Το γεγονός αυτό, τονίζει την θέση του Airoidi (2021) σχετικά με την κουλτούρα εντός του κώδικα. Η οικονομική εκμετάλλευση του συστήματος κοινωνικής ασφάλισης αναγνωρίζεται κατά κύριο λόγο στα πρόσωπα των κοινωνικοοικονομικών μειονοτήτων, των ξένων στο ευρωπαϊκό κράτος της Ολλανδίας πολιτών. Η αρνητική απάντηση ως παράγοντας διακινδύνευσης στην μεταβλητή ολλανδική ιθαγένεια, θέτει την ύπαρξη ολλανδικής ιθαγένειας ως τεκμήριο οικονομικής και εν τέλει πολιτισμικής ασφάλειας. Το σύστημα ανίχνευσης ρίσκου, οι υπεύθυνοι διαχείρισής και οι χειριστές του, αποτελούν και οι ίδιοι τμήμα του μεροληπτικού μηχανισμού. Η προαγωγή εθνοτικών διακρίσεων, δίχως τεκμήρια εξαπάτησης, στον βωμό της προστασίας των κρατικών πόρων, λειτουργεί αφαιρετικά του δικαιώματός του πληθυσμού αυτού για ενίσχυση. Οι εκτεταμένες μεροληπτικές διακρίσεις που προέκυψαν, δεν είναι μονάχα αποτέλεσμα της ξενοφοβικής κωδικοποίησης του αλγορίθμου. Η μεροληπτική εφαρμογή του δεν έγινε κατανοητή, αφού το σύστημα τεχνητής νοημοσύνης που χρησιμοποιήθηκε, από την αρχή της λειτουργίας του αποτέλεσε ένα μαύρο κουτί.

Γ.3. Το μαύρο κουτί του συστήματος ταξινόμησης ρίσκου

Η υπόθεση «Toeslagenaffaire», έγινε γνωστή αρκετά χρόνια μετά την πρώτη εμφάνιση και χρήση του συστήματος. Αρκεί κανείς να αναλογιστεί ότι η λειτουργία του αλγορίθμου διακόπηκε το 2020 ενώ ο αλγόριθμος είχε ξεκινήσει την λειτουργία ήδη από το 2013. Όπως και οι υπόλοιπες διακινδυνεύσεις, έτσι και αυτές του μαύρου κουτιού των συστημάτων τεχνητής νοημοσύνης διατηρούνται πολλές φορές χάρη στην «αορατότητα» τους. Από την πραγμάτευση του Πασκάλ (2016), σχετικά με την Κοινωνία του Μαύρου Κουτιού έγινε κατανοητή η πολύπλευρη διαδικασία συγκρότησης ενός μαύρου κουτιού. Έτσι και στην περίπτωση του συστήματος ταξινόμησης ρίσκου που εφαρμόστηκε στο πρόγραμμα των ολλανδικών παροχών παιδικής φροντίδας, η διαδικασία μετατροπής του αυτοματοποιημένου συστήματος, αφενός δεν είναι στιγμιαία, αφετέρου σχετίζεται και με άλλους δράστες πέρα από το σύστημα αυτό κάθε αυτό.

Η μεροληπτική εφαρμογή των παροχών παιδιού δεν ξεκινάει μονάχα από την συγκρότηση του αυτοματοποιημένου συστήματος ρίσκου αλλά και από τα δεδομένα τα οποία αξιοποιήθηκαν καθώς και την πορεία αυτών σε διάφορες κρατικές πλατφόρμες. Συγκεκριμένα, τα δεδομένα σχετικά με την διπλή ιθαγένεια, τα οποία αξιοποιήθηκαν μεροληπτικά από τον αλγόριθμο οδηγώντας σε διακοπή των παροχών από πολλές οικογένειες, διατηρήθηκαν στην βάση δεδομένων του προγράμματος και αξιοποιήθηκαν παράτυπα. Η πορεία της διατήρησής τους για πολύ καιρό αποτελούσε ένα μαύρο κουτί τόσο στο ευρύ κοινό όσο και σε ορισμένους υπαλλήλους. Οι πληροφορίες σχετικά με αυτήν αντλούνται μέσω της αναφοράς της ολλανδικής Αρχής Δεδομένων (2020).

Η διπλή ιθαγένεια, η οποία αποτέλεσε μεταβλητή που συνυπολογίστηκε παράτυπα στην βαθμολόγηση ρίσκου, διατηρήθηκε σαν προσωπικό δεδομένο σε πολλαπλές υπηρεσίες μέχρι να μεταβεί στην υπηρεσία που ήταν αρμόδια για τις παροχές παιδικής φροντίδας. Η πορεία της μεταβλητής της «διπλής ιθαγένειας» ξεκινάει από την βάση δεδομένων του μητρώου πληθυσμού GBA. Σε αυτή, η διπλή ιθαγένεια αναφερόταν επιπλέον της ολλανδικής ιθαγένειας. Ωστόσο, το GBA στις 6 Ιανουαρίου του 2014 διαδέχθηκε το BRP, στο οποίο όμως στους πολίτες με ολλανδική ιθαγένεια δεν αναφέρεται η διπλή τους ιθαγένεια, παρά μόνο η ολλανδική.

Τα στοιχεία από τις παραπάνω υπηρεσίες αξιοποιήθηκαν από το σύστημα της εφορίας BVR το οποίο αφορούσε προσωπικά δεδομένα και δεδομένα διευθύνσεων

(Dutch Data Protection Office, 2020). Το BVR πέρα από το τόπο διαμονής των πολιτών διέθετε πληροφορίες για τις σχέσεις των ατόμων, όπως την οικογενειακή κατάσταση. Το σύστημα BVR ουσιαστικά χαρτογραφούσε το αν οι πολίτες διαμένουν νόμιμα στις Κάτω Χώρες αλλά και το αν τα άτομα που είχαν διάφορες μορφές σχέσης μεταξύ τους διέμεναν νόμιμα σε αυτές. Προκειμένου λοιπόν το σύστημα BVR να έχει πληρέστερο έλεγχο για την νομιμότητα διαμονής αξιοποιούσε δεδομένα από το σύστημα RNI, το σύστημα του «Μητρώου Μη Κατοίκων». Βάσει των προηγούμενων λοιπόν, το BVR διέθετε όλα τα απαραίτητα δεδομένα σχετικά με την νόμιμη ή όχι διαμονή των πολιτών στην Ολλανδία, η οποία αποτελούσε και προϋπόθεση συμμετοχής. Επιπλέον, όλα τα υπάρχοντα δεδομένα σχετικά με την διπλή ιθαγένεια των μη Ολλανδών πολιτών, όπως ισχυρίστηκε ο διευθυντής παροχών αφαιρέθηκαν από το σύστημα του BVR στις αρχές του 2015.

Η μεταφορά των προσωπικών δεδομένων των πολιτών εντός των πλατφόρμων των υπηρεσιών της Ολλανδικής κυβέρνησης δεν ολοκληρώνεται ακόμα. Το σύστημα που αξιοποιήθηκε από το πρόγραμμα παροχών παιδικής μέριμνας ήταν το «Toeslagen Verstrekkingen Systeem» ή αλλιώς TVS. Σύμφωνα με την αναφορά (2020), το TVS ενώ για τους μη Ολλανδούς υπηκόους η διπλή ιθαγένεια είχε αφαιρεθεί, η ιθαγένεια των Ολλανδών πολιτών συνέχισε να διατηρείται εντός του συστήματος. Η διατήρηση της διπλής ιθαγένειας στο TVS αποτελεί και την αφορμή, όπου η διπλή ιθαγένεια αλλά και ευαίσθητες πληροφορίες, που εμφανίζονταν παράλληλα με αυτή, αποτέλεσαν φαινόμενο μεροληπτικής διαχείρισης από το αυτοματοποιημένο αλγοριθμικό σύστημα.

Η σημασία της πορείας των προσωπικών δεδομένων των πολιτών μέσω των κρατικών βάσεων δεδομένων και των συστημάτων έχει μεγάλη σημασία για μια κοινωνιολογική προσέγγιση των μαύρων κουτιών των σύγχρονων κοινωνιών των μεγάλων δεδομένων και της τεχνητής νοημοσύνης. Αυτό διότι το σύστημα τεχνητής νοημοσύνης που αξιοποιήθηκε από τον αλγόριθμο βασίστηκε στην αξιολόγηση ρίσκου βάσει αφενός του σχεδιασμού του και αφετέρου βάσει των δεδομένων που αξιοποίησε. Το πρόγραμμα παροχής παιδικής μέριμνας ουδέποτε προβλεπόταν ότι θα αξιολογούσε ως επικίνδυνη την διπλή ιθαγένεια ολλανδών πολιτών (Dutch Data Protection Office, 2020; Amnesty International, 2021) .

Εφόσον λοιπόν η μεταβλητή Ιθαγένεια δεν αποτελούσε προϋπόθεση συμμετοχής, παρά μόνο η μόνιμη διαμονή, η διπλή ιθαγένεια δεν θα έπρεπε να ληφθεί υπόψιν ούτε στην περίπτωση των ολλανδών πολιτών, ούτε και σε αυτή των μη

ολλανδών πολιτών. Επομένως, η διατήρηση πληροφοριών οι οποίες σύμφωνα με τις προβλέψεις του προγράμματος δεν ήταν απαραίτητες, συγκρότησε ένα κλίμα άγνοιας, τόσο για τους πολίτες, οι οποίοι δεν γνώριζαν τον λόγο απόρριψής τους εφόσον δεν υπήρχε κάποιο λάθος στα στοιχεία που παρέθεταν, όσο και για τους χειριστές τέτοιων υποθέσεων οι οποίοι έπρεπε στη διεξαγωγή έρευνας που πραγματοποιήθηκε να αιτιολογήσουν την ύπαρξη και τους τρόπους διαχείρισης μιας τέτοια πληροφορίας. Η πορεία των προσωπικών δεδομένων της διπλής ιθαγένειας των ολλανδών αποτελεί τμήμα του μαύρου κουτιού του των παροχών παιδικής μέριμνας. Από την GBA, στην BRP, μετέπειτα στην BVR και εν τέλει στην TVS, η πορεία αυτής της μεταβλητής πραγματοποιήθηκε υπό συνθήκες αορατότητας και άγνοιας των πολιτών παρά το γεγονός ότι αυτά αξιοποιήθηκαν εν γνώσει των χειριστών και των υπευθύνων.

Η πραγμάτευση των Πασκάλ (2016) και Beck (2015) σχετικά με την σημασία της γνώσης ως σημαντική μορφή κεφαλαίου, βρίσκεται στο επίκεντρο της παρούσας περίπτωσης. Η γνώση της λειτουργίας του συγκεκριμένου κοινωνικοτεχνικού συστήματος αποτελεί απαραίτητη προϋπόθεση για την πρόληψη από τις μεροληψίες του και τις αρνητικές του επιπτώσεις. Πέρα από την πορεία της μεταβλητής της ιθαγένειας, η μεθοδολογία βαθμολόγησης κινδύνου από το αυτοματοποιημένο σύστημα ρίσκου, αποτελούσε ένα μαύρο κουτί για τους πολίτες και τους δημοσίους υπαλλήλους. Μάλιστα οι δεύτεροι ενώ αναλάμβαναν την εξέταση των περιπτώσεων με υψηλή βαθμολογία κινδύνου, δεν είχαν πληροφορηθεί σχετικά με τον τρόπο αξιολόγησης κινδύνου του συστήματος (Amnesty International, 2021). Το γεγονός αυτό, δηλαδή ο τρόπος κατανομής της εξουσίας, αποτελεί σημαντική προβληματική που τίθεται από τις κριτικές σπουδές των αλγορίθμων.

Όπως τονίζεται από τις CAS, τα συστήματα κατανομής ρίσκου αποτελούν παράλληλα συστήματα κατανομής ισχύος και εξουσίας (Benjamin, 2019; Crawford, 2021). Στα συστήματα αυτά η γνώση συνιστά την ύψιστη μορφή εξουσίας. Στην περίπτωση του εν λόγω συστήματος τεχνητής νοημοσύνης, η γνώση σχετικά με τον τρόπο τον οποίο αυτό αξιολογεί την διακινδύνευση παρέμενε άγνωστη τόσο στους πολίτες όσο και στους δημοσίους υπαλλήλους. Ωστόσο η εξουσία μεταξύ υπαλλήλων και πολιτών δεν κατανεμόταν με τον ίδιο τρόπο. Ακόμα και αν υπήρχε άγνοια και από τις δύο πλευρές σχετικά με την λειτουργία του συστήματος, ο δημόσιος υπάλληλοι μπορούσαν να αξιολογούν χειροκίνητα τις αιτήσεις αλλά και να αναζητούν αιτήσεις με βάση την καταγωγή ολλανδών πολιτών που είχαν διπλή ιθαγένεια.

Η διατήρηση λειτουργίας των συστημάτων αυτών κάτω από συνθήκες αορατότητας και άγνοιας εντείνουν την άνιση κατανομή της εξουσίας (Pasquale, 2016). Τα τεχνοεπιστημονικά μαύρα κουτιά των αυτοματοποιημένων συστημάτων ταξινόμησης ρίσκου συμβάλλουν στην διεύρυνση του χώρου της επιτήρησης για τα άτομα που διαθέτουν την γνώση λειτουργίας τους, συνεπώς και την ισχύ να λάβουν αποφάσεις περί αυτών. Η ποσοτικοποίηση του κινδύνου ιδίως όταν αυτή καλύπτεται από τους ειδικούς με το προσωπείο της ουδετερότητας οδηγεί σε μία νέα μορφή διακινδύνευσης, όπως αυτές τις αντιλαμβάνεται ο Beck (2015), με τις Κοινωνίες του Μαύρου Κουτιού να μετατρέπονται σε «Κοινωνίες της Βαθμολόγησης» (Pasquale, 2016). Στις κοινωνίες αυτές η βαθμολόγηση αποτελεί μια σημαντική μορφή διακινδύνευσης όπου αυτός που κατέχει την εξουσία μπορεί ταυτόχρονα να ρυθμίζει την πραγματικότητα βαθμολογώντας την.

Έτσι και στην περίπτωση των παροχών παιδικής μέριμνας από το ολλανδικό κράτος η «επικινδυνότητα» βαθμολογήθηκε και οριοθετήθηκε κάτω από άγνωστους όρους απέναντι στους πολίτες. Από όσα προκύπτουν από τόσο από την έκθεση της Διεθνούς Αμνηστίας (2021) όσο και από την αναφορά του ολλανδικού Γραφείου Προστασίας Δεδομένων (2020), γίνεται κατανοητή η ανισοτική κουλτούρα στον κώδικα (Airoldi, 2021), η οποία διαπνέει τόσο το κράτος το οποίο διαθέτει την εξουσία να ορίζει το τι είναι επικίνδυνο, όσο και τα συστήματα που αυτό αξιολογεί και με τα οποία ασκεί τις κοινωνικές και οικονομικές πολιτικές του.

Γ.4. Τα αποτελέσματα του μαύρου κουτιού του αλγοριθμικού συστήματος παιδικής μέριμνας ως συμπεράσματα

Στο έργο του «Κοινωνίες της Διακινδύνευσης» (2015) ο Μπεκ, αναπτύσσοντας τα χαρακτηριστικά των διακινδυνεύσεων, τονίζει ότι πολλές φορές οι διακινδυνεύσεις γίνονται κατανοητές όταν τα αποτελέσματά τους γίνονται ορατά. Έτσι και στις Κοινωνίες του μαύρου κουτιού της Τεχνητής Νοημοσύνης εξαιτίας της άγνοιας σχετικά με τις λειτουργίες των συστημάτων αυτών, οι κρίσεις που αυτές δημιουργούν γίνονται ορατές όταν τα αποτελέσματά τους εμφανίζονται. Προκειμένου λοιπόν να ληφθούν γόνιμα κοινωνιολογικά συμπεράσματα για την περίπτωση των ολλανδικών παροχών παιδικής μέριμνας πρέπει να γίνει κατανοητό ότι πρέπει κάποιος να ανατρέξει στα αποτελέσματα του φαινομένου που μελετάται, καθώς αυτά αποτελούν την στιγμή όπου η αόρατη και μεροληπτική λειτουργία του συστήματος τεχνητής νοημοσύνης

γίνεται αντιληπτή. Η στιγμή αυτή αποτελεί και την έναρξη διάνοιξης του μαύρου κουτιού του συστήματος ταξινόμησης ρίσκου.

Το αντίκτυπο της μεροληπτικής βαθμολόγησης από το αυτοματοποιημένο σύστημα ταξινόμησης ρίσκου επέφερε πληθώρα προβλημάτων και κινδύνων στα άτομα που αξιολογήθηκαν αρνητικά από αυτό. Τα άτομα τα οποία είχαν αιτηθεί για αυτήν την ενίσχυση, βρίσκονταν ήδη εκτιθέμενα σε συνθήκες οικονομικές επισφάλειας και κινδύνου. Μάλιστα, πρόκειται για μια από τις πιο αδύναμες κοινωνικά ομάδες καθώς η ενίσχυση δεν αφορούσε μόνο τα άτομα τα οποία δικαιούνταν να αιτηθούν, δηλαδή γονείς και φροντιστές, αλλά και την ομάδα των παιδιών αυτών των οικογενειών. Αυτό έχει ως αποτέλεσμα η δυσχερής οικονομική θέση των γονέων να επηρεάζει άμεσα την ποιότητα ζωής των παιδιών.

Οι αιτούντες που αξιολογήθηκαν αρνητικά με την ετικέτα της «εξαπάτησης» και είχαν λάβει προηγουμένως χρήματα έπρεπε να επιστρέψουν μέρος από το χρηματικό ποσό το οποίο είχε ληφθεί (Amnesty International, 2021). Αφού όμως τα χρήματα της μέριμνας λαμβάνονταν προκαταβολικά το γεγονός ότι αυτά είχαν ξοδευτεί, δημιουργούσε έντονες δυσκολίες σχετικά με την επιστροφή του. Οι οικογένειες που μεροληπτικά και παράνομα (Heikkilä, 2022) επιλέχθηκαν από το αυτοματοποιημένο σύστημα ρίσκου στα πλαίσια οργανωμένης οικονομικής απάτης (Amnesty International, 2021), πέρα από τα σοβαρά αρνητικά προβλήματα με τα οποία ήρθαν αντιμέτωπες, γνώρισαν επιπλέον τον κίνδυνο της καταστροφής τους. Χιλιάδες παιδιά εξαιτίας της οικονομικής αδυναμίας των οικογενειών τους να ανταπεξέλθουν οικονομικά, ήρθαν αντιμέτωπα με την αλλαγή περιβάλλοντος και την απομάκρυνσή τους από την οικογενειακή τους εστία (NI.Times, 2021; Pascoe, 2022) με τα αποτελέσματα των συγκεκριμένων γεγονότων να έχουν έντονες ψυχολογικές επιπτώσεις (Boztas, 2021).

Όλες αυτές οι επιπτώσεις αναδεικνύουν τις διακινδυνεύσεις οι οποίες προκύπτουν από την κλειστότητα και αδιαφάνεια των συστημάτων λήψης αποφάσεων αλλά και γενικότερα των συστημάτων τεχνητής νοημοσύνης. Έτσι και το σύστημα που αξιοποιήθηκε στην περίπτωση των παροχών παιδικής μέριμνας στην Ολλανδία, αποτελεί ένα σημαντικό παράδειγμα μη ουδετερότητας των κοινωνικοτεχνικών συστημάτων. Ο προγραμματισμός του συστήματος ταξινόμησης ρίσκου αλλά και οι αδιαφανείς πορείες που έλαβαν τα δεδομένα της διπλής ιθαγένειας των πολιτών, τονίζουν την διάχυση των στερεοτύπων και διακρίσεων εντός του κώδικα αλλά και την μετατροπή του σε ένα μαύρο κουτί, στο οποίο αδύναμες κοινωνικά ομάδες παράλληλα

με την υπονόμηση θεμελιωδών δικαιωμάτων τους βιώνουν πολύπλευρες συνέπειες εξαιτίας του μεροληπτικού και ξενοφοβικού σχεδιασμού του.

Επίλογος-Συμπεράσματα

Όπως προκύπτει από τα παραδείγματα που αναλύθηκαν αλλά και από την περίπτωση της «Toeslagenaffaire» που μελετήθηκε, η αυτοματοποίηση της διακυβέρνησης τόσο του ιδιωτικού αλλά και του δημοσίου τομέα μέσω της αξιοποίησης συστημάτων τεχνητής νοημοσύνης και μεγάλων, προσωπικών δεδομένων, μπορεί να αποτελέσει μια σημαντική μορφή μεροληπτικής διακινδύνευσης. Η έλλειψη αντικειμενικότητας των συστημάτων τεχνητής νοημοσύνης σε συνάρτηση με την μορφή του μαύρου κουτιού, τα οποία αυτά λαμβάνουν, έχουν έντονες συνέπειες στα κοινωνικά σύνολα στα οποία αυτά εφαρμόζονται. Ωστόσο, όπως αναδείχθηκε από την παρούσα εργασία, τα τεχνοεπιστημονικά συστήματα τεχνητής νοημοσύνης δεν επηρεάζουν μονάχα την κοινωνία. Τόσο η συγκρότησή τους, όσο και ο προγραμματισμός τους επηρεάζονται από τα κοινωνικά περιβάλλοντα, στα οποία τα συστήματα αυτά αναπτύσσονται και εφαρμόζονται, αποδίδοντάς τους εύλογα τον χαρακτηρισμό «κοινωνικοτεχνικά».

Για την κοινωνιολογική ανάλυση των συστημάτων τεχνητής νοημοσύνης ιδιαίτερα σημαντική είναι η αξιοποίηση πόρων από το πεδίο των σπουδών Επιστήμης και Τεχνολογίας, του οποίου οι θεωρίες αξιοποιήθηκαν σε μεγάλο βαθμό στην παρούσα πραγματεία. Συγκεκριμένα από τον χώρο των STS αναδείχθηκε η σημασία του κοινωνικού κονστρουκτιβισμού και της ιδέας της κοινωνικής κατασκευής της τεχνολογίας και της επιστήμης. Τα επιστημονικά ευρήματα, μέχρι και τα τέλη του 20ου αιώνα αντιμετωπίζονταν ως μια μορφή αυθύπαρκτης αλήθειας, η οποία δεν χωρούσε κάποια μορφή αμφισβήτησης καθώς χαρακτηριζόταν από γραμμικότητες και αιτιότητες (Γεωργοπούλου, 2010), δηλαδή από αρχές του θετικισμού οι οποίες εκείνη την χρονική περίοδο έγιναν ταυτόσημες της αλήθειας.

Για την αποφυγή της απόλυτης αναγωγής κάθε κατασκευής σε ένα κοινωνικό κατασκεύασμα όπως πραγματεύτηκαν οι Berger και Luckmann (1967), αναπτύχθηκαν θέσεις, οι οποίες διατήρησαν τον χαρακτήρα της κοινωνικής κατασκευής αλλά ταυτόχρονα οριοθετούσαν τις κονστρουκτιβιστικές πραγματεύσεις γύρω από την επιστήμη. Μια τέτοια πραγματεύση είναι αυτή του Μπλέρ, ο οποίος με τις αρχές που θέσπισε όσον αφορά την κοινωνιολογική μελέτη της επιστημονικής γνώσης επηρέασε την κονστρουκτιβιστική οπτική σχετικά με τις τεχνολογικές κατασκευές. Προσεγγίσεις όπως αυτή των Bijker και άλλων (1989) σχετικά με την Κοινωνική Κατασκευή της Τεχνολογίας, SCOT, επηρεάστηκαν από τις αρχές του Μπλέρ, όπως αυτή την

συμμετρίας, η οποία «θεωρείται ως προειδοποίηση ώστε να μην υποθέτουμε την ύπαρξη ανεξάρτητης τεχνολογίας όταν εξηγούμε τη λειτουργία των μηχανών» (Bijker, 2010, σ.66). Έτσι, ενώ η τεχνολογία πολλές φορές εκλαμβάνόταν ως ξεχωριστό πεδίο της επιστήμης ή εφαρμογή της επιστημονικής γνώσης (Sismondo, 2016), άρχισε να γίνεται αντιληπτή και εκείνη ως μια κοινωνική κατασκευή, πραγματοποιώντας παράλληλα μια πρωτόγνωρη μονιστική προσέγγιση, μια «διάσπαση» της δυϊστικής θεώρησης τεχνολογίας και επιστήμης (Bijker και άλλοι, 1989).

Σε αυτή την μονιστική προσέγγιση της τεχνολογίας και της επιστήμης, αναδεικνύεται καθοριστική η συνεισφορά των Μπρούνο Λατούρ και Στιβ Γούλγκαρ. Οι ρηξικέλευθες προσεγγίσεις των διανοούμενων αυτών, μετέβαλαν σημαντικά την κοινωνιολογική φαντασία σχετικά με την τεχνολογία και την επιστήμη ή όπως προσδιορίζουν οι ίδιοι το «τεχνοεπιστημονικό». Η Θεωρία Δρώντων Δικτύου και η οντολογική κατανόηση της υβριδικής φύσης των δραστών, καθώς και η δυναμική διαμόρφωση της τεχνολογίας μέσω δικτύων στα οποία συμμετέχουν τεχνολόγοι, αντικείμενα και επιστήμονες δίχως περιορισμούς όσον αφορά τις αρμοδιότητές τους, αποτέλεσαν μια νέα οπτική στον τρόπο κατανόησης των κοινωνικών φαινομένων (Law, 2008; Γεωργοπούλου, 2010). Υπό το πρίσμα της θεωρίας αυτής αλλάζει και ο τρόπος μελέτης των τεχνοεπιστημονικών φαινομένων, αφού όπως τονίζουν τα ευρήματα από τον χώρο των «Μελετών Εργαστηρίου», οι επιστημονικές και τεχνολογικές διαδικασίες έχουν έντονα κοινωνικό χαρακτήρα (Latour & Woolgar, 1986; Knorr-Cetina, 2013). Σε αυτά τα πλαίσια τα όρια τεχνολογίας και επιστήμης είναι περισσότερο δυσδιάκριτα παρά διακριτά.

Οι τεχνοεπιστημονικές προσεγγίσεις μέσα από την Θεωρία Δρώντων Δικτύου και τις Μελέτες Εργαστηρίων, επηρέασαν την προσέγγισή του Γούλγκαρ (1985) σχετικά με την τεχνητή νοημοσύνη. Οι αναδυόμενες τεχνολογίες των «expert systems» καθώς και οι ανάμειξη άλλων επιστημονικών πεδίων μαζί τους, προβλημάτισε τον Γούλγκαρ σχετικά με την θέση της Κοινωνιολογίας απέναντι σε ένα τόσο δυναμικά αναπτυσσόμενο επιστημονικό τομέα. Τόσο οι τρόποι προσέγγισης που πρότεινε, όσο και η οντολογικές προβληματικές που έθεσε σχετικά με την ικανότητα των μηχανών να φτιάχνουν οι ίδιες προβλήματα, αποτέλεσαν τις πρώτες κοινωνιολογικές προσεγγίσεις σχετικά με τα τεχνοεπιστημονικά συστήματα τεχνητής νοημοσύνης. Παρά τους χειμώνες που αντιμετώπισε ο επιστημονικός κλάδος της τεχνητής νοημοσύνης (Mutzner, 2020; Schuchmann, 2021; Glover, 2022), τα αυτοματοποιημένα αλγοριθμικά συστήματα, τα συστήματα μηχανικής και βαθιάς μάθησης, βιώνουν πλέον

μια παγκόσμια αναγνώριση και αξιοποίηση σε όλους τους τομείς της κοινωνικής ζωής των πολιτών.

Τα συστήματα τεχνητής νοημοσύνης σε συνάρτηση με την αξιοποίηση των μεγάλων δεδομένων χρησιμοποιούνται σε όλες τις κοινωνίες του εκσυγχρονισμού. Η εκμετάλλευσή τους και η διαπλοκή τους ανάμεσα σε φορείς του ιδιωτικού και του δημοσίου στα πλαίσια διακυβέρνησης, συγκροτούν μια νέα μορφή υβριδικών χώρων στους οποίους αυτοματοποιημένοι αλγόριθμοι αξιοποιούν πληθώρα προσωπικών δεδομένων εν αγνοία των πολιτών στους οποίους αυτά ανήκουν. Οι «Κοινωνίες του Μάυρου Κουτιού» (Pasquale, 2016) αποτελούν μια σύγχρονη, αναστοχαστική μορφή των «Κοινωνιών της Διακινδύνευσης» (Beck, 2015) στις οποίες, η αξιοποίηση αυτοματοποιημένων συστημάτων λήψης αποφάσεων προτάσσουν την ποσοτικοποίηση της ανθρώπινης ζωής (O'neil, 2016) μέσω των βαθμολογιών της ανθρώπινης συμπεριφοράς. Οι «Κοινωνίες της Βαθμολογίας» (Pasquale, 2016) στην προσπάθειά τους να οχυρωθούν από φαινόμενα όπως οικονομικές απάτες και εγκληματικότητα, εντείνουν την αυτοματοποίησή τους με αποτέλεσμα την μαζική παραγωγή μεροληψιών και διακρίσεων.

Η αόρατη διάσταση των συστημάτων τεχνητής νοημοσύνης εντείνει τις μεροληψίες κατά των αδύναμων και Υποεκπροσωπούμενων κοινωνικών ομάδων (O'Neil, 2016; Eubanks, 2018; Noble, 2018; Benjamin, 2019; Crawford, 2021). Οι μεροληπτικές αποφάσεις που λαμβάνονται εντός αυτών των συστημάτων αποτελούν αποτέλεσμα τόσο των δεδομένων που αυτά αξιοποιούν, όσο και του μεροληπτικού σχεδιασμού τους. Τα δεδομένα που χρησιμοποιούνται από τις τεχνολογίες της τεχνητής νοημοσύνης πολλές φορές συλλέγονται με τα ίδια μεροληπτικά φίλτρα με τα οποία ταξινομούνται από τους αλγορίθμους που τα διαχειρίζονται. Τόσο τα δεδομένα, όσο και ο κώδικας προγραμματισμού των συστημάτων αυτών, παραγκωνίζουν την όποια ουδετερότητα που επικαλούνται οι διαχειριστές τους, διοχετεύοντας εντός αυτών τις αξίες των κοινωνιών που τους περιβάλλουν. Στη συνέχεια μεροληπτικά δεδομένα τα οποία παράγονται, αξιοποιούνται εν συνεχεία από άλλα συστήματα, συγκροτώντας έναν κύκλο βλαβών (Beck, 2015), μια μεροληπτική ανακύκλωση στερεοτύπων, με συχνούς αποδέκτες των αρνητικών επιπτώσεών τους τις κοινωνικά Υποεκπροσωπούμενες ομάδες, οι οποίες διαφέρουν ως προς το φύλο, την φυλή, το χρώμα, την εθνότητα, την θρησκεία και το κοινωνικοοικονομικό υπόβαθρο.

Η περίπτωση των ολλανδικών παροχών παιδικής μέριμνας αποτελεί ένα χαρακτηριστικό παράδειγμα των μεροληψιών που παράγονται όταν τεχνολογίες

τεχνητής νοημοσύνης αξιοποιούνται άκριτα στα πλαίσια της δημόσιας, αυτοματοποιημένης διακυβέρνησης. Το σύστημα ταξινόμησης ρίσκου που αξιοποιήθηκε για τον εντοπισμού κινδύνου, προγραμματίστηκε με τέτοιο τρόπο ώστε να αντιλαμβάνεται ως επικίνδυνους προς εξαπάτηση τους ξένους, μη ολλανδούς πολίτες, δίχως να υπακούει στους κανόνες πάνω στους οποίους συγκροτήθηκε, καταπατώντας παράλληλα ανθρώπινα δικαιώματα (Amnesty International, 2021). Το συγκεκριμένο σύστημα ρίσκου, εξαιτίας της «κλειστής» στους πολίτες φύσης του, μετατράπηκε σε ένα μαύρο κουτί ταξινόμησης της εξουσίας, στο οποίο το προνόμιο εντοπισμού του κινδύνου και απόρριψης των πολιτών μονοπωλούσε ο αλγόριθμος και οι χειριστές του. Οι διακινδυνεύσεις που βίωσαν οι πολίτες εξαιτίας του μεροληπτικού μαύρου κουτιού του αυτοματοποιημένου συστήματος δεν περιορίστηκαν μονάχα σε οικονομικές απώλειες αλλά και σε οικογενειακές, με πολλούς ανθρώπους να έρχονται αντιμέτωποι με ψυχολογικές συνέπειες αυτών των συνθηκών.

Ως επιστέγασμα από το προαναφερθέντα γίνεται κατανοητό ότι η αξιοποίηση συστημάτων τεχνητής νοημοσύνης δεν εγγυάται την αμεροληψία και την ουδετερότητα, παρά το γεγονός ότι συνήθως αυτή αξιοποιείται σε αυτά τα πλαίσια. Οι κριτικές σπουδές αλγορίθμων, χάρη στην επιρροή των θεωριών από το επιστημονικό πεδίο των σπουδών Επιστήμης και Τεχνολογίας και με την συμβολή των κριτικών προσεγγίσεων της κοινωνιολογίας, συμβάλλουν στην ολιστική προσέγγιση των συστημάτων τεχνητής νοημοσύνης, τα οποία ξεπερνούν την ιδιότητα του τεχνοεπιστημονικού. Υπό το πρίσμα των CAS, οι τεχνολογίες τεχνητής νοημοσύνης γίνονται αντιληπτές ως κοινωνικοτεχνικές κατασκευές. Αυτό διότι η συγκρότησή τους συνιστά μια κοινωνική διαδικασία, η οποία από το στάδιο εισαγωγής δεδομένων εντός αυτών έως και τα αποτελέσματα που παράγονται από την χρήση τους, επηρεάζει και επηρεάζεται από την κοινωνία. Τα δεδομένα αυτά, έχουν ως αποτέλεσμα οι Κοινωνίες της Τεχνητής Νοημοσύνης, να βιώνουν μια διαρκή αναπαραγωγή κοινωνικών προκαταλήψεων, στερεοτύπων και διακρίσεων εξαιτίας της ανεξέλεγκτης διακυβέρνησης με την χρήση τεχνολογιών τεχνητής νοημοσύνης οι οποίες καταλήγουν να αποτελούν μαύρα κουτιά . Στις Κοινωνίες του Μαύρου Κουτιού η αλόγιστη εφαρμογή αυτοματοποιημένων, μεροληπτικών πολιτικών καθώς και των διακινδυνεύσεων που (ανα)παράγονται από αυτές, καθιστά επιτακτική την αξιοποίηση μιας κοινωνιολογίας, η οποία αντιμετωπίζει κριτικά όλα τα παραπάνω ζητούμενα.

Προτάσεις

Η διευρυμένη αξιοποίηση τεχνολογιών τεχνητής νοημοσύνης και οι διακινδυνεύσεις οι οποίες παράγονται μέσω της μεροληπτικής αξιοποίησης και εφαρμογής τους έχουν γίνει αντιληπτές σε διεθνές επίπεδο. Η διακινδύνευση αξιοποίησης αυτοματοποιημένων συστημάτων εντοπισμού και ταξινόμησης ρίσκου όπως στο παράδειγμα «Toeslagenaffaire» ξεπερνούν τα σύνορα των χωρών και ηπείρων. Η Ευρωπαϊκή Ένωση, αντιλαμβανόμενη την ραγδαία αύξηση μεροληπτικών περιστατικών μέσω της μετατροπής συστημάτων τεχνητής νοημοσύνης σε μαύρα κουτιά, επιτάσσει την νομική ρύθμιση των συστημάτων ώστε οι τεχνολογίες αυτές να λειτουργούν κάτω από συνθήκες διαφάνειας (Aszódi, 2023 ; European Parliament [Press Releases], 2023; Deutsche Welle, 2023). Ωστόσο, από την πραγμάτευση που προηγήθηκε, γίνεται κατανοητό ότι παρά την νομοθετική ρύθμιση τέτοιων ζητημάτων, η φύση των μαύρων κουτιών των τεχνολογιών τεχνητής νοημοσύνης, παρακάμπτει τους κανόνες και τα πρωτόκολλα κάτω από τα οποία αυτές πρέπει να λειτουργούν, λαμβάνοντας μεροληπτικές αποφάσεις κατά αδύναμων κοινωνικών ομάδων.

Η Ευρωπαϊκή Ένωση βιώνει μια πρωτόγνωρη, ταχύρρυθμη εισχώρηση τεχνολογιών τεχνητής νοημοσύνης σε ιδιωτικό αλλά και δημόσιο τομέα. Από την βιβλιογραφική ανασκόπηση που πραγματοποιήθηκε για την εκπόνηση της παρούσας εργασίας, δεν εντοπίστηκε κάποιος σημαντικός αριθμός μελετών, ιδίως κοινωνιολογικών, οι οποίες να προσεγγίζουν ολιστικά τέτοιες τεχνολογίες και να ερευνούν τους τρόπους λειτουργίας τους. Στα πλαίσια αυτά, τέτοιες τεχνολογίες αξιοποιούνται και στην Ελλάδα, συγκροτώντας ένα γόνιμο πεδίο έρευνας και εξέτασης των αυτοματοποιημένων συστημάτων τεχνητής νοημοσύνης. Η κοινωνιολογική πραγμάτευση τέτοιων ζητημάτων μέσω των προσεγγίσεων των κριτικών σπουδών αλγορίθμων μπορούν να συνεισφέρουν στην αποτροπή και πρόληψη φαινομένων μεροληπτικής επιτήρησης και κοινωνικών διακρίσεων.

Η Ελλάδα, ιδίως μετέπειτα από την πανδημία του Covid-19, υιοθέτησε την εφαρμογή πολλών τεχνολογιών οι οποίες αξιοποιούν τις δυνατότητες της τεχνητής νοημοσύνης. Ο Χελιουδάκης, (2020) στην αναφορά της «Algorithm Watch» σχετικά με τις αυτοματοποιημένες κοινωνίες την περίοδο της πανδημίας ξεχωρίζει τρεις σχετικές περιπτώσεις. Αρχικά αναφέρεται στην περίπτωση του «COVID19 Symptom Checker» ενός συστήματος, το οποίο αξιοποιεί βιογραφικές πληροφορίες και ιατρικά

δεδομένα, με στόχο την κατάταξη του πολίτη σε ομάδες ανάλογα με τον κίνδυνο έκθεσής τους στον ιό. Ανάλογα με την ταξινόμησή τους, τα άτομα αυτά λαμβάνουν συμβουλές σχετικά με τον τρόπο που πρέπει να κινηθούν και να αντιμετωπίσουν σχετικά περιστατικά. Μια ακόμα περίπτωση που αναφέρει, αφορά την χρήση drone για την εποπτεία των φυσικών μετακινήσεων, στα πλαίσια της προστασίας των πολιτών από τον ιό.

Οι δύο παραπάνω περιπτώσεις εγείρουν πολλαπλά ερωτήματα σχετικά με τους τρόπους διαχείρισης των προσωπικών δεδομένων που αξιοποιούν οι συγκεκριμένες τεχνολογίες καθώς και για το αν αυτές λειτουργούν ουδέτερα. Ειδικότερα, το σύστημα του «COVID19 Symptom Checker» μπορεί να εξεταστεί σχετικά με το ποιες παραμέτρους λάμβανε υπόψη στην διαδικασία ταξινόμησης καθώς και για το αν οι προτάσεις του ενώ λειτουργούσαν στα πλαίσια προστασίας των πολιτών από τον ιό, ταυτόχρονα περιθωριοποιούσαν ορισμένες κοινωνικές ομάδες μέσω του στιγματισμού τους ως επικίνδυνες για την δημόσια υγεία. Η τεχνητή νοημοσύνη μπορεί πράγματι να συνδράμει στην προάσπιση της ανθρώπινης ζωής σε περιόδους υγειονομικών κινδύνων. Ωστόσο, με την ίδια ευκολία μπορεί να αποβεί μοιραία μέσω των μεροληπτικών αποφάσεων που αυτή μπορεί να λάβει, όπως στην περίπτωση αξιοποίησης ψηφιακής εφαρμογής στην πόλη Guangzhou στην Κίνα, η οποία δίχως σαφή τεκμήρια ταξινόμησε ως «επικίνδυνους» αφρικανικούς πληθυσμούς (Brevini & Pasquale, 2020).

Η αξιοποίηση drone στα πλαίσια διακυβέρνησης επιτάσσει πέρα από νομικής και κοινωνιολογικής εξέτασης των χρήσεών της. Όπως όλες οι διακινδυνεύσεις έτσι και η πανδημία του Covid-19 συνέβαλε στην συγκρότηση νέων μορφών κινδύνων και μαύρων τεχνολογικών κουτιών. Με πρόφαση την επίβλεψη των μετακινήσεων των πολιτών κατά την περίοδο της πανδημίας τα drone της Ελληνικής Αστυνομίας μπορούσαν να πραγματοποιούν πτήσεις και να καταγράφουν μέσω καμερών υψηλής τεχνολογίας. Η μελέτη του οργανισμού «Homo Digitalis» (2020), πέρα από την άγνοια σχετικά με την διαχείριση και αξιοποίηση των προσωπικών δεδομένων, αναδεικνύει το νομικό πλαίσιο πίσω από την εφαρμογή πολιτικών επιτήρησης μέσω του οποίου οι τεχνολογίες αυτές μπορούν δίχως την απαραίτητη, προσεκτική εφαρμογή τους να μετατραπούν σε μαύρα κουτιά.

Ενώ λοιπόν τα νομικά πλαίσια σχετικά με την αξιοποίηση των drone δεν λειτουργούν στα πλαίσια ενός μαύρου κουτιού, δίχως απαραίτητα να σέβονται τα προσωπικά δεδομένα των πολιτών, η Ευρωπαϊκή Ένωση προτάσσει την αυτοματοποίηση των συνόρων μέσω τεχνολογιών όπως αυτή του «ROBORDER». Σε αυτήν, πέρα από την αξιοποίηση drone θα παρέχονται δυνατότητες ανίχνευσης για τον έγκαιρο εντοπισμό εγκληματικών δραστηριοτήτων και επικίνδυνων περιστατικών καθώς και παροχή γρήγορης επισκόπησης της κατάστασης με μια ματιά στους χειριστές, υποστηρίζοντας την αυτοματοποιημένη λήψη αποφάσεων. Ενώ το σύστημα αυτό εφαρμόζεται ακόμα πιλοτικά, χρίζει κοινωνιολογικής εξέτασης καθώς η αλόγιστη αξιοποίησή του μπορεί να οδηγήσει στην ανάπτυξη αυτού που η Zuboff (2019) και ο Πασκάλ (2016), χαρακτηρίζουν ως «Κοινωνίες της Επιτήρησης».

Συμπερασματικά, η Ευρωπαϊκή Ένωση, κατ' επέκταση και η Ελλάδα, βιώνουν την έντονη επανάσταση της τεχνητής νοημοσύνης καθώς και των συστημάτων ταξινόμησης και βαθμολόγησης που αυτές στελεχώνουν και ενισχύουν. Η αξιοποίηση τέτοιων τεχνολογιών στις (μετα)νεωτερικές κοινωνίες αποτελεί μια αναπόφευκτη συνθήκη. Η εφαρμογή τους καθιστά επιτακτική την ρύθμιση σχετικά με την διαφανή λειτουργία αυτών των συστημάτων. Η εξέταση των κοινωνικοτεχνικών συστημάτων που προτάθηκαν προηγουμένως, καθώς και νέων τεχνολογιών που θα αναδυθούν στο μέλλον, μέσω της σκοπιάς της Κοινωνιολογίας της Τεχνητής Νοημοσύνης, μπορεί να λειτουργήσει αποτρεπτικά των μεροληψιών, διακρίσεων και της επιτήρησης που παράγονται εξαιτίας της αλόγιστης και ωφελμιστικής χρήσης τους.

Βιβλιογραφία

Ελληνική Βιβλιογραφία

Βαμβακά, Ε., Λογοθέτη, Α., Κροντήρης, Γ., Τερζίδου, Κ., Κακαβούλης, Κ., Χελιουδάκης, Λ., Μπάκα, Μ., & Ζαχαροπούλου, Μ. (2020). COVID-19 & Ψηφιακά Δικαιώματα Στην Ελλάδα. In <https://www.homodigitalis.gr/>. Homo Digitalis. Retrieved April 28, 2023, from https://www.homodigitalis.gr/wp-content/uploads/2020/04/HomoDigitalis_Report_COVID19_and_Digital_Rights_in_Greece_22.04.2020_Final.pdf

Bauer, E., Meleghy, Preglau, Niedenzu, Staubmann, & Morel. (2014). Κοινωνιολογική Θεωρία: Επισκόπηση των βασικών εκπροσώπων της (Ν. Ναγόπουλος, Επιμ.; 1st ed.). Προπομπός.

Beck, U. (2015). *Κοινωνία της Διακινδύνευσης: Καθ' οδόν προς μια άλλη νεωτερικότητα* (Ν. Πατηνιώτης, Επιμ.). Πεδίο.

Sismondo, S. (2016). *Εισαγωγή στις Σπουδές Επιστήμης και Τεχνολογίας* (Μ. Πατηνιώτης, Επιμ.). Liberal Books.

Γεωργοπούλου, Π. (2010). *Η στροφή της Κοινωνικής σκέψης στις επιστήμες της πολυπλοκότητας* (1η έκδοση). Εκδόσεις Κριτική.

Μανχαϊμ, Κ. (1997). *Ιδεολογία Και Ουτοπία* (Π. Κονδύλης, Επιμ.; Γ. Ανδρουλιδάκης, Trans.; 1η). Εκδόσεις «Γνώση»

Ναγόπουλος, Ν. (2015). *Γνώση Μέθοδος Και Κοινωνική Πράξη : Από τη Γνωσιοθεωρία στην Κοινωνιολογία της Γνώσης*. Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών.

Κάλλας, Γ. (2006). *Η Κοινωνία της Πληροφορίας και ο Νέος Ρόλος των Κοινωνικών Επιστημών*. ΝΕΦΕΛΗ.

Κουρούτζας, Χ. (2018). *Εγκληματολογία της Γενετικής*. Πεδίο.

Ξενόγλωσση Βιβλιογραφία

- Airoidi, M. (2021). *Machine Habitus: Toward a Sociology of Algorithms*. John Wiley & Sons.
- Amnesty International. (2021, October 25). *Xenophobic machines: Discrimination through unregulated use of algorithms in the Dutch childcare benefits scandal - Amnesty International*. <https://www.amnesty.org/en/documents/eur35/4686/2021/en/>
- Aszódi, N. (2023, June 15). *EU Parliament vote on AI Act: Lawmakers chose to protect people against harms of AI systems - AlgorithmWatch*. AlgorithmWatch. Retrieved June 15, 2023, from <https://algorithmwatch.org/en/eu-parliament-vote-aiact-june-2023/>
- Azam, N. (2023). A brief Review of ChatGPT: Limitations, Challenges and Ethical-Social Implications. *Zenodo (CERN European Organization for Nuclear Research)*. <https://doi.org/10.5281/zenodo.7629888>
- Baron, L. F., & Gomez, R. S. (2016). The Associations between Technologies and Societies: The Utility of Actor-Network Theory. *Science Technology & Society*, 21(2), 129–148. <https://doi.org/10.1177/0971721816640615>
- Barthle, C. (2023, May 4). *Humans + Machines: A Look Behind Spotify's Algorithmic Playlists - Spotify Engineering*. Spotify Engineering. <https://engineering.atspotify.com/2023/04/humans-machines-a-look-behind-spotifys-algorithmic-playlists/>
- Beaulieu, A., & Leonelli, S. (2021). *Data and Society: A Critical Introduction*. SAGE.
- Beck, S. U. (1992). *Risk Society: Towards a New Modernity*. SAGE.
- Beer, D. G. (2017). The social power of algorithms. *Information, Communication & Society*, 20(1), 1–13. <https://doi.org/10.1080/1369118x.2016.1216147>
- Benjamin, R. (2016). Catching Our Breath: Critical Race STS and the Carceral Imagination. *Engaging Science, Technology, and Society*, 2, 145–156. <https://doi.org/10.17351/ests2016.70>
- Benjamin, R. (2019). *Race After Technology: Abolitionist Tools for the New Jim Code*. John Wiley & Sons.
- Berger, S. L., & Luckmann, T. (1967). *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Anchor.

- Bhageshpur, K. (2019, November 15). Data Is The New Oil -- And That's A Good Thing. *Forbes*. <https://www.forbes.com/sites/forbestechcouncil/2019/11/15/data-is-the-new-oil-and-thats-a-good-thing/>
- Bieri, Σ. (1988). Thinking Machines: Some Reflections on the Turing Test. *Poetics Today*. <https://doi.org/10.2307/1772893>
- Bijker, W. E. (2010). How is technology made?--That is the question! *Cambridge Journal of Economics*, 34(1), 63–76. <https://doi.org/10.1093/cje/bep068>
- Bijker, W. E., Hughes, T. Σ., & Pinch, T. J. (1989). *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*. MIT Press.
- Bijker, W. E., & Law, J. (1992). *Shaping Technology/building Society: Studies in Sociotechnical Change*. The MIT Press.
- Blok, A., & Jensen, T. E. (2011). *Bruno Latour: Hybrid Thoughts in a Hybrid World*. Routledge.
- Blum, A. M., Hazlett, C., & Posner, D. N. (2021). Measuring Ethnic Bias: Can Misattribution-Based Tools from Social Psychology Reveal Group Biases that Economics Games Cannot? *Political Analysis*, 29(3), 385–404. <https://doi.org/10.1017/pan.2020.37>
- Boztas, S. (2021, May 12). *Good lawyers, jobs and their good names back: child benefit victims don't want just cash - DutchNews.nl*. DutchNews.nl. Retrieved June 2, 2023, from <https://www.dutchnews.nl/2021/05/good-lawyers-jobs-and-their-good-names-back-child-benefit-victims-dont-want-just-cash/>
- Brevini, B., & Pasquale, F. A. (2020). Revisiting the Black Box Society by rethinking the political economy of big data. *Big Data & Society*, 7(2), 205395172093514. <https://doi.org/10.1177/2053951720935146>
- Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification [Proceedings of Machine Learning Research 81:1–15,2018]. <http://gendershades.org/>, 77–91. <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>

- Cassauwers, T. (2020, December 1). *Opening the 'black box' of artificial intelligence*. Horizon Magazine. <https://ec.europa.eu/research-and-innovation/en/horizon-magazine/opening-black-box-artificial-intelligence>
- Castells, M. (2002). *The Internet Galaxy: Reflections on the Internet, Business, and Society*. Oxford University Press, USA.
- Castelvecchi, D. (2016). Can we open the black box of AI? *Nature*, 538(7623), 20–23. <https://doi.org/10.1038/538020a>
- Chelioudakis, E. (2020). *Automated Decision-Making Systems in the COVID-19 Pandemic: A European Perspective/ Automating Society Report 2020*. Algorithm Watch. Retrieved May 5, 2023, from <https://algorithmwatch.org/en/wp-content/uploads/2020/08/ADM-systems-in-the-Covid-19-pandemic-Report-by-AW-BSt-Sept-2020.pdf>
- Chow, A. R., & Perrigo, B. (2023, February 17). The AI Arms Race Is Changing Everything. *Time*. <https://time.com/6255952/ai-impact-chatgpt-microsoft-google/>
- Cirillo, D., Solarz, S. C., & Guney, E. (2022). *Sex and Gender Bias in Technology and Artificial Intelligence: Biomedicine and Healthcare Applications*. Academic Press.
- Coletta, C., & Kitchin, R. (2017). Algorhythmic governance: Regulating the 'heartbeat' of a city using the Internet of Things. *Big Data & Society*, 4(2), 205395171774241. <https://doi.org/10.1177/2053951717742418>
- Collins, H. (1991). Artificial experts: social knowledge and intelligent machines. *Choice Reviews Online*, 28(11), 28–6285. <https://doi.org/10.5860/choice.28-6285>
- Collins, H. M. (1983). The Sociology of Scientific Knowledge: Studies of Contemporary Science. *Annual Review of Sociology*, 9(1), 265–285. <https://doi.org/10.1146/annurev.so.09.080183.001405>
- Comte, A. (1853). *The Positive Philosophy of Auguste Comte*.
- Coser, L. A. (1977). *Masters of Sociological Thought: Ideas in Historical and Social Context*. Houghton Mifflin Harcourt Σ.
- Crawford, K. (2021). *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.

- Darwin, C. (2003). *On the Origin of Species*. Broadview Press.
- Dastin, J. (2018, October 10). Amazon scraps secret AI recruiting tool that showed bias against women. *U.S.* <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>
- Deutsche Welle. (2023, June 14). EU lawmakers lay groundwork for “historic” AI regulation. *dw.com*. Retrieved June 15, 2023, from <https://www.dw.com/en/eu-lawmakers-lay-groundwork-for-historic-ai-regulation/a-65909881>
- D’Ignazio, C., & Klein, L. F. (2020). *Data Feminism*. MIT Press.
- Dolan, S. (2023). How mobile users spend their time on their smartphones in 2023. *Insider Intelligence*. <https://www.insiderintelligence.com/insights/mobile-users-smartphone-usage/>
- Dutch Data Protection Office. (2020). De verwerking van de nationaliteit van aanvragers van kinderopvangtoeslag (. In *www.rijksoverheid.nl*. Retrieved June 1, 2023, from <https://www.rijksoverheid.nl/documenten/rapporten/2020/07/17/de-verwerking-van-de-nationaliteit-van-aanvragers-van-kinderopvangtoeslag>
- Ensmenger, N. (2012, February). *Is chess the drosophila of artificial intelligence? A social history of an algorithm on JSTOR*. Retrieved February 3, 2023, from <https://www.jstor.org/stable/23210226>
- Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin’s Press.
- European Parliament [Press Releases]. (2023, November 5). *AI Act: a step closer to the first rules on Artificial Intelligence | News | European Parliament*. European Parliament. Retrieved June 15, 2023, from <https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence>
- Eynon, R., & Young, E. E. (2021). Methodology, Legend, and Rhetoric: The Constructions of AI by Academia, Industry, and Policy Groups for Lifelong Learning. *Science, Technology, & Human Values*, 46(1), 166–191. <https://doi.org/10.1177/0162243920906475>

- Floridi, L., Cowsls, J., Beltrametti, M., Chatila, R., Chazerand, Σ., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, Σ., & Vayena, E. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Geburu, T. (2020). Race and Gender. *The Oxford Handbook of Ethics of AI*, 251–269. <https://doi.org/10.1093/oxfordhb/9780190067397.013.16>
- Giddens, A. (1991). *Modernity and Self-Identity: Self and Society in the Late Modern Age*. Stanford University Press.
- Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media*. Yale University Press.
- Glover, E. (2022). What Is AI Winter? *Built In*. <https://builtin.com/artificial-intelligence/ai-winter>
- Heikkilä, M. (2022, April 13). Dutch scandal serves as a warning for Europe over risks of using algorithms. *POLITICO*. <https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/>
- Hey, A. J. G., Tansley, S., & Tolle, K. M. (2009). *The Fourth Paradigm: Data-intensive Scientific Discovery*.
- Hills, R. L. (2002). *James Watt: His time in Scotland, 1736-1774*.
- Jordan, M. I. (2019). Artificial Intelligence—The Revolution Hasn't Happened Yet. *Harvard Data Science Review*. <https://doi.org/10.1162/99608f92.f06c6e61>
- Joyce, K., Smith-Doerr, L., Alegria, S., Bell, S. S., Cruz, T. M., Hoffman, S., Noble, S. U., & Shestakofsky, B. (2021). Toward a Sociology of Artificial Intelligence: A Call for Research on Inequalities and Structural Change. *Socius*, 7, 237802312199958. <https://doi.org/10.1177/2378023121999581>
- Juillet, R. (2022, October 27). *What Are The Storage Methods For Big Data?* Bocasay. <https://www.bocasay.com/how-to-store-big-data/>
- Kazemier, B. H. (1946). Remarks on Logical Positivism. *Synthese*, 5(7/8), 327–332. <http://www.jstor.org/stable/20113894>

- Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1), 205395171452848. <https://doi.org/10.1177/2053951714528481>
- Knorr-Cetina, K. (2013). *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science*. Elsevier.
- Kuhn, T. S. (1968). *The Structure of Scientific Revolutions*.
- Latour, B. (1996). On actor-network theory: A few clarifications. *Nomos Verlagsgesellschaft mbH*, 47(4), 369–381. <http://www.jstor.org/stable/40878163>
- Latour, B. (1999). *Pandora's Hope: Essays on the Reality of Science Studies*. Harvard University Press.
- Latour, B., & Woolgar, S. (1986). *Laboratory Life: The Construction of Scientific Facts*. Princeton University Press.
- Law, J. (2008). On Sociology and STS. *The Sociological Review*, 56(4), 623–649. <https://doi.org/10.1111/j.1467-954x.2008.00808.x>
- Leurs, K., & Shepherd, T. (2017). 15. Datafication & Discrimination. In *Amsterdam University Press eBooks* (σσ. 211–232). <https://doi.org/10.1515/9789048531011-018>
- Lindsey, L. L. (2021). *Gender: Sociological Perspectives* (7th ed.). Routledge.
- Liu, J. (2022). Social data governance: Towards a definition and model. *Big Data & Society*, 9(2), 205395172211113. <https://doi.org/10.1177/20539517221111352>
- Liu, Z. (2021). Sociological perspectives on artificial intelligence: A typological reading. *Sociology Compass*, 15(3). <https://doi.org/10.1111/soc4.12851>
- Merton, R. K. (1973). *The Sociology of Science: Theoretical and Empirical Investigations*. University of Chicago Press.
- Merton, R. K. (1976). *Sociological Ambivalence and Other Essays*. Simon and Schuster.
- Milmo, D. (2023, March 10). Google v Microsoft: who will win the AI chatbot race? *The Guardian*. <https://www.theguardian.com/technology/2023/feb/10/google-v-microsoft-who-will-win-the-ai-chatbot-race-bard-chatgpt>

Mutzner, N. (2020). *Why not a Sociology of Artificial Intelligence? A Social Approach to the Potential Symbiosis of Sociology and AI Research*. [Thesis]. University of Zurich.

Nl.Times. (2021, October 19). Over 1,100 children taken from homes of benefits scandal victims. *NL Times*. Retrieved March 20, 2023, from <https://nltimes.nl/2021/10/19/1100-children-taken-homes-benefits-scandal-victims>

Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.

Ntoutsis, E., Fafalios, Σ., Gadiraju, U., Iosifidis, V., Nejdil, W., Vidal, M., Ruggieri, S., Turini, F., Papadopoulou, S., Krasanakis, E., Kompatsiaris, I., Kinder-Kurlanda, K., Wagner, C., Karimi, F., Fernandez, M., Alani, H., Berendt, B., Kruegel, T., Heinze, C., . . . Staab, S. (2020). Bias in data-driven artificial intelligence systems—An introductory survey. *WIREs Data Mining and Knowledge Discovery Published by Wiley Periodicals, Inc.*, 10(3). <https://doi.org/10.1002/widm.1356>

O’Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Penguin UK.

Pascoe, R. (2022, May 11). *Childcare benefit scandal: more children were removed from their homes - DutchNews.nl*. DutchNews.nl. Retrieved May 15, 2023, from <https://www.dutchnews.nl/2022/05/childcare-benefit-scandal-more-children-were-removed-from-their-homes/>

Pasquale, F. A. (2016). The Black Box Society: The Secret Algorithms That Control Money and Information. *Contemporary Sociology*, 45(3), 367–368. <https://doi.org/10.1177/0094306116641409c>

Pinch, T. (1992). Opening Black Boxes: Science, Technology and Society. *Social Studies of Science*, 22(3), 487–510. <https://doi.org/10.1177/0306312792022003003>

Platonova, Anastasia. (n.d.). *Artificial Intelligence: A New Trojan Horse for Undue Influence on Judiciaries?* Retrieved May 23, 2023, from https://www.unodc.org/dohadeclaration/en/news/2019/06/artificial-intelligence_-a-new-trojan-horse-for-undue-influence-on-judiciaries.html

Puntel, L. B. (1999). On the Logical Positivists’ Theory of Truth: The Fundamental Problem and a New Perspective on JSTOR. In *Journal for General Philosophy of*

Science / Zeitschrift für allgemeine Wissenschaftstheorie (30th ed., Vol. 1, pp. 101–130). <https://www.jstor.org/stable/25171138>

Raji, I. D., Gebru, T., Mitchell, M., Buolamwini, J., Lee, J., & Denton, E. (2020). *Saving Face*. <https://doi.org/10.1145/3375627.3375820>

Reiners, B. (2021). What Is Gender Bias in the Workplace? *Built In*. <https://builtin.com/diversity-inclusion/gender-bias-in-the-workplace>

Ross, Σ. N., & Maynard, K. (2021). Towards a 4th industrial revolution. *Intelligent Buildings International*, 13(3), 159–161. <https://doi.org/10.1080/17508975.2021.1873625>

Rudin, C., & Radin, J. (2019). Why Are We Using Black Box Models in AI When We Don't Need To? A Lesson From An Explainable AI Competition. *Harvard Data Science Review*, 1(2). <https://doi.org/10.1162/99608f92.5a8a3a3d>

Ruitt, E. R. T. (2021, November 24). *Surveillance, Companionship, and Entertainment: The Ancient History of Intelligent Machines*. The MIT Press Reader. <https://thereader.mitpress.mit.edu/the-ancient-history-of-intelligent-machines/>

Runting, R. K., Phinn, S. R., Xie, Z., Venter, O., & Watson, J. E. M. (2020). Opportunities for big data in conservation and sustainability. *Nature Communications*, 11(1). <https://doi.org/10.1038/s41467-020-15870-0>

Schuchmann, S. (2021, December 10). History of the first AI Winter - Towards Data Science. *Medium*. <https://towardsdatascience.com/history-of-the-first-ai-winter-6f8c2186f80b>

Schwab, K. (2023, May 31). *The Fourth Industrial Revolution | Essay by Klaus Schwab*. Encyclopedia Britannica. Retrieved June 1, 2023, from <https://www.britannica.com/topic/The-Fourth-Industrial-Revolution-2119734>

Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). *Fairness and Abstraction in Sociotechnical Systems*. <https://doi.org/10.1145/3287560.3287598>

Seldam, B., & Brenninkmeijer, A. (2021, April 30). *The Dutch benefits scandal: a cautionary tale for algorithmic enforcement*. <https://eulawenforcement.com/>. Retrieved June 1, 2023, from

<https://eulawenforcement.com/?s=The+Dutch+benefits+scandal%3A+a+cautionary+tale+for+algorithmic+enforcement>.

Silva, S., & Kenney, M. (2018.). Algorithms, Platforms, and Ethnic Bias: An Integrative Essay. *Phylon*, 55(1 & 2), 9–37. <https://www.jstor.org/stable/26545017>

Sinur, J. (2019, September 30). AI & Big Data; Better Together. *Forbes*. <https://www.forbes.com/sites/cognitiveworld/2019/09/30/ai-big-data-better-together/>

Stephan De Spiegeleire, Matthijs Maas and Tim Sweijjs. (2017). *What Is Artificial Intelligence? From Artificial Intelligence And The Future Of Defense: Strategic Implications For Small- And Medium-Sized Force Providers On Jstor*. <https://www.jstor.org/stable/resrep12564.7>

Suresh, H., & Guttag, J. V. (2021). Understanding Potential Sources of Harm throughout the Machine Learning Life Cycle. *Computer Science and Artificial Intelligence Laboratory, MIT*. <https://doi.org/10.21428/2c646de5.c16a07bb>

Szaniawska-Schiavo, G. (2023). Love in the Age of AI Dating Apps [2023 Statistics]. *Tidio*. <https://www.tidio.com/blog/ai-dating-apps/>

Team, U. (2020, August 14). *Machine Learning for Big Data*. Udacity. Retrieved February 14, 2023, from <https://www.udacity.com/blog/2020/08/machine-learning-for-big-data.html>.

The Editors of Encyclopaedia Britannica. (2020, June 17). *Law of three stages | Comte, Definition, & Significance*. Encyclopedia Britannica. Retrieved January 15, 2023, from <https://www.britannica.com/topic/law-of-three-stages>.

Turing, A. M. (2004). Computing Machinery and Intelligence (1950). In *Oxford University Press eBooks*. <https://doi.org/10.1093/oso/9780198250791.003.0017>

Verdegem, Σ. (2021). *AI for Everyone?: Critical Perspectives*. University of Westminster Press.

Wehbé, A. (2017, May 5). *OPM Data Breach Case Study: Mitigating Personnel Cybersecurity Risk*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2964071

Weidhorn, M. (2001). The Great Paradigm Shift. *Journal of Thought*, 36(3), 7–19. <http://www.jstor.org/stable/42589665>

WHAT IS ARTIFICIAL INTELLIGENCE? from ARTIFICIAL INTELLIGENCE AND THE FUTURE OF DEFENSE: STRATEGIC IMPLICATIONS FOR SMALL- AND MEDIUM-SIZED FORCE PROVIDERS on JSTOR. (n.d).
<http://www.jstor.com/stable/resrep12564.7>

Winner, L. (1993). Upon Opening the Black Box and Finding It Empty: Social Constructivism and the Philosophy of Technology. *Science, Technology, & Human Values*, 18(3), 362–378. <https://doi.org/10.1177/016224399301800306>

Woolgar, S. (1985). Why not a Sociology of Machines? The Case of Sociology and Artificial Intelligence. *Sociology*, 19(4), 557–572.
<https://doi.org/10.1177/0038038585019004005>

Yonego, J. T. (2015, August 7). Data Is the New Oil of the Digital Economy. *WIRED*.
<https://www.wired.com/insights/2014/07/data-new-oil-digital-economy/>

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Profile Books.