



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Αλγόριθμοι και Ηθική

Επιλεγμένα Θέματα

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Μαρία Φωστήνη

Επιβλέπων: Πέτρος Στεφανέας

Επ. Καθηγητής ΕΜΠ

Αθήνα, Ιανουάριος 2021



Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών
Υπολογιστών

Τομέας Τεχνολογίας Πληροφορικής και Υπολογιστών

Αλγόριθμοι και Ηθική

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Μαρία Φωστήνη

Επιβλέπων: Πέτρος Στεφανέας

Επίκουρος Καθηγητής ΕΜΠ

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 31η Ιανουαρίου 2021

(υπογραφή)

(υπογραφή)

(υπογραφή)

Πέτρος Στεφανέας
Επίκουρος Καθηγητής
ΕΜΠ

Αριστείδης Παγουρτζής
Αναπληρωτής Καθηγητής

Σταφυλοπάτης
Ανδρέας - Γεώργιος
Καθηγητής

Αθήνα, Ιανουάριος 2021



Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών
Υπολογιστών

Τομέας Τεχνολογίας Πληροφορικής και Υπολογιστών

Copyright ©—All rights reserved Μαρία Φωστήνη, 2021. Με την επιφύλαξη παντός δικαιώματος.

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

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

(Υπογραφή)

.....

Μαρία Φωστήνη

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Περίληψη

Η σημερινή κοινωνία διαμορφώνεται σε μεγάλο βαθμό και επηρεάζεται από την εργασία αλγορίθμων σε πολλαπλά επίπεδα. Σχεδόν πανταχού παρόν σε οποιαδήποτε δεδομένα, λειτουργίας ή επεξεργασίας, οι αλγόριθμοι διαδραματίζουν σημαντικό ρόλο σε διάφορα θέματα που καλύπτουν σχεδόν όλα όσα συμβαίνουν στο Internet και στις ευρύτερες επιστήμες υπολογιστών, συλλογή κυβερνητικών δεδομένων το επίπεδο κατανόησης της λειτουργίας τους. Αυτή η διατριβή στοχεύει να δώσει μια εικόνα για τα αποτελέσματα των αλγοριθμικών εφαρμογών και να διαλευκάνει το «μαύρο κουτί» των λειτουργιών τους αμφισβητώντας την κατάστασή τους ως ουδέτερα εργαλεία ή τεχνολογίες. Αυτή η προσέγγιση θα ληφθεί ειδικά εστιάζοντας στις ηθικές επιπτώσεις των αλγορίθμων, οι οποίες θα αξιολογηθούν σε δύο ενότητες που αντιστοιχούν στις ηθικές παραδόσεις του συνεπειασκού χαρακτήρα και της δεοντολογίας. Το πρώτο από τα δύο αυτά κεφάλαια, θα διευκρινίσει πώς οι τρέχουσες εφαρμογές αλγορίθμων συμπεριλαμβανομένων των προαναφερθέντων τεχνολογιών και πεδίων που προωθούν — έχουν αρνητικές συνέπειες. Σε αυτή την ενότητα, θα εισαχθεί ένα θεωρητικό πλαίσιο που περιγράφει τους δομικούς τρόπους με τους οποίους η πρακτική αλγοριθμική λειτουργία παράγει δυσμενή αποτελέσματα, μετά από τα οποία θα δοθούν ορισμένα παραδείγματα για την απεικόνιση της τρέχουσας κατάστασης και των συνακόλουθων πραγματικών ηθικών ζητημάτων. Μετά από αυτό, θα παρασχεθούν διάφοροι τρόποι αντιμετώπισης αυτής της δυναμικής, ιδίως προσανατολισμένοι σε μέτρα για την αύξηση της ευαισθητοποίησης, της λογοδοσίας και της εξισορρόπησης των σχέσεων ισχύος. Το δεύτερο από τα δύο αυτά κεφάλαια θα γίνει διεύρυνση της ηθικής διάστασης των αλγορίθμων από την σκοπιά της δεοντολογίας, θα απομακρύνει την ανάλυση από τις πρακτικές ανησυχίες, προς την αξιολόγηση της επιλογής για την αποδοχή των απαιτούμενων συνθηκών για την αλγοριθμική επεξεργασία βασικά. Δηλαδή, για να είναι σε θέση οι αλγόριθμοι να λειτουργούν σωστά, χρειάζεται να υιοθετηθεί ένα συγκεκριμένο πρίσμα μελέτης ή μια γνωσιολογική στάση, η οποία από μόνη της μπορεί να είναι ηθικά αμφισβητήσιμη. Θα δηλωθούν οι σχετικές προϋποθέσεις, που αποτελούν τη θεωρητική συνιστώσα, μετά από την οποία θα εισαχθούν και πάλι ορισμένα παραδείγματα για την απεικόνιση της καταλληλότητας αυτής της διάστασης. Με αυτό, επιτυγχάνεται ο κύριος στόχος της ενημέρωσης της συνεχούς συζήτησης σε πολλαπλά επίπεδα για την αντιμετώπιση των τεθέντων ζητημάτων.

Λέξεις Κλειδιά

Αλγόριθμοι, Ηθική, Δεοντολογία, Κυβερνητική, Τεχνολογία Διπλής Χρήσης, Κοινωνιολογία, Κοινωνικά Δίκτυα, Μηχανές Αναζήτησης

Abstract

Present day society is heavily shaped and influenced by the workings of algorithms on multiple levels. Near-ubiquitous in any data-processing operation, algorithms play an important role in diverse topics ranging from virtually everything that is happening on the internet and in wider computer sciences, to health-care and medicine development, traffic (control), meteorological models, financial markets, educational settings. This thesis aims to give an idea of the results of algorithmic applications and to clarify the "black box" of their functions by questioning their status as neutral tools or technologies. This approach will be taken specifically by focusing on the ethical implications of the algorithms, which will be assessed in two sections corresponding to the ethical traditions of consistency and ethics. The first of these two chapters will explain how current algorithm applications, including the aforementioned technologies and fields they promote, have negative consequences. In this section, a theoretical framework will be introduced describing the structural ways in which practical algorithmic operation produces adverse effects, after which some examples will be given to illustrate the current situation and the consequent real ethical issues. After that, various ways will be provided to address this dynamic, in particular measures aimed at raising awareness, accountability and balancing power relations. The second of these two chapters will expand the ethical dimension of algorithms from an ethical point of view, move the analysis away from practical concerns, to evaluate the option of accepting the conditions required for algorithmic processing basically. That is, in order for algorithms to work properly, it is necessary to adopt a specific prism of study or an epistemological attitude, which in itself can be morally questionable. The relevant conditions, which constitute the theoretical component, will be stated, after which some examples will be introduced again to illustrate the appropriateness of this dimension. This achieves the main goal of updating the ongoing debate at multiple levels to address the issues raised.

Key Words

Algorithms, Ethics, Deontology, Cyberethics, Dual Use Technologies, Sociology, Social Networks, Search Engines

Ευχαριστίες

Με την ολοκλήρωση της διπλωματικής μου εργασίας, θα ήθελα να ευχαριστήσω όλους όσους συνέβαλαν με οποιοδήποτε τρόπο στην επιτυχή εκπόνηση της.

Πρώτα απ' όλα θα ήθελα να ευχαριστήσω τον καθηγητή κ. Πέτρο Στεφανέα για την επίβλεψη αυτής της διπλωματικής εργασίας, για την υπομονή που έκανε, την ανοχή που έδειξε και κυρίως στην κατανόηση της κατάστασης και την ηθική υποστήριξη που μου προσφερε.

Τον καθηγητή κ. Στάθη Ζάχο που ήταν πάντα εκεί να μου θυμίζει ότι κάποτε υπήρχε κάτι σε αυτή τη σχολή που μου άρεσε. Την αδερφή μου και τους γονείς μου που ήταν το στήριγμα μου όλα αυτά τα χρόνια, παρόλες τις εντάσεις μας. Και τέλος τους κ. Ηλια Αγγελόπουλο και τον αγαπημένο μου θείο κ. Γκιώνη.

Στους γονείς μου

Ελένη & Μανώλη

Περιεχόμενα

Κεφάλαιο 1

Αλγόριθμοι

- 1.1 Σύντομη ιστορική επισκόπηση
- 1.2 Πως μπορούμε να εξηγήσουμε τη θεωρία αλγορίθμων
 - 1.2.1 Τα χαρακτηριστικά των αλγορίθμων
 - Παράδειγμα - Ο αλγόριθμος quicksort:
 - 1.2.2 Πιο σύγχρονα ερωτήματα
 - 1.2.3 Αλγόριθμοι & Προγράμματα
 - 1.2.4 Αποδοτικότητα
- 1.3 Ζητήματα Φιλοσοφίας του Υπολογισμού
 - 1.3.1 Η θέση Church - Turing (thesis)
 - Υπολογισιμότητα
 - Εκδοχές της Church - Turing Thesis
 - Ισχυρή εκδοχή της θέσης Church-Turing:
 - 1.3.2 Πέρα από τις Μηχανές Turing
 - 1.3.3 Υπολογιστική Πολυπλοκότητα

Κεφάλαιο 2

Ηθική

- 2.1 Σύντομη Ιστορία
 - 2.1.1 Προέλευση της Ηθικής
 - 2.1.2 Οι Σοφιστές
 - 2.1.3 Ο Σωκράτης
 - 2.1.4 Ο Πλάτων
 - 2.1.5 Η Αριστοτελική Ηθική
 - Η έννοια του αγαθού στον Αριστοτέλη
 - Η αρετή στον Αριστοτέλη
 - Η διανοητική αρετή στον Αριστοτέλη
 - Η ευδαιμονία στον Αριστοτέλη
 - Η δικαιοσύνη στον Αριστοτέλη
 - Η ανισότητα στον Αριστοτέλη
 - Η μεσότητα στον Αριστοτέλη
 - 2.1.6 Ο Kant
 - Η Μεταφυσική των Ηθών ως Θεμέλιο της Ηθικής

[2.1.7 Η Αναλυτική φιλοσοφία](#)

[Ο G.E Moore](#)

[Principia Ethica](#)

[2.2 Cybernetics - Η Κυβερνηθική](#)

[2.2.1 Ο Nobert Wiener - The human use of human beings](#)

[2.2.2 Θεωρία Πληροφορίας & Στοιχεία από το Shannon](#)

[2.2.3 Εξέλιξη και Εντροπία](#)

[2.3 Ηθική και Τεχνητή Νοημοσύνη](#)

[2.3.1 ELIZA](#)

[2.4 Η Κυβερνηθική στις μέρες μας](#)

[2.5 Συμπεράσματα](#)

[Κεφάλαιο 3](#)

[Εφαρμογές και Επιπτώσεις](#)

[3.1 Η Συνεπειοκρατική Κανονιστική Ηθική](#)

[3.2 Η Τεχνολογία Διπλής Χρήσης](#)

[3.3 Μερικές Ηθικές Ανησυχίες Σχετικά με Αλγορίθμους](#)

[3.3.1 Μοτίβα των ηθικών πτυχών των αλγορίθμων](#)

[3.3.2 Πτυχές του Gillespie](#)

[Μοτίβα ένταξης:](#)

[Κύκλοι πρόβλεψης:](#)

[Η αξιολόγηση της συνάφειας:](#)

[Υπόσχεση αλγοριθμικής αντικειμενικότητας:](#)

[Δέσμευση με πρακτικές:](#)

[Παραγωγή υπολογισμένης πολιτικής:](#)

[3.4 Οι Περιπτώσεις των Ηθικών Πτυχών των Αλγορίθμων](#)

[3.4.1 Συνάφεια: Μηχανές Αναζήτησης](#)

[3.4.2 Κοινωνικά Δίκτυα](#)

[3.4.3 Η Αλγοριθμική Ιθαγένεια](#)

[3.4.4 Η Αναγνώριση Προσώπων](#)

[3.5 Αντιμετώπιση των συνεπειών](#)

[3.6 Συμπεράσματα](#)

[Κεφάλαιο 4](#)

[Αλγοριθμική Επιλογή](#)

[4.1 Δεοντολογία](#)

[4.2 Τα χαρακτηριστικά των αλγορίθμων από την σκοπιά της Δεοντολογίας](#)

[4.2.1 Η καλά ορισμένη Είσοδος](#)

[4.2.2 Κωδικοποίηση και αυτοματοποίηση](#)

[4.2.3 Υπολογιστικός Θετικισμός](#)

[4.3 Περιπτώσεις Δεοντολογικών Ηθικών Θεμάτων](#)

[4.3.1 Η περίπτωση της Google](#)

[4.3.2 Αυτόνομα Οχήματα και Στρατιωτικά Drones.](#)

[4.3.3 Τεχνητή Νοημοσύνη και Robots](#)

Εισαγωγή

Η σύγχρονη κοινωνία βασίζεται και επηρεάζεται κυρίως από τους αλγόριθμους σε πολλαπλά επίπεδα. Σχεδόν πανταχού παρόν σε οποιαδήποτε λειτουργία επεξεργασίας δεδομένων, οι αλγόριθμοι διαδραματίζουν σημαντικό ρόλο σε διάφορα θέματα που κυμαίνονται από σχεδόν όλα όσα συμβαίνουν στο διαδίκτυο και στις ευρύτερες επιστήμες των υπολογιστών, έως την ανάπτυξη της υγειονομικής περίθαλψης και της ιατρικής, την κυκλοφορία (έλεγχος), τα μετεωρολογικά μοντέλα, τα οικονομικά αγορές, εκπαιδευτικά περιβάλλοντα κ.λπ.

Οι κύριες δραστηριότητες που λαμβάνουν χώρα σε αυτούς τους τομείς θα ήταν απλώς αδύνατες χωρίς αλγόριθμους και ορισμένα μέρη δεν θα είχαν αναπτυχθεί ποτέ. Κατά επέκταση, πολλά από τα αντικείμενα με τα οποία αλληλεπιδρούμε, όπως τα κινητά (έξυπνα) τηλέφωνα και οι υπολογιστές, βασίζονται στη λειτουργία αλγορίθμων με διάφορους τρόπους, δημιουργώντας μεγάλες ποσότητες δεδομένων τα οποία με τη σειρά τους επεξεργάζονται αλγοριθμικά. Η ανάπτυξη αυτών των αλγορίθμων και των εφαρμογών τους έχει ανοίξει πολλές νέες δυνατότητες μέσω της αύξησης της αποτελεσματικότητας, της ευκολίας χρήσης και της δημιουργίας και ανάλυσης μεγάλων συνόλων δεδομένων. Όλο και περισσότερο, οι αλγόριθμοι έχουν επίσης φτάσει να επηρεάζουν και να διαμορφώνουν τον τρόπο με τον οποίο ενημερώνουμε τους εαυτούς μας και με τη σειρά τους πώς να αξιολογούμε τις διαδικασίες που οι ίδιες εξαρτώνται από αλγοριθμικές διαδικασίες. Τα αλγοριθμικά ταξινομημένα αποτελέσματα αναζήτησης και τα πολύπλοκα στατιστικά μοντέλα είναι δύο κύρια παραδείγματα για το πώς όχι μόνο δημιουργούνται αλγοριθμικά μέρη του ίδιου του περιεχομένου, αλλά η κατάσταση και η λειτουργία του ταξινομούνται, αναλύονται και αξιολογούνται περαιτέρω από αλγόριθμους. Δεν είναι τυχαίο ότι τέτοια πεδία που ποικίλλουν βασίζονται σε αλγόριθμους, καθώς οι αλγόριθμοι είναι ένα ζωτικό δομικό στοιχείο που απαιτείται για τη λειτουργία τους. Δεν είναι θέμα (αυθαίρετης) επιλογής η χρήση αλγορίθμων, καθώς δεν υπάρχει βιώσιμη εναλλακτική από άποψη τεχνολογίας για την εκτέλεση των ίδιων λειτουργιών με συγκριτικά αποτελεσματικό τρόπο.

Κεφάλαιο 1

Αλγόριθμοι

Προκειμένου να αναλύσουμε την ηθική των αλγορίθμων, είναι σημαντικό να κατανοήσουμε τι εννοούμε με την έννοια αλγόριθμος. Η ηθική πλευρά των αλγορίθμων και η μελέτη της συνδέεται άρρηκτα με την ορθή κατανόηση των σχετικών εννοιών και όρων.

Είναι κρίσιμο να κατανοήσουμε τι μπορούν και τι δεν μπορούν να υλοποιήσουν οι αλγόριθμοι τόσο από την υπολογιστική πλευρά, όσο και από τις εφαρμογές τους. Επομένως, πρέπει να δώσουμε έναν ορισμό συγκεκριμένους αλλά ταυτόχρονα και εξειδικευμένο. Οι προσεγγίσεις μας ενδεικτικά θα είναι τριών κατηγοριών:

1. Ιστορική πλευρά: Η διερεύνηση της προέλευσης και τον τρόπο με τον οποίο αρχικά χρησιμοποιήθηκαν και την χρήση τους μέχρι τις ημέρες μας.
2. Περιγραφή μέσω της φυσικής γλώσσας: Πώς μπορούμε να εξηγήσουμε τη λειτουργία των αλγορίθμων.
3. Μελέτη φιλοσοφικών πλευρών: Από τεχνολογικής αλλά κι θεωρητικής άποψης τις αρχές που σχετίζονται με την λειτουργία τους.

Το κεφάλαιο αυτό βασίζεται στο άρθρο *ethics of algorithms*.

Σύμφωνα με το [4]: “Αντί να αμφισβητούμε την ασάφεια γύρω από τους αλγορίθμους ως όρο, ίσως πρέπει να το εναγκαλιάσουμε”. Σε αυτό το κεφάλαιο θα επικεντρωθούμε κυρίως στους ίδιους τους αλγόριθμους, και στα επόμενα κεφάλαια θα δούμε τη δυναμική λειτουργία τους και τη διπλή χρήση τους που αφορά τους ανθρώπους. Έτσι από τους δύο άξονες (τρόπους) στους οποίους βασίζεται αυτή η διπλωματική εργασία, δηλαδή *δυνατότητα* (*possibility*) και *επιθυμία* (*desirability*), αυτό το κεφάλαιο θα αναφερθεί στη δυνατότητα.

1.1 Σύντομη ιστορική επισκόπηση

Στα αρχαία ελληνικά μαθηματικά μια από τις πιο χαρακτηριστικές μεθόδους (αλγορίθμους) ήταν ο αλγόριθμος του Ευκλείδη. Ο **αλγόριθμος του Ευκλείδη** ή **Ευκλείδειος αλγόριθμος** είναι μια αποτελεσματική μέθοδος για τον υπολογισμό του μέγιστου κοινού διαιρέτη (ΜΚΔ) δύο ακεραίων αριθμών.

Η λέξη ‘αλγόριθμος’ συνδέεται ετυμολογικά με τον Abu Jafar Muhammad Ibn Musa al-Khwarizmi, μαθηματικό, ο οποίος είναι γνωστός για το παλαιότερο γνωστό βιβλίο άλγεβρας τον 9ο αιώνα μ.Χ. Το τελευταίο μέρος από το όνομα του al-Khwarizmi — το οποίο

μεταφράζεται από το Khardwarezm, μια περιοχή της Κασπίας Θάλασσας στην Κεντρική Ασία, έχει υποστεί “αλλοιωμένη μεταγραφή” με αποτέλεσμα να μετασχηματίσει την τρέχουσα λέξη ‘αλγόριθμος’ από λατινικά γράμματα. Άλλες παραλλαγές της λέξης εμφανίζονται στην αγγλική γλώσσα από τον 18ο αιώνα, και μόνο τον 20ο αιώνα ο αλγόριθμος έγινε βασική έννοια που υποδηλώνει ένα σύνολο υπολογιστικών διαδικασιών. Στην ουσία η έννοια περιγράφει διαδικασίες που επαναλαμβάνονται με μηχανικό τρόπο. Το έργο του al-Khwarizmi μεταφράστηκε τον 12ο αιώνα σε ευρωπαϊκές γλώσσες και εισήγαγε νέες μεθόδους προσέγγισης προβλημάτων αριθμητικής, αντικαθιστώντας σταδιακά μεθόδους που βασίζονται σε έναν πίνακα μέτρησης ή άβακα. Ο τύπος αλγορίθμου που περιγράφει ο al-Khwarizmi είναι για απλούς υπολογισμούς, για παράδειγμα για την εύρεση της τιμής του x σε εξισώσεις όπως $ax^2 + bx = \gamma$. Πρόκειται για ένα σχεδιασμό συγκεκριμένων βημάτων που με συγκεκριμένη είσοδο a, b, γ μας δίνει συγκεκριμένη έξοδο x . Οι επαναλαμβανόμενες διαδικασίες αποτελούν χαρακτηριστικό στοιχείο των αλγοριθμικών εφαρμογών διότι οι μαθηματικές παραστάσεις που περιγράφουν τις διαδικασίες σε αυτή τη μορφή είναι ένα τυπικό μαθηματικό σύστημα.

Ο καθηγητής Donald Knuth [\[62\]](#), σημείωσε ότι “μέχρι το 1950, ο αλγόριθμος ως λέξη σχετίζεται περισσότερο με τον αλγόριθμο του Ευκλείδη” που αναφέραμε στην αρχή. Ωστόσο, η χρήση της έννοιας επεκτάθηκε μετά τον ραγδαίο αριθμό θεμελιωδών ανακαλύψεων στον τομέα της θεωρίας του υπολογισμού και την έλευση των ηλεκτρονικών υπολογιστών. Η εφαρμογή τους οδήγησε στην ανάπτυξη ενός μεγάλου αριθμού διαφορετικών λειτουργιών. Οι αλγόριθμοι ταυτίζονται από πολλούς σήμερα μια μορφή τεχνολογίας ή μια συγκεκριμένη μορφή λήψης αποφάσεων ή μια επιστημολογία.

Τρεις κομβικές προσωπικότητες των μαθηματικών έχουν παίξει σημαντικό ρόλο στη θεμελίωση των αλγορίθμων. Ο Alonzo Church, ο Alan Turing και ο Andrey Markov, Jr.

Ο Alonzo Church (14 Ιουνίου 1903 – 11 Αυγούστου 1995) ήταν διαπρεπής Αμερικανός μαθηματικός και λογικός που συνέβαλε σημαντικά στη μαθηματική λογική και στα θεμέλια της θεωρητικής επιστήμης των υπολογιστών. Είναι πιο γνωστός για το λ-λογισμό, τη Church-Turing thesis, την αποδειξη της μη επιλυσιμότητας του προβλήματος Entscheidungsproblem, την οντολογία Frege – Church, και το γνωστό θεώρημα Church-Rosser. Ασχολήθηκε επίσης στη φιλοσοφία της γλώσσας και ήταν ο επιβλέπων καθηγητής του διδακτορικού του Alan Turing.

Ο Alan Turing (23 Ιουνίου 1912 – 7 Ιουνίου 1954) ήταν Άγγλος μαθηματικός, καθηγητής της λογικής, κρυπτογράφος και θεωρητικός βιολόγος θεωρείται ο «πατέρας της επιστήμης υπολογιστών» αλλά και της τεχνητής νοημοσύνης. Το λεγόμενο Turing test, προτάθηκε ως ένας τρόπος για να διαπιστωθεί εμπειρικά εάν μία μηχανή έχει επαρκείς γνωστικές δυνατότητες και “μπορεί να σκεφτεί”. Το έργο του από τη δεκαετία του '30 προσέδωσε στην ως τότε άτυπη έννοια του αλγορίθμου μία επίσημη, αυστηρή μαθηματική διατύπωση μέσω

της λεγόμενης Μηχανής Turing. Ακόμα, ο Turing διατύπωσε από κοινού με τον Alonzo Church την περίφημη εικασία του, ευρέως αποδεκτή, σύμφωνα με την οποία οποιοδήποτε μαθηματικό μοντέλο υπολογισμού είναι είτε ισοδύναμο είτε υποδεέστερο της Καθολικής Μηχανής Turing, επομένως αυτή περιγράφει τον ευρύτερο δυνατό υπολογιστή γενικού σκοπού: είναι θεωρητικά ικανή να υπολογίσει ό,τι είναι δυνατό να υπολογιστεί αλγοριθμικά

Ο σοβιετικός μαθηματικός Andrey Markov, Jr (22 Σεπτεμβρίου 1903 – 11 Οκτωβρίου 1979) ορίσε αυτό που σήμερα αποκαλείται αλγόριθμος Markov στην επιστήμη των υπολογιστών, δηλαδή ένα σύστημα επανεγγραφής συμβολοσειρών που χρησιμοποιεί κανόνες που μοιάζουν με γραμματική για να λειτουργεί σε συμβολοσειρές συμβόλων. Ο ορισμός κάθε (κανονικού) αλγορίθμου αποτελείται από δύο μέρη: τον ορισμό του αλφαβήτου του αλγορίθμου (ο αλγόριθμος θα εφαρμοστεί στις συμβολοσειρές αυτών των συμβόλων αλφαβήτου) και τον ορισμό του σχήματος του. Το σχήμα ενός κανονικού αλγορίθμου είναι ένα πεπερασμένο σύνολο των λεγόμενων τύπων αντικατάστασης, καθένας από τους οποίους μπορεί να είναι απλός ή τελικός. Οι απλοί τύποι αντικατάστασης αντιπροσωπεύονται από συμβολοσειρές της μορφής $L \rightarrow D$ όπου L και D είναι δύο αυθαίρετες συμβολοσειρές στο αλφάβητο του αλγορίθμου. Ομοίως, οι τελικοί τύποι αντικατάστασης αντιπροσωπεύονται από συμβολοσειρές της μορφής $L \rightarrow \cdot D$ όπου L και D είναι δύο αυθαίρετες συμβολοσειρές στο αλφάβητο του αλγορίθμου. Οι βοηθητικοί χαρακτήρες \rightarrow και $\rightarrow \cdot$ δεν ανήκουν στο αλφάβητο του αλγορίθμου. Ένα παράδειγμα ενός (κανονικού) σχήματος αλγορίθμου στο αλφάβητο πέντε γραμμάτων $| * a b c$:

$$\left\{ \begin{array}{l} |b \rightarrow ba| \\ ab \rightarrow ba \\ b \rightarrow \\ *| \rightarrow b* \\ * \rightarrow c \\ |c \rightarrow c \\ ac \rightarrow c| \\ c \rightarrow \cdot \end{array} \right.$$

Πολλοί αντιλαμβάνονται τους αλγόριθμους ως μια συγκεκριμένη μορφή ορθολογισμού, ως μέρος ενός γενικού μοντέλου κοινωνικής οργάνωσης, δηλαδή ως κοινωνικο-τεχνική διαδικασία. Σε αυτή την εργασία, θα χρησιμοποιήσουμε στοιχεία κάθε ερμηνείας για να ενισχύσουμε την ανάλυσή μας, καθώς κάθε ερμηνεία προσφέρει ενδιαφέρουσες πτυχές των αλγορίθμων. Η ανάπτυξη και η ενσωμάτωση των αλγορίθμων, έχει αναμφισβήτητα προχωρήσει πολύ από τη δεκαετία του 1950, και ξεπερνά όλες τις ηλεκτρονικές τεχνολογίες, καθώς και έχει φέρει επανάσταση σε τρόπους σκέψης και δομής του κόσμου.

1.2 Πως μπορούμε να εξηγήσουμε τη θεωρία αλγορίθμων

Υπάρχει μια ευρεία γκάμα εφαρμογών των αλγορίθμων. Είναι φυσιολογικό να οδηγηθούμε στο ερώτημα τι είναι αυτό που κάνει τους αλγόριθμους να έχουν τέτοια επιτυχία και τέτοιο εύρος εφαρμογής.

1.2.1 Τα χαρακτηριστικά των αλγορίθμων

Σύμφωνα με τον Knuth: “οι αλγόριθμοι είναι ένα σύνολο πεπερασμένων κανόνων που δίνει μια ακολουθία λειτουργιών για την επίλυση ενός συγκεκριμένου τύπου προβλήματος”. Ο Knuth διαχώρισε τους αλγόριθμους από άλλες έννοιες όπως “συνταγή”, “διαδικασία”, “μέθοδος”, “τεχνική”, “διαδικασία”, “ρουτίνα” [\[62\]](#) μέσω πέντε βασικών χαρακτηριστικών:

1. Περαιτότητα (finiteness) — Ένας αλγόριθμος πρέπει να τερματίσει, δηλαδή να σταματήσει να λειτουργεί, μετά από έναν πεπερασμένο αριθμό βημάτων, ανεξάρτητα από την (αποδεκτή) είσοδο.
2. Καθοριστικότητα (definiteness) — Κάθε βήμα της λειτουργίας του αλγορίθμου πρέπει να είναι “καλά καθορισμένο”, δηλαδή να μην περιέχει πιθανή ασάφεια. Δηλαδή, οι αλγόριθμοι δεν μπορούν να υπάρχουν χωρίς μια επίσημη γραμματική, χωρίς περιβάλλον που ορίζει και περιορίζει όλες τις σχέσεις μεταξύ των πιθανών συνόλων συμβολοσειρών.
3. Είσοδος (input) — Ένας αλγόριθμος μπορεί να διαθέτει μηδενικές ή περισσότερες εισόδους. Δηλαδή, ο αλγόριθμος μπορεί να εκκινήσει με βάση μια ποσότητα δεδομένων πριν από το πρώτο βήμα.
4. Έξοδος (output) — ο αλγόριθμος παράγει μία ή περισσότερες εξόδους. Οι έξοδοι προκύπτουν από την είσοδο ή τις εισόδους, μέσω πεπερασμένων και καλά καθορισμένων βημάτων.
5. Αποτελεσματικότητα (effectiveness) — Αφορά την πολυπλοκότητα που περιλαμβάνεται σε κάθε βήμα. Τα μη λογικά ή λογικά ασυνεπή βήματα δεν αποτελούν μέρος ενός κατάλληλου αλγορίθμου.

Παράδειγμα - Ο αλγόριθμος quicksort:

Αυτός ο αλγόριθμος χρησιμοποιείται για την εσωτερική ταξινόμηση πεπερασμένων συνόλων πινάκων, όπως φαίνεται παρακάτω, όπου n το μήκος του πίνακα [\[105\]](#):

1. Επιλέγεται ένα στοιχείο *pivot* από τον πίνακα. Μπορούμε να επιλέξουμε οποιοδήποτε στοιχείο από τον πίνακα ως *pivot*. Στο παράδειγμα μας θα πάρουμε το δεξιότερο (δηλαδή το τελευταίο στοιχείο) του πίνακα ως *pivot*.

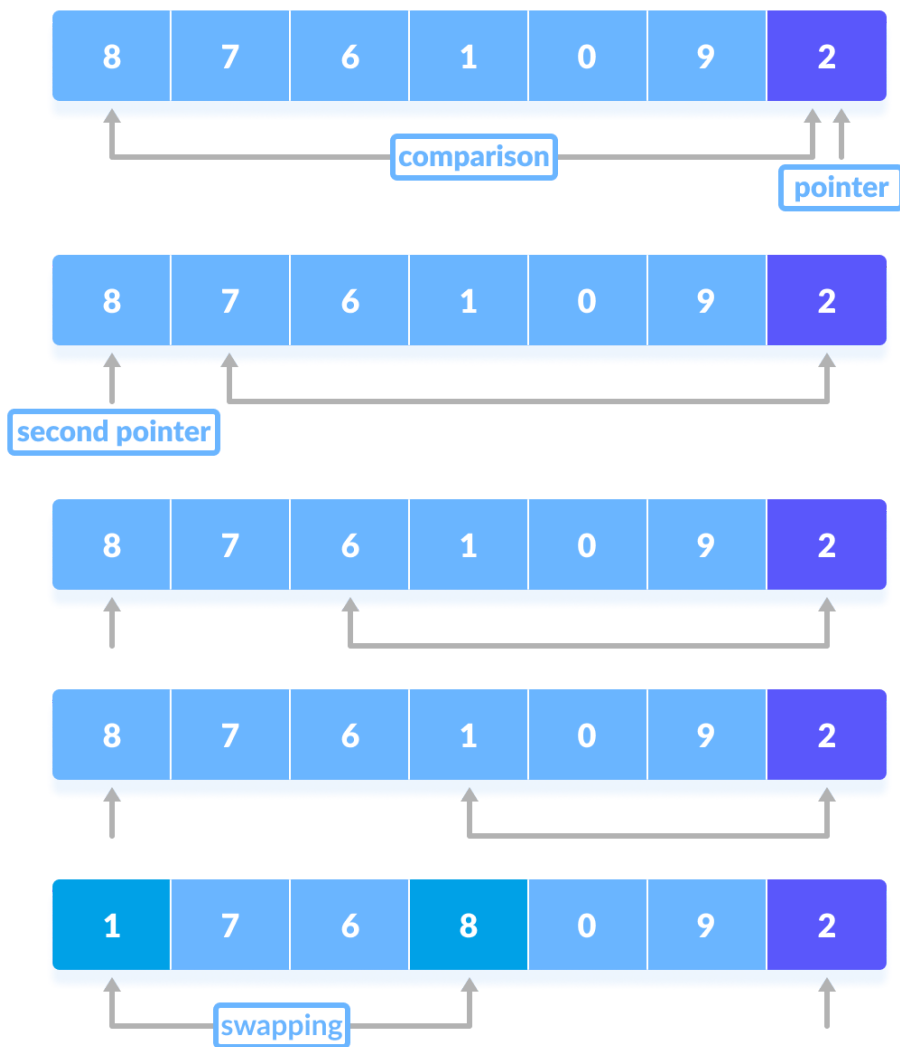


2. Τα στοιχεία μικρότερα από το ρινότ τοποθετούνται στα αριστερά και τα στοιχεία μεγαλύτερα από το ρινότ τοποθετούνται στα δεξιά.



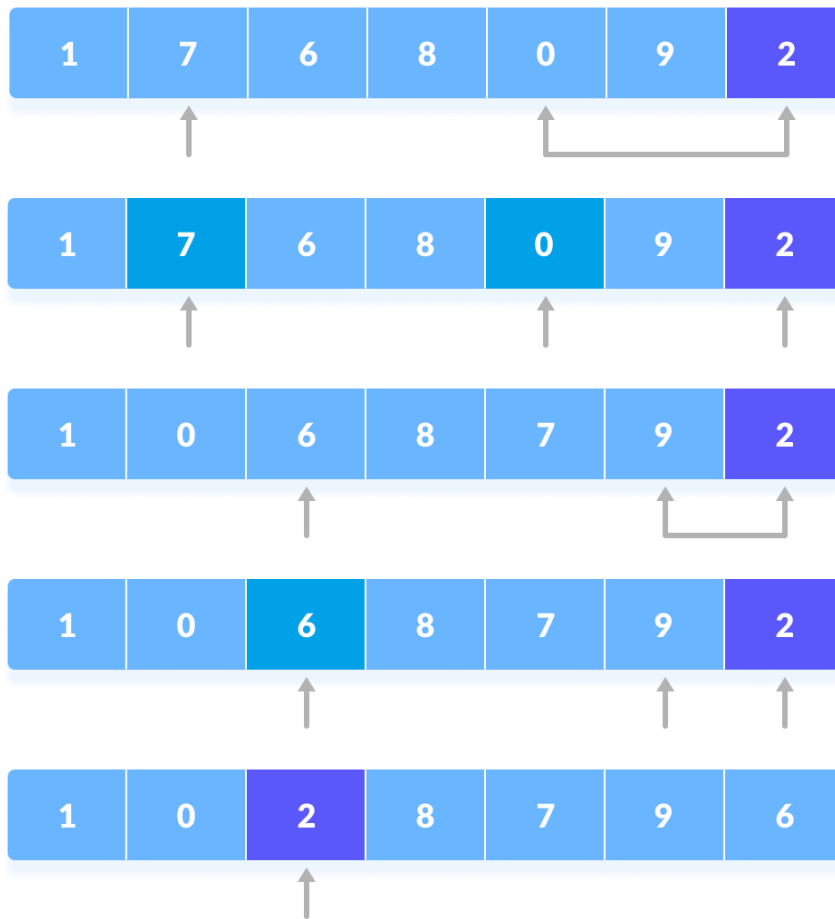
Η παραπάνω ρύθμιση επιτυγχάνεται με τα ακόλουθα βήματα.

- Ένας δείκτης είναι σταθερός στο ρινότ. Το ρινότ συγκρίνεται με τα στοιχεία που ξεκινούν από τον πρώτο δείκτη. Εάν το στοιχείο που δείχνει ο αριστερός δείκτης είναι μεγαλύτερο από το ρινότ, ορίζεται ένας δεύτερος δείκτης για αυτό το στοιχείο.
- Τώρα, το ρινότ συγκρίνεται με τα άλλα στοιχεία (τρίτος δείκτης). Εάν επιτευχθεί ένα στοιχείο μικρότερο από το ρινότ, το μικρότερο στοιχείο ανταλλάσσεται με το μεγαλύτερο στοιχείο που βρέθηκε νωρίτερα.



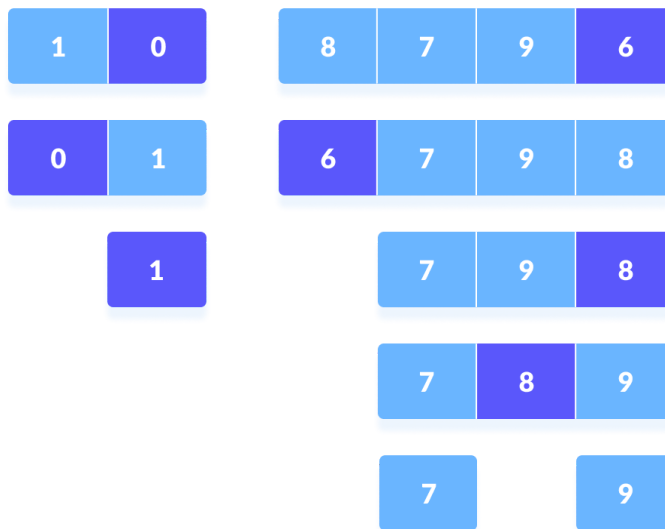
- Η διαδικασία συνεχίζεται έως ότου επιτευχθεί το δεύτερο τελευταίο στοιχείο.

Τέλος, το ρινοί εναλλάσσεται με το δεύτερο δείκτη.



- Τώρα τα αριστερά και τα δεξιά τμήματα αυτού του ρινοῦ λαμβάνονται για περαιτέρω επεξεργασία στα παρακάτω βήματα.

3. Τα ρινοῦ επιλέγονται ξανά για τα αριστερά και τα δεξιά υπο-τμήματα ξεχωριστά. Μέσα σε αυτά τα υπο-μέρη, τα περιστρεφόμενα στοιχεία τοποθετούνται στη σωστή τους θέση. Στη συνέχεια, το βήμα 2 επαναλαμβάνεται.



4. Τα υπο-τμήματα χωρίζονται και πάλι σε μικρότερα υπο-τμήματα έως ότου κάθε τμήμα αποτελείται από ένα μόνο στοιχείο.
5. Σε αυτό το σημείο, ο πίνακας έχει ήδη ταξινομηθεί.

Από αυτό το σύντομο παράδειγμα του αλγορίθμου Quicksort προκύπτει ότι έχει τις παρακάτω ιδιότητες:

- Περαιτότητα: ο αριθμός των βημάτων εξαρτάται σε μεγάλο βαθμό από την είσοδο του μήκους n του πίνακα
- Καθοριστικότητα : κάθε βήμα είναι καλά καθορισμένο και έγκυρο για κάθε επιτρεπόμενη είσοδο και παράγει έξοδο ίσου μήκους n .
- Αποτελεσματικότητα : αυτός ο αλγόριθμος είναι καλά καθορισμένος ώστε να ανήκει στην κατηγορία $O(n \log n)$ πολυπλοκότητας.

Quicksort εμφανίστηκε και εφαρμόστηκε στο Unix ως η προεπιλεγμένη συνάρτηση ταξινόμησης βιβλιοθηκών, από όπου έδωσε το όνομά της συνάρτηση της γλώσσας προγραμματισμού C (στην standard library) qsort και αντίστοιχα στη Java.

Μια υλοποίηση της quicksort σε Python, όπου επιλέγουμε ως ρινοτ το τελευταίο στοιχείο του πίνακα: [\[104\]](#)

```
# Python program for implementation of Quicksort Sort
# This function takes last element as pivot, places
```

```

# the pivot element at its correct position in sorted
# array, and places all smaller (smaller than pivot)
# to left of pivot and all greater elements to right
# of pivot
def partition(arr,low,high):
    i = ( low-1 )          # index of smaller element
    pivot = arr[high]     # pivot

    for j in range(low , high):

        # If current element is smaller than the pivot
        if arr[j] < pivot:

            # increment index of smaller element
            i = i+1
            arr[i],arr[j] = arr[j],arr[i]

    arr[i+1],arr[high] = arr[high],arr[i+1]
    return ( i+1 )

# The main function that implements QuickSort
# arr[] --> Array to be sorted,
# low --> Starting index,
# high --> Ending index

# Function to do Quick sort
def quickSort(arr,low,high):
    if low < high:

        # pi is partitioning index, arr[p] is now
        # at right place
        pi = partition(arr,low,high)

        # Separately sort elements before
        # partition and after partition
        quickSort(arr, low, pi-1)

```

```

        quickSort(arr, pi+1, high)

# Driver code to test above
arr = [10, 7, 8, 9, 1, 5]
n = len(arr)
quickSort(arr, 0, n-1)
print ("Sorted array is:")
for i in range(n):
    print ("%d" %arr[i]),

```

1.2.2 Πιο σύγχρονα ερωτήματα

Υπάρχουν πολλοί που πιστεύουν ότι τα πέντε χαρακτηριστικά του Knuth θεωρούνται ξεπερασμένα [\[23\]](#) [\[121\]](#). Υποστηρίζουν ότι δεν δίνουν έναν γενικό ορισμό της “διαδικασίας”, καθιστώντας το ανοιχτό σε ερμηνεία. Επίσης ισχυρίστηκαν ότι ο Knuth προβάλλει συγκεκριμένους ισχυρισμούς σχετικά με αλγορίθμους που δεν ισχύουν πλέον, όπως ότι ένας αλγόριθμος πρέπει να σταματάει για κάθε είσοδο. Ο Dean προσπάθησε να δώσει μια εναλλακτική λίστα “κοινών παρατηρήσεων” που αφορούν τους αλγορίθμους [\[23\]](#):

1. Μαθηματικές διαδικασίες — Οι αλγορίθμοι μπορούν να θεωρηθούν ως διαδικασίες που δρουν σε μαθηματικά αντικείμενα, δεδομένου ότι παίρνει είσοδο και επιστρέφει μια έξοδο.
2. Μαθηματικά προβλήματα — Οι αλγορίθμοι ουσιαστικά εισάγονται για να εξυπηρετήσουν ένα συγκεκριμένο μαθηματικό πρόβλημα.
3. Περατότητα — Οι αλγόριθμοι διακρίνονται από τις μαθηματικές διαδικασίες με τη γενική έννοια, πιο συγκεκριμένα:
 - Οι αλγόριθμοι δεσμεύονται από μια πεπερασμένη δυνατότητα προσδιορισμού ή από έναν πεπερασμένο αριθμό πρωταρχικών εκφράσεων σε γλώσσες προγραμματισμού.
 - Η φύση των ίδιων των πρωταρχικών εκφράσεων πρέπει να οδηγήσει σε μείωση των πόρων που απαιτούνται για τον τερματισμό το αλγορίθμου.
 - Είναι σχεδόν πάντα απαραίτητο ο αλγόριθμος να εκτελεί έναν πεπερασμένο αριθμό βημάτων πριν από την ολοκλήρωση.
4. Επαναληψιμότητα (repetition) - οι περισσότεροι γνωστοί αλγόριθμοι σχετίζονται με μεθόδους μαθηματικών παραστάσεων.
5. Αφαιρετικότητα (abstraction) — Οι αλγόριθμοι καθορίζονται συμβατικά ως “υποχρεωτικές δηλώσεις πάνω σε μια φυσική ή τυπική γλώσσα”, αυτό δεν πρέπει να συγχέεται με τον πραγματικό αλγόριθμο.

1.2.3 Αλγόριθμοι & Προγράμματα

Ένα πρόγραμμα παρουσιάζει τα πέντε χαρακτηριστικά του Knuth. Η πιο συνηθισμένη τοποθέτηση είναι να ισχυριστούμε ότι οι αλγόριθμοι και τα προγράμματα ταυτίζονται, δεδομένου ότι και τα δύο αποτελούνται από βήματα - εντολές. Ένα στοιχείο διαφοράς μεταξύ αλγορίθμων και προγραμμάτων είναι ότι οι αλγόριθμοι μπορούν να εκφραστούν σε φυσική γλώσσα, ενώ τα προγράμματα διατυπώνονται σε γλώσσες προγραμματισμού και είναι εκτελέσιμα από ηλεκτρονικούς υπολογιστές. Ισχυριζόμαστε ότι “ένας αλγόριθμος είναι μια γενική τεχνική για την επίλυση ενός προβλήματος (δηλαδή, προσανατολισμένη στο πρόβλημα), ενώ ένα πρόγραμμα είναι το συγκεκριμένη διατύπωση ενός αλγορίθμου, όπως απαιτείται για την εκτέλεση από έναν υπολογιστή (και επομένως είναι προσανατολισμένος στη μηχανή)”. Επιπλέον, μπορούμε να συνοψίσουμε ότι “ο αλγόριθμος μπορεί να θεωρηθεί ως η καρδιά του προγράμματος”. Επιπλέον, οι αλγόριθμοι και τα προγράμματα διαφέρουν από το εύρος του λεξιλογίου, καθώς οι γλώσσες προγραμματισμού έχουν περιορισμένο λεξιλόγιο. Ο ίδιος αλγόριθμος μπορεί να υλοποιηθεί με διαφορετικά προγράμματα σε διαφορετικές γλώσσες.

1.2.4 Αποδοτικότητα

Ως αποδοτικότητα ορίζεται ο αριθμός των πόρων που απαιτεί ο αλγόριθμος για να τερματίσει και να παράγει έξοδο. Οι “πόροι” αφορούν τόσο ζητήματα χρόνου όσο και χώρου.

- Ο χρόνος μετριέται από την αρχή ενός αλγορίθμου έως τον τερματισμό του, δηλαδή ο χρόνος που απαιτείται για τη μετάβαση από την είσοδο στην έξοδο.
- Ο χώρος αναφέρεται στους υπολογιστικούς πόρους που απαιτούνται για τον τερματισμό και ορίζεται ως επί το πλείστον σε σχέση με τη μνήμη RAM. Τέσσερα διαφορετικά στοιχεία που εμπλέκονται στον τρόπο με τον οποίο ένας αλγόριθμος “χρησιμοποιεί” χώρο RAM:
 1. Το μέγεθος του χώρου μνήμης που καταλαμβάνει ο ίδιος ο κωδικός του αλγορίθμου
 2. Η συνολική ποσότητα χώρου μνήμης που καταλαμβάνει η είσοδος του αλγορίθμου
 3. Η συνολική ποσότητα χώρου μνήμης που καταλαμβάνει η έξοδος του αλγορίθμου
 4. Το μέγεθος του χώρου μνήμης που απαιτεί η λειτουργία κατά την εκτέλεση του αλγορίθμου.

Σε ορισμένες περιπτώσεις, η ταχύτητα (ελκυστική στον άξονα του χρόνου) είναι απαραίτητη, όπως για τον σκοπό της γρήγορης (ιατρικής) διάγνωσης. Αυτή η ταχύτητα έρχεται σε βάρος της υπολογιστικής ισχύος, ανεξάρτητα από τις ιδιαιτερότητες του ίδιου του αλγορίθμου. Η

ισορροπία στη δημιουργία μετεωρολογικών μοντέλων έχει συχνά την τάση να δίνει προτεραιότητα προς τη βελτιστοποίηση της υπολογιστικής ισχύος, καθώς ο χρόνος εκτέλεσης του αλγορίθμου δεν είναι περιοριστικός. Οι αλγόριθμοι μπορούν να τροποποιηθούν για να δοθεί προτίμηση σε κάθε δεδομένο τύπο βελτιστοποίησης πόρων.

1.3 Ζητήματα Φιλοσοφίας του Υπολογισμού

Στην ενότητα αυτή θα ασχοληθούμε με την Υπολογισιμότητα, τη Church-Turing Thesis και τις Μηχανές Turing.

1.3.1 Η θέση Church - Turing (thesis)

Η θέση Church - Turing αφορά ότι κάθε αποτελεσματικός υπολογισμός μπορεί να πραγματοποιηθεί από μια μηχανή Turing. Ωστόσο υπάρχουν διάφορες ισοδύναμες διατυπώσεις της θέσης Church-Turing.

Υπολογισιμότητα

Η μελέτη της υπολογισιμότητας αφορά το τι είναι θεωρητικά υπολογίσιμο ή όχι. Το έτος 1936, όταν οι Church Alonzo και ο Alan Turing δημοσίευσαν, ανεξάρτητες εργασίες σχετικά με την υπολογισιμότητα σε συνέχεια του σχετικού προβλήματος (Entscheidungsproblem) που είχε τεθεί από τους David Hilbert και Wilhelm Ackermann το 1928. Το πρόβλημα αφορά την δυνατότητα να υπάρχει ένας αλγόριθμος που να μας απαντά με "Ναι" ή "Όχι" (ανάλογα με το εάν ένα πρόβλημα είναι υπολογίσιμο ή όχι).

Σύμφωνα με τον Church: [\[18\]](#)

“Από το Entscheidungsproblem σε ένα σύστημα συμβολικής λογικής είναι κατανοητό ότι είναι αναγκαίο να βρούμε μια αποδοτική αποτελεσματική μέθοδο με την οποία, δεδομένης οποιασδήποτε έκφρασης Q στη σημειογραφία του συστήματος, να μπορεί να προσδιοριστεί αν το Q είναι αποδεδειγμένο ή όχι στο σύστημα”¹.

Αυτό μας οδηγεί στο ερώτημα εάν υπάρχει μια αποτελεσματική μέθοδος για να διαπιστωθεί εάν μια λογική έκφραση είναι καθολικά έγκυρη. Οι λέξεις “αποτελεσματική μέθοδος” σε αυτό το απόσπασμα είναι ένας άτυπος τρόπος έκφρασης, όπου πρέπει να υπάρχει μια διαδικασία ή αλγόριθμος για να γίνει αυτό σε έναν πεπερασμένο αριθμό βημάτων. Οι Church και Turing προσέγγισαν το Entscheidungsproblem από διαφορετικές οπτικές γωνίες [\[109\]](#). Δεδομένου ότι οι απαντήσεις των Church και Turing είναι ισοδύναμες, θα εξετάσουμε

¹ “By the Entscheidungsproblem of a system of symbolic logic here is understood the problem to find an effective method by which, given any expression Q in the notation of the system, it can be determined whether or not Q is provable in the system.”

μόνο μία από τις δύο απαντήσεις. Επιλέγουμε για λόγους απλότητας την απάντηση του Turing, καθώς η απάντηση του Church περιλαμβάνει λ-λογισμό και είναι πιο σύνθετη. Πριν την απάντηση του Turing στο Entscheidungsproblem, απαιτούνται δύο βήματα:

1. Η “αποτελεσματική μέθοδος” του Turing
2. Η μελέτη των μηχανών Turing που τυποποιούν τη μέθοδο του.

Σύμφωνα με το [\[20\]](#), μια αποτελεσματική μέθοδος M περιέχει τα ακόλουθα στοιχεία:

1. Το M ορίζεται σε όρους πεπερασμένου αριθμού ακριβών οδηγιών και κάθε οδηγία εκφράζεται μέσω ενός πεπερασμένου αριθμού συμβόλων
2. Εάν το M πραγματοποιηθεί χωρίς σφάλμα, θα παράγει το επιθυμητό αποτέλεσμα σε έναν πεπερασμένο αριθμό βημάτων
3. Το M μπορεί στην πράξη να εκτελείται από έναν άνθρωπο χωρίς βοήθεια από οποιαδήποτε μηχανήματα εκτός από χαρτί και μολύβι
4. Ο M δεν απαιτεί καμία διορατικότητα ή εφευρετικότητα από την πλευρά του ανθρώπου που το εκτελεί.

Σε συνέχεια των παραπάνω ο Turing πρότεινε μια μαθηματική δομή που ονομάζεται υπολογιστική μηχανή και στη συνέχεια πήρε το όνομα του (Turing Machines (TMs)). Στην ουσία πρόκειται για μια μαθηματική τυποποίηση με τη βοήθεια της λογικής που αφορά τη φυσική περιγραφή ενός αλγορίθμου. Αυτή η έννοια είναι λογικά παρόμοια με τη φυσική περιγραφή ενός αλγορίθμου στην προηγούμενη ενότητα. Η προσέγγιση του Turing ήταν μια προσπάθεια να ορίσουμε τη διαδικασία υπολογισμού από τα πιο βασικά στοιχεία και να φτάσουμε στην καρδιά του υπολογισμού. Οι μηχανές Turing αποτελούνται από έναν σαρωτή και μια απεριόριστη ταινία μνήμης που κινείται πάνω στο σαρωτή. Η ταινία χωρίζεται σε τμήματα, δηλαδή ξεχωριστά στοιχεία, καθένα από τα οποία είναι είτε είναι το κενό ή περιέχει ένα μόνο σύμβολο, συνήθως είτε 0 είτε 1. Μπορεί να χρησιμοποιηθεί σε πολλά αλφάβητα δηλαδή σε οποιοδήποτε πεπερασμένο σύνολο στοιχείων μιας γλώσσας. Ο σαρωτής είναι σε θέση να εξετάζει ένα στοιχείο κάθε φορά και να λειτουργεί ανάλογα αλλάζοντας την κατεύθυνση της ταινίας και το περιεχόμενο κάθε διακριτού χώρου και προς τις δύο κατευθύνσεις.

Επιπλέον, ο σαρωτής έχει τη δυνατότητα διαγραφής και εκτύπωσης συμβόλων κατά τη σάρωση της ταινίας. Η Μηχανή Turing μπορεί, θεωρητικά, να εκτελέσει κάθε υπολογισμό που θα μπορούσε να υλοποιηθεί μηχανικά δηλαδή αλγοριθμικά. Υπάρχουν πολλές παραλλαγές με διάφορους βαθμούς αυστηρότητας της μηχανής Turing.

Εκδοχές της Church - Turing Thesis

Οι Λογικές Υπολογιστικές Μηχανές - Logical Computing Machines LCMs² μπορούν να κάνουν οτιδήποτε θα μπορούσε να περιγραφεί ως “κανόνας” ή “καθαρά μηχανική διαδικασία”.

Στη θεωρία υπολογισμού, το πρόβλημα τερματισμού - halting problem μπορεί να οριστεί ως εξής: "Δοθείσης μιας περιγραφής ενός αυθαίρετου υπολογιστικού προγράμματος αποφάσισε αν το πρόγραμμα θα σταματήσει να τρέχει ή αν θα συνεχίσει να τρέχει για πάντα". Το παραπάνω πρόβλημα είναι ισοδύναμο με το πρόβλημα της απόφασης, δοθέντος ενός προγράμματος και μιας εισόδου, αν αυτό τελικά θα σταματήσει εφόσον έχει τρέξει πεπερασμένες φορές ή αν θα συνεχίσει επ' άπειρον. Το halting problem δηλαδή είναι ένα πρόβλημα που αποφασίζει εκ των προτέρων εάν ένα πρόβλημα είναι δυνατό να τερματίσει. Ένα παράδειγμα από ένα non-halting TM θα υπολόγιζε τον αριθμού π με ολοένα αυξανόμενη ακρίβεια. Δεδομένου ότι ο αριθμός π είναι ένας άρρητος αριθμός, τα ψηφία που πρέπει να υπολογιστούν είναι ατελείωτα χωρίς να προκύψει κάποιο επαναλαμβανόμενο μοτίβο. Ένα TM που υπολογίζει αυτά τα ψηφία θα τρέχει για πάντα χωρίς τερματισμό. Το halting problem συνίσταται στο γεγονός ότι είναι αδύνατο σε ορισμένες περιπτώσεις να αποφασιστεί εκ των προτέρων εάν το TM θα τερματίσει ή όχι. Η λειτουργία διακοπής δεν μπορεί να υπολογιστεί από μια μηχανή Turing" [19].

Ισχυρή εκδοχή της θέσης Church-Turing:

Στην ισχυρή εκδοχή της θέσης Church Turing ισχυρίζεται ότι η κατηγορία των καλά καθορισμένων προβλημάτων μπορεί να υλοποιηθεί από οποιοδήποτε Turing Machine. [103]

1.3.2 Πέρα από τις Μηχανές Turing

Το μοντέλο υπολογισμού των μηχανών Turing είχε εξαιρετική επιρροή, και δικαίως θεωρείται ως θεμελιώδες για τους ηλεκτρονικούς υπολογιστές. Ωστόσο, όπως οι Goldin και Wegner [45] αμφισβήτησαν, θα ήταν λάθος όλα όσα εκτελούνται από τους σημερινούς υπολογιστές να τα εξισώσουμε με τις μηχανές Turing. Αυτό δεν σημαίνει ότι οι αλγόριθμοι δεν αποτελούν σημαντικό παράγοντα με θεμελιώδη έννοια, καθώς αυτή η λογική υποστηρίζει σχεδόν όλες τις εφαρμογές, αλλά κάπως ότι η σημασία της ισοδυναμίας TM δεν πρέπει να υπερεκτιμάται.

1.3.3 Υπολογιστική Πολυπλοκότητα

² LCM: Turing's expression for Turing machines.

Η θεωρία πολυπλοκότητας είναι το μέρος εκείνο της θεωρίας υπολογισμού, το οποίο ασχολείται με την κοστολόγηση των πόρων που απαιτούνται για την αλγοριθμική επίλυση ενός προβλήματος. Επομένως η θεωρία πολυπλοκότητας αποτελεί βασικό δομικό λίθο της ανάλυσης αλγορίθμων και κεντρικό γνωστικό πεδίο της επιστήμης υπολογιστών.

Οι συνηθέστερες προσεγγίσεις όσο αφορά την πολυπλοκότητα είναι:

- η πολυπλοκότητα ως προς το χρόνο, οπότε μιλάμε για τη χρονική πολυπλοκότητα του αλγορίθμου, δηλαδή πόσα «βήματα» χρειάζεται να εκτελέσει ο αλγόριθμος συναρτήσει της εισόδου του.

Παράδειγμα η εύρεση του μέγιστο στοιχείου μιας λίστας που δεν είναι ταξινομημένη έχει χρονική πολυπλοκότητα $O(n)$ όπου n το μήκος της λίστας.

- πολυπλοκότητα χώρου, οπότε μιλάμε για τη χωρική πολυπλοκότητα, δηλαδή πόσο χώρο (μνήμη) χρειάζεται ο αλγόριθμος συναρτήσει της εισόδου του.

Στην περίπτωση του πολυπλοκότητας χώρου είναι σημαντικό να λάβουμε υπόψιν μας ότι παίζει σημαντικό ρόλο η γλώσσα προγραμματισμού, ο μεταγλωττιστής ή και το μηχάνημα που χρησιμοποιείται. Στην περίπτωση της QuickSort που αναφέραμε παραπάνω, ενώ η χρονική πολυπλοκότητα είναι $O(n \log n)$, η χωρική πολυπλοκότητα είναι $O(n)$ όσο δηλαδή για να δημιουργηθεί ο ταξινομημένος πίνακας.

Εκτός από αυτές τις δύο αυτές βασικές προσεγγίσεις μπορεί να μας ενδιαφέρει επίσης πόσοι επεξεργαστές χρειάζονται για να λυθεί ένα πρόβλημα. [\[124\]](#)

Η θεωρία υπολογισιμότητας εξετάζει αν μπορεί να λυθεί κάποιο πρόβλημα με έναν υπολογιστή ισοδύναμο με TM. Το πεδίο της υπολογιστικής θεωρίας πολυπλοκότητας μελετά το επίπεδο πολυπλοκότητας των προβλημάτων και τα ταξινομεί ως προς τον χρόνο που απαιτείται από έναν βέλτιστο αλγόριθμο για την επίλυσή του. Η πιο σχετική διάκριση στον τομέα αυτό είναι μεταξύ των δύο κατηγοριών που ονομάζονται P και NP.

P: περιλαμβάνει όλα τα υπολογιστικά προβλήματα που μπορούν να λυθούν σε πολυωνυμικό χρόνο.

NP: Όλα τα προβλήματα για οποία η λύση μπορεί να επαληθευτεί σε πολυωνυμικό χρόνο. Διαφορετική περιγραφή: Όλα τα προβλήματα τα οποία μπορούν να λυθούν σε πολυωνυμικό χρόνο από μια μη-ντετερμινιστική μηχανή Turing. Μη-ντετερμινιστική μηχανή Turing είναι μια μηχανή Turing που σε κάθε βήμα μπορεί και μαντεύει σωστά όταν πρέπει να επιλέξει μεταξύ περισσότερες από μια επιλογές. Μια μη-ντετερμινιστική μηχανή Turing είναι ένα (μη ρεαλιστικό) πολύ ισχυρό υπολογιστικό μοντέλο.

Τα προβλήματα που ανήκουν στο P μπορούν να υπολογιστούν απο ντετερμινιστικές μηχανές Turing σε πολυωνυμικό χρόνο, όπως για παράδειγμα:

$$n^2, \forall n \in \mathbb{N}$$

Δεδομένης μιας αρκετά μεγάλης τιμής για το n , αυτό σαφώς θα είναι προβληματικό σε πρακτικούς όρους για επίλυση, αλλά αυτή η πολυπλοκότητα εξακολουθεί να είναι πολυωνυμική για υπολογισμό και είναι διαφορετική διάταξη από την NP.

Αντίστοιχα, ένα παράδειγμα που ανήκει στα προβλήματα τάξης NP είναι:

$$2^n, \forall n \in \mathbb{N}$$

Η μη-ντετερμινιστική αποφασιστικότητα αυτής της τάξης προβλημάτων περιστρέφεται γύρω από την ιδέα ότι δεν είναι προσδιορίσιμο εάν θα επιτευχθεί λύση σε οποιοδήποτε χρονικό διάστημα, αλλά όταν έχει καθοριστεί, η επαλήθευση της λύσης είναι δυνατή σε πολυωνυμικό χρόνο.

Κεφάλαιο 2

Ηθική

Το κεφάλαιο αυτό αποτελεί μια σύντομη ιστορία της ηθικής. Η κύρια πηγή σε αυτό το κεφάλαιο είναι η Stanford Encyclopedia of Philosophy. Παραθέτουμε μια συνοπτική ιστορία της ηθικής από την αρχαία Ελλάδα, φτάνοντας μέχρι τις σύγχρονες τάσεις που επικρατούν. Θα καταλήξουμε σε σύγχρονα ερωτήματα στην Ηθική των Αλγορίθμων.

Ηθική είναι ο κλάδος της φιλοσοφίας που ασχολείται κυρίως με το ερώτημα: ποιες ανθρώπινες πράξεις είναι αποδεκτές και ποιές ορθές και ποιες ανάρμοστες και ποιές λανθασμένες. Ωστόσο, στα θέματα αυτά σπάνια υπάρχει συμφωνία θέσεων ακόμη και του ποιο είναι το θεμελιώδες ερώτημα αποτελεί αντικείμενο διχογνωμίας.

Για μεθοδολογικούς κυρίως λόγους η ηθική χωρίζεται συνήθως σε τρία πεδία ή κλάδους: σε κανονιστική, μεταηθική και περιγραφική.

Η κανονιστική ηθική διατυπώνει κανόνες του ηθικού πράττειν, η μεταηθική είναι μία κριτική δευτέρου επιπέδου, περί του πώς πρέπει να εννοούνται και να χρησιμοποιούνται οι ηθικές έννοιες και η περιγραφική ηθική περιγράφει τα ήθη και το τρόπο σκέψης μιας κοινωνίας.

Οι συνεπειοκρατικές και οι δεοντολογικές ηθικές θεωρίες θέτουν ως κεντρικό ζήτημα τα κριτήρια ορθότητα ή μη-ορθότητας των πράξεων, αν και διαφέρουν ως προς το ποια είναι αυτά τα κριτήρια.

Οι αρεταϊκές ηθικές θεωρούν ως κεντρικό ζήτημα τα στοιχεία που καθορίζουν το ήθος, τις ηθικές αρετές του χαρακτήρα ή τον τρόπο ζωής που θα πρέπει να ακολουθείται έτσι.

2.1 Σύντομη Ιστορία

2.1.1 Προέλευση της Ηθικής

Από που προέρχεται η Ηθική; Αν κάποιος έχει κατά νου την ηθική που θέλει να δώσει απάντηση στο ερώτημα τι είναι σωστό, δηλαδή, τη μελέτη του τι είναι ηθικά σωστό και λάθος, είναι σαφές ότι η ηθική θα μπορούσε να είχε υπάρξει μόνο όταν τα ανθρώπινα όντα άρχισαν να σκέφτονται για έναν καλύτερο τρόπο ζωής. Αυτό το στοχαστικό στάδιο εμφανίστηκε πολύ μετά την ανάπτυξη των ανθρώπινων κοινωνιών κάποιου είδους ηθικής, συνήθως με τη μορφή συνήθων προτύπων ορθής και λανθασμένης συμπεριφοράς.

Στο Λούβρο στο Παρίσι υπάρχει μια μαύρη βαβυλωνιακή στήλη με ανάγλυφο που δείχνει τον θεό ήλιο Shamash να παρουσιάζει τον κώδικα νόμων στο Hammurabi, ο οποίος πέθανε το 1750 π.Χ. Στην Παλαιά Διαθήκη ο Θεός δίνει τις Δέκα Εντολές στον Μωυσή, και χρονολογείται τον 14-13ο αιώνα π.Χ. Στον διάλογο του Πρωταγόρα, έργο του Πλάτωνα, υπάρχει μια προφανώς μυθική αφήγηση για το πώς ο Δίας λυπάται για τους άτυχους ανθρώπους, οι οποίοι φυσικά δεν ταιριάζουν με τα άλλα θηρία. Για να καλύψει αυτές τις ελλείψεις, ο Δίας έδωσε στους ανθρώπους μια ηθική αίσθηση και την ικανότητα για νόμο και δικαιοσύνη, έτσι ώστε να μπορούν να ζουν σε μεγαλύτερες κοινότητες και να συνεργάζονται μεταξύ τους.

2.1.2 Οι Σοφιστές

Τομή στην εξέλιξη της αρχαίας ελληνικής σκέψης αποτελεί η εμφάνιση και δράση του Σοφιστών, στα μέσα του 5ου αιώνα, τότε που η Αθήνα βρισκόταν στο απόγειο της ανάπτυξής της και δεν είχαν φανεί ακόμη οι πρώτοι καπνοί του Πελοποννησιακού πολέμου και τα πρώτα σημάδια της μετέπειτα παρακμής της. Οι σοφιστές τοποθέτησαν τον άνθρωπο στο επίκεντρο του ενδιαφέροντός τους, με κύριο όργανο τον ρητορικό λόγο διεκδίκησαν από τους φιλοσόφους τη διαπαιδαγώγηση των νέων, αμφισβήτησαν τις μέχρι τότε αδιαμφισβήτητες αξίες και πρότειναν καινούργια πρότυπα σκέψης και δράσης. Δίχασαν τους Αθηναίους σε φανατικούς οπαδούς και σε ορκισμένους αντιπάλους, κατηγορήσαν και κατηγορήθηκαν και αντί για διδάσκαλοι της πολιτικής αρετής στο τέλος θεωρήθηκαν απατεώνες και ψεύτες.

Οι σοφιστές δεν θέλησαν να ορίσουν τις ηθικές αρχές και να δημιουργήσουν έναν κώδικα ηθικής συμπεριφοράς, ωστόσο ο Σωκράτης είναι εκείνος που προσπάθησε με μέτρα καθολικά και αναλλοίωτα να ορίσει τις ηθικές έννοιες και να ρυθμίσει την ηθική αγωγή του ανθρώπου, ώστε να προσεγγίσει το ύψιστο αγαθό, την ευδαιμονία. Κινήθηκε πέρα από τα ηθικά φαινόμενα στην απόπειρά του να ανακαλύψει την αρχή κάθε ηθικής έννοιας, την ουσία της ηθικής. Αντίθετα με τους σοφιστές, απέρριψε το σχετικό και αναζήτησε το απόλυτο. Ταύτισε την αρετή με τη γνώση και υποστήριξε ότι ούδεις έκων κακός. Οι πλέον περίφημοι σοφιστές της αρχαιότητας ήταν οι Γοργίας, Πρωταγόρας, Ιππίας.

2.1.3 Ο Σωκράτης

Ο Σωκράτης θεωρείται ως ο θεμελιωτής της ηθικής φιλοσοφίας επειδή έστρεψε το φως της σκέψης και του προβληματισμού του από τη φύση στην ψυχή και την ηθική συμπεριφορά του ανθρώπου. Το «παράδοξο» είναι ότι τούτο το πέτυχε δίχως ποτέ ο ίδιος να γράψει

τίποτα ή να οργανώσει κάποια σχολή ή κάποιο φιλοσοφικό σύστημα. Απλά με τη δύναμη του κριτικού στοχασμού και με την καθαρότητα του προσωπικού του παραδείγματος ζωής - μιας ζωής που την οικοδόμησε σαν έργο τέχνης- άφησε τόσο υλικό για τους συγχρονους του και φίλους του ώστε αυτό αρκούσε για να γράψουν κατόπιν εκείνοι.

2.1.4 Ο Πλάτων

Ακολουθώντας τον δρόμο που χάραξε ο Σωκράτης, ο Πλάτων με τους διαλόγους του πρόβαλε και επέκτεινε τις σωκρατικές θεωρίες και προσπάθησε να ανατρέψει τις σοφιστικές αποκαλύπτοντας την πλάνη και τη φαινομενικότητα που τις χαρακτήριζε. Κινήθηκε όμως σε αμιγώς θεωρητικό επίπεδο και διαχωρίζοντας την ψυχή από το σώμα, το οποίο καταδίκασε σε δεσμοτήριο της ψυχής, δημιούργησε έναν ουτοπικό κόσμο, τον κόσμο των ιδεών, στον οποίο αναζήτησε την εὐδαιμονίαν.

2.1.5 Η Αριστοτελική Ηθική

Σύμφωνα με την Britannica Encyclopedia: “ο Αριστοτέλης υπήρξε ο πρώτος γνήσιος επιστήμονας στην ιστορία, και κάθε κατοπινός επιστήμονας του οφείλει κάτι” [29], υπήρξε μαθητής του Πλάτωνα, τον οδήγησε στις απόψεις του πλατωνισμού, αργότερα, όμως, μετά τον θάνατο του Πλάτωνα, οδηγήθηκε, περισσότερο, σε εμπειρικές μελέτες και μετατοπίζεται από τον πλατωνισμό στον εμπειρισμό. Πίστευε ότι οι ιδέες και οι γνώσεις όλων των λαών βασιζόταν, τελικά, στην αντίληψη. Στις απόψεις του, για τις φυσικές επιστήμες, βασίστηκαν πολλά έργα του.

Μαζί με το δάσκαλό του Πλάτωνα αποτελεί σημαντική μορφή της φιλοσοφικής σκέψης του αρχαίου κόσμου, και η διδασκαλία του διαπερνούσε βαθύτατα τη δυτική φιλοσοφική και επιστημονική σκέψη μέχρι και την Επιστημονική Επανάσταση του 17ου αιώνα. Υπήρξε φυσιοδίφης, φιλόσοφος, δημιουργός της λογικής και ο σημαντικότερος από τους διαλεκτικούς της αρχαιότητας.

Σύμφωνα με την εγκυκλοπαίδεια φιλοσοφίας του Stanford, τα Ηθικά Νικομάχεια είναι μία από τις τρεις πραγματείες που αναλύουν κυρίως τις μορφές και τις συνθήκες της ηθικής και διανοητικής αρετής του Αριστοτέλη. Στα Ηθικά Νικομάχεια ο Αριστοτέλης επισημαίνει ότι οι Αρετές είναι δύο ειδών: η διανοητική και η ηθική αρετή. Η διανοητική προκύπτει από την διδασκαλία και θέλει χρόνο και πείρα για να αποκτηθεί. Η ηθική αρετή είναι αποτέλεσμα του "έθους" (συνήθεια).

Ο Αριστοτέλης απαντά με ειλικρινή απλότητα ότι η καλή ζωή είναι η ευτυχισμένη ζωή. Εξετάζει στα “Ηθικά” του όχι πως να γίνονται οι άνθρωποι καλοί άνθρωποι, όπως ο Πλάτων,

αλλά πως να γίνονται ευτυχείς. Όλα τα άλλα πράγματα εκτός της ευδαιμονίας, αναζητούνται με κάποιο άλλο σκοπό. Μόνο η ευτυχία αναζητείται για τον εαυτό της. Ορισμένα πράγματα είναι αναγκαία για την διαρκή ευτυχία: η καλή γέννηση, η καλή υγεία, το κάλλος, η καλή τύχη, η καλή υπόληψη, οι καλοί φίλοι, τα καλά χρήματα και η καλοσύνη.

Η έννοια του αγαθού στον Αριστοτέλη

Κατά τον Πλάτωνα (στα βιβλία του Φαίδων και Πολιτεία) απαιτείται συγκεκριμένη γνώση του αγαθού που ενδιαθέτα συνδέεται με την σαφή κατανόηση των πραγμάτων και υπαγορεύει ένα δίκαιο τρόπο ζωής. Υπάρχει στην ψυχή και δεν ταυτίζεται με τα οικονομικά αγαθά. Χωρίζει τα αγαθά σε τρεις κατηγορίες, στα επαινετά, στα τίμια και στις δυνάμεις που αντιστοιχούν τα εξωτερικά αγαθά (τα δευτερεύοντα). Τα αγαθά είναι εκείνα που εξασφαλίζουν το “ζην”: κοινωνική θέση, περιουσία, σωματική ακεραιότητα και ασφάλεια.

Η αρετή στον Αριστοτέλη

Για τον Αριστοτέλη η αρετή είναι ύψιστο αγαθό. Η ενέργεια του νου και ο σκοπός της πράξης είναι αλληλένδετα και αλληλοεξαρτώμενα αλλά και το μέσον που πραγματώνει το σκοπό. Η ενέργεια του νου έχει δύο όψεις είναι και σκοπός και μέσο. Ως σκοπός η ενέργεια είναι ευδαιμονία, ως μέσον η ενέργεια ονομάζεται αρετή, έτσι, αρετή είναι η ενέργεια που υποκινείται από το νου και γι αυτό είναι σύμφωνη με την έλλογη φύση, που είναι το ιδιαίτερο γνώρισμα του ανθρώπου. Η αρετή δεν είναι φυσική προδιάθεση, αλλά κάτι που προκύπτει από δραστηριότητα και επίμονη άσκηση και δεν αποκτάται με τη διδασκαλία, αλλά με την πρακτική. Το παιδί μπορεί να μάθει μαθηματικά, αλλά δεν μπορεί να διδαχτεί την ηθική, αφού δεν έχει πείρα από τη ζωή και τις πράξεις. Επειδή οι αρετές σχετίζονται με την ανθρώπινη δραστηριότητα, η κατάταξη τους θα πρέπει να γίνει με βάση το πόσες είναι οι βασικές ανθρώπινες δραστηριότητες. Δύο είναι οι βασικές ανθρώπινες δραστηριότητες: η λογική και η δραστηριότητα που συμμορφώνεται με τη λογική. Η πρώτη αντιστοιχεί με τη διανοητική ενέργεια και η δεύτερη με την ηθική βούληση.

Η διανοητική αρετή στον Αριστοτέλη

Οι αρετές κατατάσσονται σε δύο κατηγορίες τις διανοητικές, δηλαδή τη σοφία και τη φρόνηση και τις ηθικές που συνδέονται με τη βούληση και γενικά τον ηθικό βίο. Οι διανοητικές αρετές δημιουργούνται από τη διδασκαλία, ενώ οι ηθικές από τον συνεχή εξάσκηση, ας μην ξεχνάμε ότι έθος σημαίνει συνήθεια.

Οι διανοητικές αρετές (π.χ. σοφία, φρόνηση, σύνεση) σχετίζονται με τη λογική και ανήκουν, σύμφωνα με το χωρισμό του Αριστοτέλη, στο καθαρά λογικό μέρος της ψυχής, το «λόγον ἔχον» μέρος. Παράγοντες που συμβάλλουν στη γένεση και την επαύξησή τους είναι κατά κύριο λόγο η διδασκαλία, η οποία απαιτεί εμπειρία και χρόνο

Η ευδαιμονία στον Αριστοτέλη

Ο Αριστοτέλης δίνει ιδιαίτερη βαρύτητα στην ευδαιμονία γιατί σε τελική ανάλυση κανείς δεν θέλει να είναι δυστυχής. Οι δραστηριότητες που αποτελούν την ευδαιμονία απαιτούν ανθρώπινη νόηση που έχει ωριμάσει μέσα από την πρακτική. Πράγματι: “Η ευδαιμονία μπορεί να είναι και κάτι κοινό σε όλους τους ανθρώπους, γιατί είναι δυνατόν να υπάρχει σε όλους τους ανθρώπους, που δεν είναι ανίκανοι να φτάσουν στην αρετή με τη βοήθεια κάποιας εκμάθησης και επιμέλειας”. Η ευδαιμονία είναι ταυτόχρονα και το άριστο και το πιο ωραίο και το πιο ευχάριστο, κι αυτά τα χαρακτηριστικά της δεν διαχωρίζονται σύμφωνα με το Δηλιακό επίγραμμα, που λέει ότι:

Το πιο δίκαιο είναι το πιο ωραίο απ' όλα και η υγεία το πιο ευάρεστο, κι από τη φύση το πιο ευχάριστο το να βρει κανείς εκείνο που αγαπά.

Το παραπάνω χωρίο επιχειρεί να δείξει πως είναι λογικό να υποθέσουμε ότι η ευδαιμονία είναι αποτέλεσμα προσπάθειας μάλλον παρά τύχης. Ο γενικός κανόνας στη φύση είναι ότι τα καλύτερα αποτελέσματα σε κάθε τομέα, προκύπτουν πάντα από κάποιο είδος σκόπιμης προσπάθειας. Πέραν αυτών ο Αριστοτέλης εξετάζει τα εξωτερικά αγαθά, το ρόλο που μπορεί να παίζουν στην ανθρώπινη ζωή, τον τρόπο και τον βαθμό που η ευ-τυχία ή η δυσ-τυχία καθορίζουν την ευδαιμονία. Η άποψη του Αριστοτέλη είναι πως εκείνο που καθορίζει την αξία της ζωής είναι οι πράξεις. Δεν πρέπει όμως να ξεχνάμε ότι ο άνθρωπος δεν μπορεί να πετύχει το “ευ ζην” χωρίς κανένα υλικό, όπως ο υποδηματοποιός δεν μπορεί να φτιάξει υποδήματα χωρίς δέρμα.

Τις πρώτες αναφορές για τους ανθρώπινους χαρακτήρες όπου αναπτύσσεται το Αριστοτελικό αξίωμα της μεσότητας. Πάντα, κατά τον Αριστοτέλη, η υπερβολή και η έλλειψη είναι δυο ακρότητες ανάμεσα στις οποίες πρέπει να εφαρμόζουμε το σωστό μέτρο. Για παράδειγμα η ανδρεία είναι το σωστό μέτρο ανάμεσα στο θράσος (παράτολμο) και στη δειλία, η σωφροσύνη ανάμεσα στην αναισθησία (απάθεια, αδιαφορία) και στην ακολασία, η γενναιοδωρία ανάμεσα στη φιλαργυρία (τσιγκουνιά) και στην σπατάλη. Η πραότητα είναι μεσότητα ως προς την οργή. Επειδή όμως το μέσον ως προς την οργή δεν έχει ιδιαίτερο όνομα, σχεδόν δεν έχουν και τα δύο άκρα, αποδίδουμε ως ιδιαίτερη ονομασία την πραότητα, αν και αυτή αποκλίνει προς την έλλειψη οργής που κι αυτή είναι χωρίς όνομα. Η υπερβολική οργή θα μπορούσε να ονομαστεί οξυθυμία, όπως οι πικρόχολοι που είναι αδιάλλακτοι και κρατούν την οργή τους για μεγάλο χρονικό διάστημα, ενώ ονομάζουμε δύστροπους τους ανθρώπους εκείνους που στενοχωριούνται και οργίζονται για πράγματα που δεν πρέπει. Επομένως οι αναφορές του Αριστοτέλη στους ανθρώπινους χαρακτήρες αφορούν την άσκηση της αρετής σχετικά με τον πλούτο, τα περιουσιακά στοιχεία, τις εν γένει συναναστροφές, τη συμβίωση και την επικοινωνία με λόγια και πράξεις (έν δὲ ταῖς ὀμιλίαις καὶ τῷ συζῆν καὶ λόγων καὶ πραγμάτων κοινωνεῖν).

Η δικαιοσύνη στον Αριστοτέλη

Η δικαιοσύνη ορίζεται ως αγαθό που δεν στοχεύει στην ευδαιμονία όποιου την ασκεί, αλλά στον άλλον άνθρωπο. Ο Αριστοτέλης διακρίνει τη δικαιοσύνη σε τρία είδη:

- Τη διανεμητική που απονέμεται με βάση την αξία του ανθρώπου. Για τους δημοκρατικούς όμως αξία αποτελεί η ελευθερία, για τους ολιγαρχικούς ο πλούτος ή η καταγωγή και για τους αριστοκρατικούς η αρετή. Το διανεμητικό δίκαιο δεν περιορίζεται στην αναγνώριση των δικαιωμάτων τους ενός ή του άλλου, αλλά συνιστά εφαρμογή, πραγματοποιήσιμη της δικαιοσύνης. Προϋποθέτει επομένως την ύπαρξη δημόσιων αρχών που πρέπει να κυριαρχούνται από την ιδέα της δικαιοσύνης και τη θέληση να την κάνουν να επικρατήσει σε κάθε περίπτωση.
- Τη διορθωτική που απονέμεται με βάση την αρχή όλα τα άτομα είναι ίσα μεταξύ τους. Η άνιση μεταχείριση όμως των ίσων είναι εξίσου άδικη της ίσης μεταχείρισης των ανίσων.
- Της αμοιβαιότητας που οφείλεται στην ελεύθερη βούληση των μελών της κοινωνίας και είναι ιδιαίτερα σημαντική γιατί συντελεί στην ενότητα της

Η ανισότητα στον Αριστοτέλη

Όλοι οι άνθρωποι γεννιούνται άνισοι. “Η ισότητα είναι δίκαιη, αλλά μόνο μεταξύ ίσων”. Και οι ανώτερες τάξεις θα εξεγερθούν εξίσου εύκολα, αν επιβληθεί αφύσικη ισότητα, όσο και οι κατώτερες τάξεις, όταν η ανισότητα επενεργεί ως αφύσικη ακρότητα. Όταν η δημοκρατία διέπεται από τις κατώτερες τάξεις, οι πλούσιοι φορολογούνται για να προμηθεύσουν χρήματα στους φτωχούς. “Οι πτωχοί τα παίρνουν και τα ζητούν και πάλι και είναι σαν να χύνει κανείς νερό μέσα σε ένα κόσκινο”. Επειδή όλοι είναι με τον ίδιο τρόπο ελεύθεροι, θεωρούν πως είναι με τον ίδιο τρόπο ίσοι. Η ισότητα είναι ισότητα προς την αξία του καθενός, όχι ισότητα μεταξύ στοιχείων με διάφορη αξία.

Η μεσότητα στον Αριστοτέλη

Σύμφωνα με το Stanford Encyclopedia of Philosophy έχουμε έναν πίνακα των αρετών σύμφωνα με τη θεωρία της Μεσότητας του Αριστοτέλη. Στον παρακάτω πίνακα φαίνεται πώς εννοεί ο Αριστοτέλης τη μεσότητα και πώς την έλλειψη και την υπερβολή, πώς εφαρμόζει την ιδέα της μεσότητας.

Θέμα	Έλλειψις	Μεσότης - Αρετή	Υπερβολή
Φόβος και	Δειλία	Ανδρεία	Θρασύτης

Θάρρος	Το να φοβάται κανείς, η ανανδρία	Η ανδροπρέπεια, η γενναία πράξη	Η αυθάδεια, η προπετεία
Ηδονές	Αναισθησία Η απάθεια, η αδιαφορία	Σωφροσύνη Η Εγκράτεια	Ακολασία Η παραλυσία, η φιληδονία
Χρήμα και Πλούτος	Ανελευθερία Η δουλικότητα, η φιλαργυρία	Ελευθεριότης Η γενναιοδωρία, η απλοχεριά	Ασωτία Η σπατάλη, η κατάχρηση
	Ζημία Όταν χάνεις στις συναλλαγές	Δίκαιον Όταν στις συναλλαγές ούτε χάνεις ούτε κερδίζεις	Κέρδος Όταν στις συναλλαγές κερδίζεις
	Μικροπρέπεια Χορηγία με δισταγμό και συνεχή παράπονα, του δωρητή	Μεγαλοπρέπεια Μεγαλείο γενναιοδωρίας. Μεγάλες χορηγίες για το δημόσιο καλό	Απειροκαλία Η επίδειξη πλούτου, η ακαλαισθησία, η κακογουστιά
Τιμές	Μικροψυχία Η διεκδίκηση μικρότερων έργων και αγαθών, σε σχέση με την αξία που έχεις	Μεγαλοψυχία Η αξιοσύνη για τα μεγάλα έργα και τιμές. Η γενναιοφροσύνη	Χαυνοτης Η διεκδίκηση μεγάλων έργων και αγαθών χωρίς αληθινή αξιοσύνη
	Αφιλοτιμία Η μη επιθυμία προς διάκριση	Φιλοτιμία Η υγιής φιλοδοξία, η αγάπη προς τις τιμές	
Οργή	Αοργησία Η ηρεμία, η ψυχραιμία	Πραότης Η γλυκύτητα, η ηπιότητα	Οργιλότης Η οξυθυμία, η αφάδα
Αλήθεια	Ειρωνεία Η κρυψίνοια, η προσποιητή άγνοια	Αλήθεια Η πραγματικότητα, η ακρίβεια, η ειλικρίνεια	Αλαζονεία Ο κομπασμός, η έπαρση, η μεγαλοστομία
Διασκέδαση και Γλυκύτητα Βίου	Αγροικία Η τραχύτητα τρόπων	Ευτραπελία Η ευφυής αστειότης	Βωμολοχία Η χυδαιολογία
	Δυσκολία Η δυστροπία, η εναντιότης, η δυσαρέσκεια.	Φιλία Η φιλική διάθεση ως αρμόζει, η αγάπη, η εύνοια	Κολακεία Η υπερβολική φιλοφροσύνη
Συναισθήματα σε σχέση με τους άλλους	Αναισχυντία Η αναίδεια, η αδιαντροπία	Αιδώς Η ντροπή, ο σεβασμός, η συστολή, η σεμνότητα	Κατάπληξις Ο ισχυρός θαυμασμός, το σάστισμα
	Επιχαιρεκακία Η ανικανότητα για λύπη, η χαρά για τα ξένα παθήματα	Νέμεσις Η δίκαιη αγανάκτηση η λύπη για όσους ευτυχούν αδίκως	Φθόνος Η λύπη γενικώς για την ευτυχία των άλλων

2.1.6 Ο Kant

Η Καντιανή ηθική σύμφωνα με την Stanford Encyclopedia of Philosophy [46] αναφέρεται σε μια δεοντολογική ηθική θεωρία που αποδίδεται στον Γερμανό φιλόσοφο Immanuel Kant. Η θεωρία, που αναπτύχθηκε ως αποτέλεσμα του ορθολογισμού του Διαφωτισμού, βασίζεται στην άποψη ότι το μόνο εγγενώς καλό πράγμα είναι η καλή θέληση. Η αρχή της καθολικότητας απαιτεί ότι, για να επιτρέπεται μια ενέργεια, πρέπει να είναι δυνατή η εφαρμογή της σε όλους τους ανθρώπους χωρίς να υπάρχει αντίφαση. Εάν προκύψει αντίφαση, η πράξη παραβιάζει την έννοια “μη αντίφαση” του Αριστοτέλη, η οποία δηλώνει ότι οι πράξεις δεν μπορούν να οδηγήσουν σε αντιφάσεις. Η διατύπωση του Καντ στο δεύτερο κεφάλαιο του έργου του “Κατηγορηματική Προσταγή”, δηλώνει ότι:

“ως αυτοσκοπός οι άνθρωποι δεν απαιτείται ποτέ να αντιμετωπίζονται τους άλλους απλώς ως μέσο για τον σκοπό, αλλά πάντα ως αυτοσκοπός”.

Ο Καντ έκανε επίσης διάκριση μεταξύ τέλειων και ατελών καθηκόντων. Ένα τέλειο καθήκον, όπως “το καθήκον να μην ψεύδεται, ισχύει πάντα, ενώ ένα ατελές καθήκον, όπως το καθήκον να είμαστε φιλόστοργοι, μπορεί να γίνει ευέλικτο και να εφαρμοστεί σε συγκεκριμένο χρόνο και τόπο”.

Ο Καντ δημιούργησε τη βάση ενός ηθικού νόμου με την έννοια του καθήκοντος. Θεμελίωσε την ηθική του θεωρία υποστηρίζοντας ότι η μόνη αρετή που μπορεί να είναι ανεπιφύλακτα καλή είναι η καλή θέληση. Οι υπόλοιπες αρετές μπορούν να χρησιμοποιηθούν για την επίτευξη ανήθικων σκοπών. Για παράδειγμα, η αρετή της πίστης δεν είναι καλή αν κάποιος είναι πιστός σε ένα κακό άτομο. Θεμέλιο κάθε ηθικής φαίνεται να θεωρεί ο Καντ την απόλυτα καλή θέληση, η οποία δεν εξαρτάται από κανένα εξωτερικό κίνητρο, ούτε από την επίτευξη κάποιου επιδιωκόμενου σκοπού, αλλά είναι στην ουσία καλή και η οποία ‘είναι απαραίτητη για να είναι κανείς άξιος ευτυχίας. Με άλλα λόγια, μπορεί να αποτελέσει το κίνητρο των ηθικών μας πράξεων η αναζήτηση ευτυχίας, δεν είναι όμως ο αίτημα για ευτυχία ηθικό αξίωμα ούτε σκοπός ηθικότητας η επίτευξη της ευτυχίας. Επισημαίνει ευθέως ότι: *“Η κατεξοχήν ηθική δεν είναι διδασκαλία του πως θα γίνουμε ευτυχισμένοι, αλλά του πως θα γίνουμε άξιοι της ευτυχίας”.*

Η αντίληψη του Καντ για το καθήκον εμπεριέχει την έννοια της απόλυτα καλής θέλησης, πράγμα που αποτελεί το μοναδικό δρόμο για την τέλεση μιας πραγματικά ηθικής πράξης. Ισχυρίζεται λοιπόν ότι το μοναδικό πράγμα το οποίο μπορεί να δώσει σε μια ενέργεια ή



Immanuel Kant
Πορτραίτο του Johann Gottlieb
Becker, 1768

πράξη ηθική αξία, είναι να γίνεται από αίσθηση καθήκοντος. Γιατί αν κανείς πράττει μόνο και μόνο γιατί κάτι του αρέσει δεν έχει ηθική αξία. Ο Καντ εφιστά την προσοχή μας ώστε να μη γίνεται σύγχυση μιας πράξης που εκτελείται από καθήκον με μια πράξη που γίνεται σύμφωνα με το καθήκον, αφού η πρώτη παράγει την ηθικότητα και η δεύτερη τη νομιμότητα. Για να έχει όμως μια πράξη αληθινή ηθική αξία, θα πρέπει να γίνεται από καθήκον, δηλαδή μόνο και μόνο επειδή τέτοια είναι η ηθική προσταγή και όχι επειδή έτσι θα προκύψουν κάποια ατομικά οφέλη. Κατά τον φιλόσοφο, με άλλα λόγια, οι πράξεις μας πρέπει να συμφωνούν κατά βάθος με τον ηθικό νόμο και να πηγάζουν από τη συναίσθηση του καθήκοντος. Με αυτό τον τρόπο χρησιμοποιεί το παράδειγμα ενός εμπόρου που παρά την έντονη κίνηση και ζήτηση των καταναλωτών δεν αυξάνει την τιμή των προϊόντων του αλλά τη διατηρεί σταθερή και προσιτή σε όλους, χωρίς να σκέφτεται τους ευκαιριακούς ρυθμούς οικονομικής ανόδου που έχει με τη δυνατή υπερτιμολόγηση.

Η απάντηση στο ερώτημα επομένως αν είναι μια καλή πράξη απαραίτητα και ηθική είναι αρνητική. Επειδή ένας άνθρωπος που του αρέσει να συμπεριφέρεται καλά στους άλλους λόγω του ότι έχει τέτοια προδιάθεση, είναι μεν τυχερός γιατί γεννήθηκε με κοινωνικά επιθυμητό χαρακτηριστικό αλλά συμπεριφέρεται σύμφωνα με την προδιάθεση του με παρόμοιο τρόπο όπως και τον άλλον που εκμεταλλεύεται τους άλλους. Συνεπώς, εδώ ο φιλόσοφος κάνει μια διάκριση στις πράξεις εκείνες που είναι σύμφωνες με τον καθήκον και εκείνες που γίνεται από καθήκον. Οι πρώτες δεν παραβιάζουν τον ηθικό νόμο παρά το ότι δεν γίνονται για χάρη του, αλλά δεν είναι όμως και ενάρετες. Ενώ οι δεύτερες γίνονται από καθήκον και είναι ενάρετες. Όπως για παράδειγμα. Το άτομο που σκορπίζει γύρω του την ευτυχία από συναίσθηση καθήκοντος, το κάνει γιατί ξέρει ότι αυτό ακριβώς είναι εκείνο που απαιτεί ο ηθικός νόμος.

Στη συνέχεια ο Καντ, θα προχωρήσει στη διατύπωση της θέσης ότι: *“η ηθική αξία μιας πράξης που γίνεται από καθήκον δεν έγκειται στον σκοπό, ο οποίος επιδιώκεται με αυτή την πράξη, αλλά στον γνώμονα, σύμφωνα με τον οποίο η πράξη αποφασίζεται. Η ηθική αξία δεν εξαρτάται έτσι από την πραγματοποίησης και από τα αποτελέσματα της πράξης, παρά μόνο από το αξίωμα της θέλησης, σύμφωνα με το οποίο η πράξη εκτελείται άσχετα από όλα όσα πετυχαίνονται ή ποθούνται”*. Σύμφωνα με αυτή την πρόταση είναι κατανοητό ότι αυτό που δίνει σε μια ενέργεια ηθική αξία είναι το αξίωμα, η υποκειμενική αρχή, με βάση της οποίας ενεργεί ο παράγων άνθρωπος. Φαίνεται λοιπόν ότι η θέληση βρίσκεται μπροστά σε ένα “σταυροδρόμι”, ανάμεσα στο a priori αξίωμα της και το a posteriori κίνητρό της. Με την έννοια a priori, ο Καντ εννοεί την αλήθεια εκείνη που είναι ανεξάρτητη από κάθε γνώση και γίνεται γνωστή απλά μέσα από τη λογική, ενώ η έννοια που φέρει η έκφραση a posteriori παραπέμπει σε κάθε γνώση που γίνεται αντιληπτή μόνο μέσα από κάποια εμπειρία.

Η Μεταφυσική των Ηθών ως Θεμέλιο της Ηθικής

Στη συνέχεια ο Καντ θα δείξει έμπρακτα ότι αποστασιοποιείται από κάθε εμπειρική σχέση ή καταβολή της ηθικής, αφού γίνεται αντιληπτό ότι ούτε η ηθικότητα μιας πράξης, ούτε και η έννοια του καθήκοντος μπορούν να διαπιστωθούν εμπειρικά. Για αυτό ασκεί δριμεία κριτική στον εμπειρισμό του David Hume, ο οποίος απέδωσε την αναγκαιότητα σύνδεσης αιτίου και αποτελέσματος στην ανθρώπινη συνήθεια. Παρατηρεί λοιπόν, ότι ο εμπειρισμός εξολοθρεύει από τη ρίζα της την ηθικότητα του φρονήματος και στη θέση της βάζει κάτι εντελώς διαφορετικό, το εμπειρικό συμφέρον, με το οποίο συνδέονται όλες οι φυσικές ροπές αντί να εγκαθιδρύσει το καθήκον. Καταλαβαίνουμε ότι ο Καντ αναζητεί την απόλυτα καθαρή πηγή των πράξεων μας. Πέρα από κάθε εμπειρία εξυψώνει αφενός την έννοια του καθήκοντος με προσήλωση και αφετέρου, την έννοια της λογικής η οποία καθορίζει τη θέληση μας.

Θα πρέπει συνεπώς δεχόμενοι την αλήθεια της ηθικής να δεχτούμε επίσης την οικουμενικότητα του ηθικού νόμου ο οποίος ισχύει απόλυτα για όλα τα έλλογα όντα τονίζοντας χαρακτηριστικά ότι τίποτα δε βλάπτει περισσότερο την ηθικότητα από την αντίληψη ότι αυτή προκύπτει από παραδείγματα. Γιατί κάθε παράδειγμα που τοποθετείται μπροστά μας πρέπει πρώτα να εξεταστεί με αξιώματα της ηθικότητας, ώστε να κριθεί αν αξίζει να χρησιμεύσει ως πρότυπο. Στην ουσία όμως κανένα από τα παραδείγματα δεν είναι σε θέση σύμφωνα με τον Καντ να προσδώσει την ύψιστη έννοια της ηθικότητας, εκτός από εκείνο που θεμελιώνεται στην καθαρή λογική ανεξάρτητα από κάθε εμπειρία.

Στόχος του Καντ είναι η θεμελίωση της ηθικής πάνω στα αξιώματα της καθαρής πρακτικής λογικής. Κάθε τι στη φύση λειτουργεί με νόμους που μόνο τα έλλογα όντα έχουν τη δυνατότητα να ενεργήσουν σύμφωνα με αξιώματα, συμπεραίνει ότι μόνο τα έλλογα όντα έχουν θέληση επειδή βασίζονται στην πρακτική λογική. Σύμφωνα με τον Καντ βασική ιδιότητα της λογικής είναι η προσταγή, που για τον ίδιο το διανοητή, είναι “ένα αντικειμενικό αξίωμα”, κατά το μέτρο που εξαναγκάζει τη θέληση, ονομάζεται προσταγή της λογικής και η διατύπωση της προσταγής ονομάζεται προστακτική.

Στη συνέχεια για τη διατύπωση των προστακτικών, επισημαίνει ότι όλες αυτές εκφράζονται με το επίρρημα “πρέπει” και στην ουσία είναι διατυπώσεις που εκφράζουν τη σχέση των αντικειμενικών νόμων της θέλησης του κάθε έλλογου όντος. Εξάλλου, οι προστακτικές αυτές είναι υποθετικές που δείχνουν τι πρέπει να εκτελεστεί εάν υπάρχει θέληση να επιτευχθεί ένας σκοπός και κατηγορικές αν διατυπώνουν ένα “πρέπει” που είναι αυτοσκοπός και όχι μέσο για κάποιο σκοπό. Ο Καντ αναγνωρίζει μόνο τις κατηγορικές από τα διάφορα είδη των προστακτικών, ως σύμφωνες με το καθήκον και την ηθική. Ωστόσο, υποπίπτει πάνω στο προβληματισμό, ότι οι προστακτικές έχουν ως αποτέλεσμα τον εξαναγκασμό της θέλησης και αυτό έχει ως αποτέλεσμα τη στέρηση κάθε ελευθερίας.

Ο Καντ φτάνει στη διατύπωση μιας και μόνο προσταγής η οποία σημαίνει “πράττε μόνο σύμφωνα με ένα τέτοιο γνώμονα, μέσω του οποίου μπορείς συνάμα να θέλεις αυτός ο γνώμονας να γίνει καθολικός νόμος”. Ουσιαστικά μας προτρέπει να σκεπτόμαστε από μόνοι

μας και να στοιχίζουμε τις πράξεις μας σύμφωνα πάντα με τον ηθικό νόμος χωρίς να εξαρτόμαστε από κάποιον άλλο. Έτσι λοιπόν σκέψεις και πράξεις χαρακτηρίζονται από πλήρη αυτονομία, η οποία όμως να έχει τα κοινωνικά της όρια. Ο περιορισμός της ελευθερίας μας είναι ότι πρέπει να σεβόμαστε την ελευθερία όλων, διαφορετικά πώς μπορεί η ελευθερία μας να προέρχεται από καθολικούς νόμους;

Αυτό που απασχολεί τον Καντ και το έργο του Μεταφυσική των Ηθών, είναι όχι η θεωρητική φιλοσοφία αλλά μια πρακτική φιλοσοφία, η οποία δεν ενδιαφέρεται για τις αιτίες όσων συμβαίνουν αλλά για τους νόμους όσων πρέπει να συμβαίνουν έστω κι αν δεν συμβαίνουν ποτέ, δηλαδή για τους αντικειμενικούς πρακτικούς νόμους. Θεμελιώνει λοιπόν την κατηγορική προσταγή, δηλαδή τον ηθικό νόμο στην ιδέα ότι ο άνθρωπος και γενικά κάθε έλλογο ον υπάρχει ως σκοπός και όχι ως μέσο, για την αυθαίρετη χρήση κάποιας θέλησης. Αν λοιπόν κάθε άνθρωπος αντιληφθεί αυτό το αξίωμα ατομικά, τότε γίνεται αντικειμενικό και καθολικό αξίωμα το οποίο συνεπάγεται ο νόμος της θέλησης. Σχετικά με αυτό διατυπώνει ο Γερμανός φιλόσοφος: *“πράττε έτσι ώστε να χρησιμοποιείς την ανθρωπότητα, τόσο στο πρόσωπό σου, όσο και στο πρόσωπο κάθε άλλου ανθρώπου, ταυτόχρονα πάντα ως σκοπό και ποτέ ως μέσο”*.

Συνεχίζοντας ο Καντ διαπιστώνει ότι παρά την υποκειμενικότητα της η έλλογη θέση δεν πρέπει να νομοθετεί υποκειμενικά, αλλά να στοχεύει προς τα έλλογα όντα ως αυτοσκοπός και να τους θέτει καθολικά έγκυρους νόμους ως ένα βασίλειο σκοπών. Επηρεασμένος κυρίως από τη θεωρία του Leibniz, ο Καντ εισάγει την έννοια ενός κράτους των σκοπών. Πρόκειται για ένα ιδεατό κράτος το οποίο αποτελεί τη συνταγματική ένωση διαφόρων έλλογων όντων μέσω κοινών νόμων. Τα έλλογα όντα απαλλαγμένα από υποκειμενικές ροπές νομοθετούν και αυτο-υποτάσσονται στον καθολικό νόμο σαν αρχηγοί μα ταυτόχρονα και σαν ισότιμα μέλη του με κοινά καθήκοντα, με αρχηγικό ρόλο μόνο στο πλαίσιο που μπορεί να νομοθετεί σαν καθολικός νομοθέτης, αφού ακόμη και στην ιδανική περίπτωση ο άνθρωπος δεν είναι πλήρως απαλλαγμένος από την εμπειρική και υποκειμενική του πλευρά. Εκείνο που δηλώνει λοιπόν εδώ, είναι η διατύπωση μιας νέας πρότασης της κατηγορικής προσταγής όπως λέει ο ίδιος: *“κάθε λογικό ον ενεργεί ως να ήταν μέλος του παγκόσμιου βασιλείου των σκοπών”*. Τέτοια προσταγή παραπέμπει στο γεγονός ότι κάθε θεώρηση σκοπών βάζει αίτημα την ύπαρξη ιδανικού κόσμου μέσα στον οποίο δεν υπάρχουν συγκρούσεις με τη λογική και το κάθε έλλογο ον είναι υποταγμένο στον επικρατούντα κανόνα ή νόμο.

Η οριστική εκδοχή αυτή της κατηγορικής προστακτικής φαίνεται ως ένας πιο ρομαντικός τρόπος έκφρασής της. Παρατηρείται επίσης και μια φαινομενική αντίθεση στην ιδέα αυτή, με την αναγνώριση ότι οι νόμοι στο κράτος των σκοπών καθορίζουν τη σχέση των όντων μεταξύ τους ως σκοπούς και ως μέσα ταυτόχρονα. Μια ιδιότυπη δηλαδή και στενή εξάρτηση μέσου και σκοπού, υποκειμενικού γνώμονα και αντικειμενικού νόμου, εξαναγκασμού και ελευθερίας, παρά την αυστηρή διάκριση σκοπού και μέσου που προηγήθηκε.

Τέλος, το αγαθό ευτυχία προϋποθέτει την ικανοποίηση των επιθυμιών του ανθρώπου, αλλά αυτό δεν είναι εφικτό λόγω του ότι ο ίδιος ο άνθρωπος δεν είναι δημιουργός του κόσμου. Επομένως, ο άνθρωπος δεν είναι σε θέση να έχει σχεδιάσει τα πράγματα έτσι ώστε να υπάρχει αντιστοιχία βούλησης και πραγματικότητας, γι αυτό το λόγο χρειάζεται ο Θεός. Με άλλα λόγια ο Θεός υπάρχει κατά τον Καντ ως αίτημα του πρακτικού λόγου.

2.1.7 Η Αναλυτική φιλοσοφία

Η Αναλυτική φιλοσοφία είναι είδος φιλοσοφίας το οποίο έγινε κυρίαρχο στις αρχές του 20ου αιώνα, με κύριους εκπροσώπους τον Bertrand Russell, τον Ludwig Wittgenstein, G. E. Moore, Gottlob Frege, και οι φιλόσοφοι του λογικού θετικισμού. Ως φιλοσοφική πρακτική, χαρακτηρίζεται από την έμφαση στην σαφήνεια της επιχειρηματολογίας, συχνά κάνοντας χρήση της λογικής, ανάλυσης και σε λιγότερο βαθμό των μαθηματικών και φυσικών επιστημών. Η αναλυτική φιλοσοφία δεν είναι κλάδος ή επιμέρους αντικείμενο της φιλοσοφίας αλλά αποτελεί τρόπο διεξαγωγής της φιλοσοφίας γενικά. Το βασικό της χαρακτηριστικό είναι ότι ως μορφή φιλοσοφίας λαμβάνει υπόψη όλα τα είδη γνώσης και όλα τα αντικείμενα των άλλων επιστημών. Από την αρχή η αναλυτική φιλοσοφία ήταν ένα κίνημα που ενσωματώνει στη φιλοσοφία τις εξελίξεις που προκύπτουν στις άλλες επιστήμες. Αρχικά ενσωμάτωσε στη φιλοσοφική πρακτική συμπεράσματα και πρακτικές από τις τυπικές επιστήμες (τη λογική και τα μαθηματικά). Συγχρόνως δέχθηκε ότι τα συμπεράσματα των εμπειρικών επιστημών (φυσική, χημεία, βιολογία, πειραματική ψυχολογία κτλ.) και οι μέθοδοί τους (παρατήρηση, πειράματα, μέτρηση, κανόνες συναγωγής) θα μπορούσαν να καταλαμβάνουν σημαντική, ίσως και κεντρική, θέση στη φιλοσοφική εικόνα του κόσμου. Πιο πρόσφατα η αναλυτική φιλοσοφία ενσωμάτωσε στη φιλοσοφική εικόνα του κόσμου και τα ευρήματα επιστημών του εγκεφάλου όπως η γνωσιακή ψυχολογία, οι νευροεπιστήμες και οι επιστήμες της όρασης.



Bertrand Russell, Ludwig Wittgenstein, George Edward Moore
Κεντρικά πρόσωπα στην ανάπτυξη της Αναλυτικής φιλοσοφίας μαζί με τον G. Frege

Η αναλυτική φιλοσοφία ανέτρεψε αυτή την κατάσταση και εξισορρόπησε τη φιλοσοφία ώστε να λαμβάνονται υπόψη οι τυπικές και εμπειρικές επιστήμες. Αρχικά αυτή η εξισορρόπηση εστίασε στις τυπικές (λογικές) και επιστημονικές πλευρές της γνώσης σε τέτοιο βαθμό που παραμέρισε άλλες πλευρές της γνώσης—με αποτέλεσμα η ίδια η πράξη εξισορρόπησης να μην είναι πολύ ισορροπημένη υπόθεση. Όμως σταδιακά η αναλυτική φιλοσοφία διόρθωσε αυτή την αρχική της τάση. Έδειξε με αυτό τον τρόπο ότι ήταν μια ευρύτερη πρακτική και όχι μια σχολή σκέψης που εξαρτάται από πάγιες φιλοσοφικές θέσεις. Σήμερα, η αναλυτική φιλοσοφία είναι πολύ ενεργή σε κάθε τομέα της φιλοσοφίας, από τον αναλυτικό Μαρξισμό μέχρι τη μεταφυσική και από την αισθητική μέχρι τη φιλοσοφία των μαθηματικών.

Ο G.E Moore

Ο George Edward Moore ήταν Άγγλος φιλοσοφος, και μαζί με τους Bertrand Russell και Ludwig Wittgenstein θεωρείται ένας από τους ιδρυτές της αναλυτικής παράδοσης στη φιλοσοφία. Σε συνεργασία με τον Russell, απομάκρυνε τη φιλοσοφία του Ηνωμένου Βασιλείου από τον ιδεαλισμό, και έγινε γνωστός για την υπεράσπιση των εννοιών της κοινής λογικής, τις συνεισφορές του στην ηθική, την επιστημολογία και τη μεταφυσική και για την “εξαιρετική του προσωπικότητα και ηθικό χαρακτήρα”. Ο Moore είναι πλέον γνωστός σήμερα για την υπεράσπιση του ηθικού μη-φυσιοκρατισμού, την έμφαση του στην κοινή λογική στη φιλοσοφική μέθοδο και το παράδοξο που φέρει το όνομά του. Τα δοκίμια του Moore είναι γνωστά για το σαφές, προσεκτικό τρόπο γραφής και για τη μεθοδική και υπομονετική του προσέγγιση στα φιλοσοφικά προβλήματα. Ήταν επικριτικός για τη σύγχρονη φιλοσοφία για την έλλειψη προόδου, την οποία πίστευε ότι ήταν σε πλήρη αντίθεση με τις ραγδαίες εξελίξεις στις φυσικές επιστήμες από την Αναγέννηση. Μεταξύ των πιο διάσημων έργων του Moore είναι το βιβλίο του Principia Ethica και τα δοκίμια του, "Η διαψεύσεις του ιδεαλισμού", "Μια υπεράσπιση της κοινής λογικής" και "Μια απόδειξη του εξωτερικού κόσμου".



Ο G. E. Moore το 1914

Principia Ethica

Το έργο του G. E. Moore, Principia Ethica άσκησε επιρροή στην εξέλιξη της ηθικής φιλοσοφίας, το οποίο συμπυκνώνει τις μετα-ηθικές αντιπαραθέσεις του προηγούμενου

αιώνα. Ο Moore χρησιμοποίησε σαν κύρια μέθοδο του την περιγραφή της λογικής συμπεριφοράς των γλωσσικών εκφράσεων της κοινής γλώσσας και την ανάλυση των σύνθετων αντικειμένων και εννοιών στα συστατικά τους. Τη γλωσσική ανάλυση την έβλεπε ως μέσον για το τι εννοούν οι διάφορες λεκτικές εκφράσεις. Πίστευε πως γνωρίζουμε ότι μια πρόταση είναι αληθινή, χωρίς να γνωρίζουμε τη σωστή ανάλυση του νοήματος της και χωρίς να μπορούμε να δώσουμε τους αναγκαίους και επαρκείς όρους για το νόημά της. Δεν θεώρησε την γλωσσική ανάλυση αυτοσκοπό, αλλά μέσο για την διακρίβωση γεγονότων. Ο Moore ως φιλοσοφία εννοούσε την προσπάθεια ανακάλυψης απαντήσεων σε ερωτήσεις, σύμφωνα με τη σωστή μέθοδο (ανάλυση), εφαρμοσμένη πάνω στη γραμματική της "καθημερινής γλώσσας".

Δεχόταν πως, παράλληλα στα υλικά αντικείμενα, "υπάρχουν" και οι συνειδησιακές (νοητικές) πράξεις, προσαρτημένες στα υλικά αντικείμενα. Εφόσον τα αντικείμενα είναι ανεξάρτητα από τη συνείδηση, η γνώση μας γι' αυτά προέρχεται από την παρατήρηση. Οι πράξεις της συνείδησης και μη φυσικές ποιότητες είναι επίσης γνωσιακά (cognitive) αντικείμενα.

Ο Moore θέτει ως στόχο κατ' αρχήν την καθαρή διάκριση δύο ειδών ερωτήσεων:

1. Τι είδους πράγματα πρέπει να υπάρχουν. Αφορά το αγαθό, αυτό που έχει αληθινή (intrinsic) αξία και που πρέπει να υπάρχει ως αυτοσκοπός και όχι ως μέσο για την ύπαρξη κάποιου άλλου
2. Τι είδους πράξεις είναι χρέος μας να εκτελούμε όσον αφορά το καθήκον.

Αντίστοιχο με τις ερωτήσεις αυτές είναι το πρόβλημα: "ποια είναι η φύση των αποδεικτικών στοιχείων (evidence) με τα οποία μπορεί να αποδειχθεί ή να ανασκευαστεί μια ηθική πρόταση"; Σχετικά με το πρώτο ερώτημα θεωρεί ότι καμία απόδειξη δεν μπορεί να προσαχθεί, ενώ κάθε απάντηση στην δεύτερη επιδέχεται απόδειξη ή αναίρεση. Πρέπει όμως να περιέχουν προτάσεις δύο ειδών: "αλήθειες σχετικές με τα αποτελέσματα της εν λόγω πράξης (αιτιακές αλήθειες) και ηθικές αλήθειες της πρώτης αναπόδεικτης τάξης".

Αποκαλεί τις ηθικές προτάσεις της πρώτης τάξης "ενοράσεις", χωρίς να θεωρεί τον εαυτό του διαισθησιοκράτη. γιατί δεν δέχεται ότι προτάσεις σχετικές με το καθήκον είναι ενοράσεις όπως οι γνήσιοι αισθησιοκράτες, αλλά θεωρεί ενοράσεις μόνο αυτές που αναφέρονται στο αγαθό. Χαρακτηρίζει δηλαδή ενοράσεις τις έννοιες και τις προτάσεις που δεν επιδέχονται ορισμό, χωρίς όμως να συζητά τον τρόπο και την προέλευση της γνώσης τους. Ούτε θεωρεί όλες τις ενοράσεις αληθινές. Σε αντιδιαστολή με την διακρίβωση της αλήθειας των ηθικών κρίσεων, των προτάσεων που περιέχουν τους κοινούς ηθικούς όρους (αρετή, κακία, χρέος κλπ.) θεωρεί ο Moore αντικείμενο της ηθικής κυρίως "τη γενική έρευνα για ό,τι είναι αγαθό". Δεν λογίζει το πρόβλημα της ηθικής συμπεριφοράς πρωταρχικό, αλλά επακόλουθο της σωστής απάντησης στα ερωτήματα "τί είναι αγαθό"; και "τί είναι συμπεριφορά" (conduct); Γιατί δεν είναι λογικά απαραίτητο κάθε είδος συμπεριφοράς να είναι ορθό, ούτε η ορθή συμπεριφορά είναι το μόνο αγαθό. Ως ορισμό (ανάλυση του αντικείμενου της σκέψης στα συστατικά του) ο Moore δεν εννοεί το λεκτικό ή λεξικογραφικό, ούτε τη σωστή χρήση της

λέξης, όπως καθιερώθηκε από τη συνήθεια, αλλά την καθημερινή και κοινή. "Το αγαθό είναι αγαθό" (ταυτολογία), δεν δέχεται ότι μπορεί να οριστεί. Οι προτάσεις για το αγαθό είναι όλες συνθετικές, ποτέ αναλυτικές. Το αγαθό (όπως η λέξη κίτρινο) είναι έννοια απλή όχι σύνθετη. Οι απλοί όροι (μέρη ενός σύνθετου) δεν ορίζονται. Μπορεί κανείς απλώς να τους σκεφτεί ή να τους αντιληφθεί, ενώ αντικείμενα που ούτε είναι ούτε υπάρχουν, όπως η χίμαιρα, μπορούν να οριστούν γιατί είναι σύνθετα. Ως έννοια απλή που δεν αναλύεται σε μερικότερες, το "αγαθό" δεν μπορεί να υποκατασταθεί από κάτι. Το "αγαθό" όμως ως ουσιαστικό, το πράγμα δηλαδή που είναι αγαθό, το όλον στο οποίο εφαρμόζεται το επίθετο "αγαθό", ορίζεται. Το όλον μπορεί να είναι ευχάριστο, έξυπνο κλπ. Αλλά, ενώ οι ιδιότητες αυτές ορίζουν "το αγαθό", δεν ορίζουν το "αγαθό" ως ποιότητα που αποδίδουμε σε κάποιο πράγμα. Υπενθυμίζουμε ότι το "αγαθό" επισημαίνει μια απλή και αόριστη ποιότητα. Η ηθική προσπαθεί να καθορίσει τις ιδιότητες που ανήκουν στα αγαθά πράγματα. Αλλά οι ιδιότητες αυτές δεν ορίζουν το "αγαθό". Η ταύτισή τους με το "αγαθό" αποτελεί την "φυσιοκρατική πλάνη". Πολλοί φιλόσοφοι υποστήριξαν ότι το "αγαθό" είναι η ηδονή ή αυτό που είναι αντικείμενο επιθυμίας. Αλλά αυτό είναι ψυχολογική και όχι ηθική κατηγορία. Όταν λένε "η ηδονή είναι αγαθό", δεν μπορεί να εννοούν πως η ηδονή είναι αναγκαίο επακόλουθο αν ηδονή σημαίνει αγαθό. Σκοπός της ηθικής θεωρεί ο Moore όχι την απλή συναγωγή αληθών συμπερασμάτων, αλλά τη συναγωγή τους με λογικά έγκυρη επιχειρηματολογία. Άμεσο αντικείμενο της ηθικής είναι γι' αυτόν η γνώση και όχι η πράξη. Όσον αφορά αυτό το μη επιδεκτικό ορισμού και μη αναλύσιμο νοητικό αντικείμενο, όλοι οι όροι της ηθικής κρίσης αναφέρονται σ' αυτό κατά δύο διαφορετικούς τρόπους: Βεβαιώνουν ότι είτε η μοναδική ιδιότητα συνδέεται πάντα με το εν λόγω πράγμα, ή ότι το εν λόγω πράγμα (το αγαθό) είναι αιτία για την ύπαρξη άλλων πραγμάτων με τα οποία συνδέεται η μοναδική αυτή ιδιότητα. Στη καθημερινή γλώσσα η διάκριση αυτή συμπίπτει με την αντιδιαστολή μεταξύ του αγαθού ως μέσου και του αγαθού καθεαυτού. Το αγαθό ως μέσον είναι δηλωτικό αιτιακών σχέσεων. Όμως δεν βρίσκουμε εύκολα αιτιακές κρίσεις καθολικά αληθινές. Είναι γενικά για μια χρονική περίοδο αληθινές, γιατί τα αποτελέσματα αλλάζουν με τις περιστάσεις. Ποτέ δεν είμαστε απόλυτα βέβαιοι για τα αποτελέσματα μιας πράξης, πρέπει επομένως να αρκούμαστε κατά τον "ισολογισμό" στο μεγαλύτερο περίσσειμα αγαθών αποτελεσμάτων ή στο μικρότερο υπόλοιπο κακών. Έτσι, λοιπόν, οι είναι δύσκολες οι ηθικές γενικεύσεις, και οι αιτιακές ηθικές κρίσεις είναι μόνο γενικά και όχι καθολικά αληθινές. Ούτε οι κρίσεις για το "τί πρέπει να πράττουμε", ούτε για το "τί σκοπούς πρέπει να προάγουν οι πράξεις μας" είναι καθарές κρίσεις αληθινής αξίας. Μια απόλυτα υποχρεωτική πράξη μπορεί να μην έχει αληθινή αξία. Στις σχετικές ερωτήσεις πρέπει να απαντήσουμε σε δύο διαφορετικά πράγματα: "τί βαθμό εσωτερικής αξίας έχουν διάφορα πράγματα"; και "πώς μπορούν αυτά να επιτευχθούν"; Η απάντηση στην ερώτηση "ποιά πράγματα έχουν αληθινή αξία" είναι πρωταρχική και αποτελεί προϋπόθεση της απάντησης στο "τί πρέπει να πράττουμε". Τα βασικά λοιπόν κατά Moore ερωτήματα είναι: τί εννοούμε με τον όρο "αγαθό", ποιά πράγματα είναι αγαθά καθαυτά, και ποιές αιτιακές σχέσεις ισχύουν μεταξύ του αγαθού καθεαυτού και άλλων πραγμάτων; Οι φυσιοκρατικές θεωρίες, θεωρούν την ηθική εμπειρική ή θετική επιστήμη, και ορίζουν το αγαθό σε αναφορά με ένα "φυσικό αντικείμενο" της εμπειρίας μας, θεωρούν το αγαθό ιδιότητα των πραγμάτων που υπάρχουν εν χρόνω. Οι μεταφυσικές θεωρίες το

ορίζουν σε αναφορά με κάτι που συνάγεται πως υπάρχει σε έναν υπεραισθητό αλλά εξίσου πραγματικό κόσμο (εκπροσώπους της μεταφυσικής ηθικής θεωρεί ο Moore κυρίως τους Στωικούς, τον Spinoza, τον Kant και τους νεοεγελιανούς). Στους φυσιοκράτες, πιστεύει ο Moore, γίνεται σύγχυση μεταξύ "φυσικού" (natural), φυσιολογικού (normal) και αναγκαίου (necessary). Άλλοτε το ιδανικό ταυτίζεται με το πίο εξελιγμένο (Spencer) ή με την ποιότητα και την ένταση της ζωής (Guyeau). Αντίθετα απ' ότι για τους Στωικούς, η φύση για τον Moore αλλά και τον Wittgenstein δεν έχει δεοντολογική σημασία. Η γνώση των ηθικών διακρίσεων είναι αυτή που ενδιαφέρει έστω και αν δεν κάνει τη ζωή του ανθρώπου καλύτερη. Πρόκειται για την ηθική ουδετερότητα της αναλυτικής ηθικής, χωρίς τη σωματική αισιοδοξία ότι "η αρετή είναι γνώση". Οι ηδονιστές (Αρίστιππος, Επίκουρος, αλλά και οι Bentham, Mill, Spencer, Sidgwick) κατά τον Moore θεωρούν όλα τα πράγματα (γνώση, αρετή, ζωή) ως μέσα για την απόλαυση της ηδονής. Ο Mill μεταπηδά αυθαίρετα από το είναι στο πρέπει να είναι, από το is στο ought, από το επιθυμητό (desired- η ηδονή) σε αυτό που πρέπει να είναι αντικείμενο επιθυμίας, και το ορίζει ως αγαθό. Η ηδονή όμως δεν είναι αντικείμενο επιθυμίας αλλά αιτία επιθυμίας. Η σχέση τους είναι αιτιακή. Κάθε κρίση της πρακτικής ηθικής ανάγεται στη μορφή: "αυτό είναι μια αιτία αυτού του αγαθού αποτελέσματος". Ισχυρισμοί όπως "αυτή η πράξη είναι ορθή" ή "αυτό είναι το καθήκον μου" είναι ισοδύναμοι με τον ισχυρισμό: "τα συνολικά αποτελέσματα της εν λόγω πράξης θα είναι τα καλύτερα δυνατά". Όλοι οι ηθικοί νόμοι είναι απλοί ισχυρισμοί ότι μερικά είδη πράξεων θα έχουν αγαθά αποτελέσματα. 30 Και όμως, στην ηθική υποτίθεται ότι το ορθό και το συμφέρον συγκρούονται. Αλλά "ορθό" σημαίνει "αίτιο αγαθών αποτελεσμάτων" και είναι το ίδιο με το "χρήσιμο". Συνεπώς καμία πράξη, που δεν δικαιώνεται από τα αποτελέσματά της, δεν είναι ορθή από ηθική άποψη. Δηλαδή, ισχυρίζεται ο Moore ότι μια πρόταση για το καθήκον δεν είναι αντικείμενο ενόρασης, δεν είναι αυτονόητη. Ως ορθή ορίζεται μια εκούσια πράξη, μόνο όταν καμία άλλη δυνατή για τον πράττοντα κάτω από τις ίδιες συνθήκες δεν θα προξενούσε μεγαλύτερη ευχαρίστηση. Ενώ οι ορθές πράξεις προξενούν ένα μέγιστο ηδονής, το χρέος ή το καθήκον προξενούν το μέγιστο της ηδονής. Απόλυτα καθολικό κριτήριο ορθότητας είναι η ηδονή (οι κακές πράξεις αντίθετα, προξενούν πόνο). Αλλά δεν είναι ορθές μόνο επειδή έχουν αυτό το χαρακτηριστικό. Πόνος και ηδονή δεν είναι τα μόνα αποτελέσματα των πράξεων. Οι συνολικές επιπτώσεις πρέπει να υπολογίζονται περισσότερο από τις άμεσες, και οι πνευματικές και αισθητικές ηδονές περισσότερο από τις αισθησιακές. Παράλληλα, η ηθική δεν μπορεί να δώσει μια "λίστα" καθηκόντων, αφού είναι αδύνατον στην στενή έννοια να ανακαλύψουμε ποιο είναι το καθήκον μας. Το καθήκον ο Moore δεν το διακρίνει ηθικά από τη σκοπιμότητα (expediency). Μη ηθική επίσης θεωρεί και τη διάκριση καθήκοντος-συμφέροντος (interest) και δεν βλέπει τη σύγκρουσή τους αναγκαία. Οι αρετές, επίσης, δεν θεωρούνται διαθέσεις αγαθές καθεαυτές ούτε αναγκαία ηθικές πράξεις καθεαυτές, αλλά αγαθές ως μέσα, διαθέσεις για την εκτέλεση αγαθών πράξεων. Βασική μέθοδο για την εκτίμηση κάποιων πραγμάτων θεωρεί ο Moore την απομόνωσή τους, την θεώρησή τους δηλαδή αν υπήρχαν εντελώς μόνα τους. Έτσι ότων τα απομονώσουμε, πίο πολύτιμα πράγματα θα μας φανούν ορισμένες συνειδησιακές καταστάσεις, όπως η απόλαυση των ωραίων πραγμάτων και η ευχάριστη ανθρώπινη επικοινωνία: οι αισθητικές, δηλαδή, απολαύσεις και οι προσωπικές σχέσεις. Η συνείδηση

των ωραίων πραγμάτων: έσχατη και θεμελιώδης αλήθεια της ηθικής. Για χατήρι αυτών των πραγμάτων και μόνο πιστεύει ότι αξίζει να ζει κανείς την δημόσια και ιδιωτική του ζωή. Αυτά, ως σύνθετα σύνολα, αποτελούν το λόγο ύπαρξης της αρετής, τον έσχατο σκοπό της ανθρώπινης πράξης, το μόνο κριτήριο της κοινωνικής προόδου. Για την σωστή βέβαια αισθητική εκτίμηση χρειάζεται γνώση των ωραίων ποιοτήτων και κατάλληλη αισθητική συγκίνηση. Η κοινή γνώμη θεωρείται αποφασιστική για τον καθορισμό των ωραίων πραγμάτων. Ως ωραίο ορίζεται "εκείνο που η γεμάτη θαυμασμό θεώρησή του είναι αγαθό καθεαυτό". Δεν είναι απλώς θέμα γούστου ή αισθήματος. Δεν είναι αγαθό καθεαυτό αλλά αναγκαίο στοιχείο του αγαθού και ορίζεται σε αναφορά με το αγαθό. Το σύνολο με το οποίο συνδέεται είναι αγαθό.

2.2 Cybernetics - Η Κυβερνηθική

Η κυβερνηθική ορίζεται σύμφωνα με την Stanford Encyclopedia of Philosophy ως: *Οι μελέτες ηθικών, νομικών και κοινωνικών ζητημάτων που αφορούν την κυβερνοτεχνολογία, η επίδραση της τεχνολογίας στον κυβερνοχώρο στα κοινωνικά, νομικά και ηθικά μας συστήματα, αξιολογώντας κοινωνικές πολιτικές και νόμους που έχουν διαμορφωθεί ως απάντηση σε ζητήματα παράγεται από την ανάπτυξη και τη χρήση του.*

Αυτό ορίζει την κυβερνηθική (cyberethics) όπου πέρα από τη γενικευμένη μελέτη της εφαρμοσμένης ηθικής στην πραγματικότητα, είναι ένας διαφορετικός τομέας μελέτης. Το ζήτημα του πώς θα πρέπει να αναπτυχθεί και να χρησιμοποιηθεί η τεχνολογία γίνεται όλο και πιο αναγκαίο και απαραίτητο για μελέτη.

Η επιτυχής αντιμετώπιση των περισσότερων ηθικών προκλήσεων απαιτεί προσεκτική ή τη βασική κατευθυντήρια γραμμή που θα πρέπει να είναι “καλός”. Τόσο η σωστή διαδικασία όσο και η “καλή συμπεριφορά” είναι κατευθυντήριες γραμμές και ηθικές αξίες που ένας εργαζόμενος ή ένα άτομο καταρτίστηκε κατά τη διάρκεια της παιδικής τους ηλικίας και των εκπαιδευτικών ετών. Φυσικά, η επανεκπαίδευση υπαλλήλων ή ατόμων σχετικά με την ηθική και τα ηθικά είναι μια επιλογή.

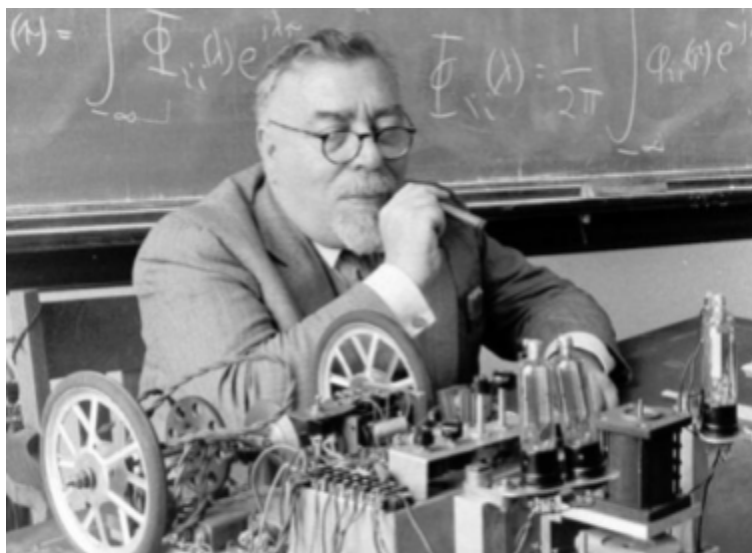
Η επιστήμη των υπολογιστών υπάρχει εδώ και 50 χρόνια και παίζει όλο και μεγαλύτερο ρόλο στη ζωή μας και προσελκύει περισσότερη προσοχή. Οι τεχνολογίες θα συνεχίσουν να επιταχύνουν την παραγωγικότητα και να διευκολύνουν την έρευνα προς έναν μεγαλύτερο στόχο. Ωστόσο, με κάθε νέα τεχνολογία, η ανησυχία γίνεται μεγαλύτερη για την ασφάλεια και τον αντίκτυπο που θα έχουν αυτές οι τεχνολογίες στα αντίστοιχα περιβάλλοντά τους.

2.2.1 Ο Norbert Wiener - The human use of human beings

Ο Norbert Wiener ήταν Αμερικάνος μαθηματικός και φιλόσοφος, και καθηγητής Μαθηματικών στο Massachusetts Institute of Technology (MIT). Ο Wiener αργότερα έγινε πρώτος ερευνητής στοχαστικών και μαθηματικών διαδικασιών θορύβου. Θεωρείται ο δημιουργός της κυβερνηθικής, μια τυποποίηση της έννοιας της ανατροφοδότησης, με επιπτώσεις στη μηχανική, τον έλεγχο συστημάτων, την πληροφορική, τη βιολογία, τη νευροεπιστήμη, τη φιλοσοφία και την οργάνωση της κοινωνίας. Ο Norbert Wiener είναι ένας από τους πρώτους που θεωρούν ότι όλη η έξυπνη συμπεριφορά ήταν αποτέλεσμα μηχανισμών ανατροφοδότησης, που θα μπορούσαν να προσομοιωθούν από μηχανές και ήταν ένα σημαντικό πρώιμο βήμα προς την ανάπτυξη της σύγχρονης Τεχνητής Νοημοσύνης.

Το βιβλίο του Wiener “The human use of human beings”, είναι ένα πρωτότυπο βιβλίο, ειδικά κατά τη διάρκεια της περιόδου που κυκλοφόρησε, όπου χαρακτηρίζεται από τις έννοιες της “πληροφορίας” και της “επικοινωνίας”. Επινόησε τον όρο “cybernetics” για να χαρακτηρίσει μια πολύ γενική επιστήμη του “control and communication in the animal and machine”.

Συγκέντρωσε έννοιες από τη μηχανική, τη μελέτη του νευρικού συστήματος και τη μηχανική - όπως την εντροπία που θα δούμε αμέσως παρακάτω. Μέσα από αυτά τα παραδείγματα, ο συγγραφέας ανέπτυξε έννοιες που έχουν διαπεραστεί μέσω της επιστήμης, ειδικά της βιολογίας και της πληροφορικής και της κοινής ομιλίας: “πληροφορία”, “μήνυμα”, “ανατροφοδότηση” και “έλεγχος”.



Ο Norbert Wiener στην τάξη διδασκαλίας του MIT

“Εάν ο δέκατος έβδομος και ο πρώτος δέκατος όγδοος αιώνας είναι η εποχή των ρολογιών, και ο επόμενος δέκατος όγδοος και δέκατος ένατος αιώνας αποτελούν την εποχή των ατμομηχανών, η παρούσα στιγμή είναι η εποχή της επικοινωνίας και του ελέγχου”.

Μετά την συμμετοχή του Wiener στον Δεύτερο Παγκόσμιο Πόλεμο, σε ερευνητικά προγράμματα του στρατού, αρνήθηκε στη συνέχεια της ζωής του να αναλάβει οποιαδήποτε άλλη εργασία που σχετίζεται με τον στρατό. Προχώρησε στην οικοδόμηση μιας πρακτικής στάσης απέναντι στην τεχνολογία που βασίζεται στις θεμελιώδεις φιλοσοφικές προοπτικές. Για αυτόν, οι τεχνολογίες δεν θεωρούνταν τόσο η εφαρμοσμένη επιστήμη, αλλά η εφαρμοσμένη κοινωνική και ηθική φιλοσοφία. Η πρόβλεψή του αφορούσε μια επικείμενη “επανάσταση στην επικοινωνία” στην οποία το “μήνυμα” θα ήταν μια κεντρική έννοια και οι σχετικές τεχνολογικές εξελίξεις θα ήταν στον τομέα της επικοινωνίας, του υπολογισμού και της οργάνωσης.

Η λέξη cybernetics αναφέρεται στη θεωρία της μετάδοσης μηνυμάτων μεταξύ ανθρώπων και μηχανών. Η κεντρική του θέση στο βιβλίο του είναι ότι:

“Η κοινωνία μπορεί να γίνει κατανοητή μόνο μέσω μιας μελέτης των μηνυμάτων και των τρόπων επικοινωνίας που ανήκουν σε αυτή και ότι στο μέλλον η ανάπτυξη αυτών των μηνυμάτων και τρόπων επικοινωνίας μεταξύ ανθρώπου και μηχανών, μεταξύ μηχανών και

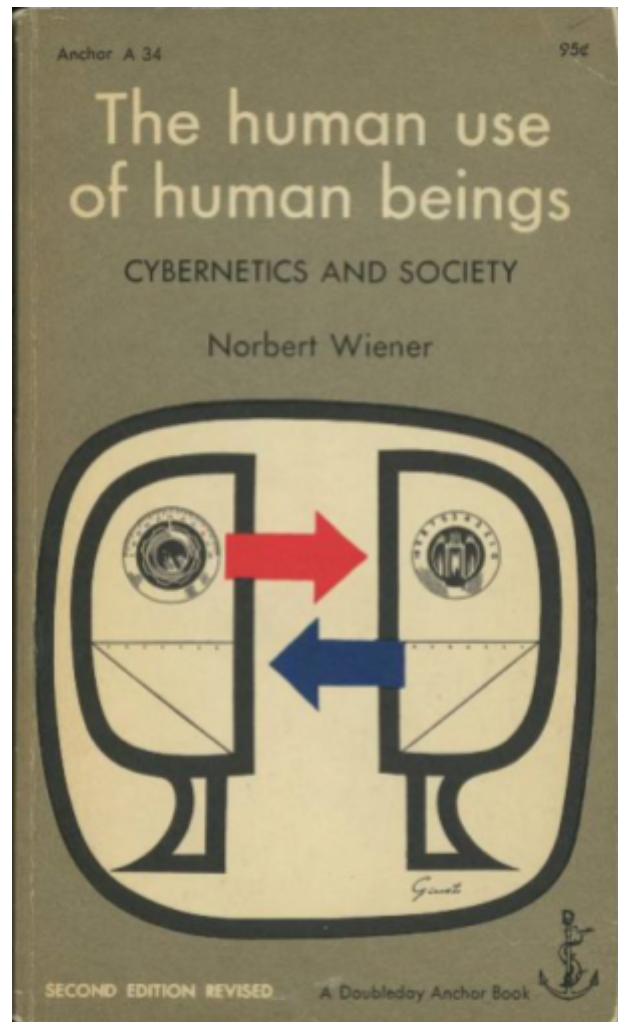
ανθρώπου, και μεταξύ μηχανής και μηχανής, προορίζονται να διαδραματίσουν έναν ολοένα αυξανόμενο ρόλο”.

Οι μέθοδοι επικοινωνίας έχουν εισέλθει σε ένα νέο πεδίο που περιλαμβάνει νέες τεχνολογίες. Είτε μια μετάδοση γίνεται μεταξύ ατόμων, είτε μεταξύ ατόμων και μηχανημάτων, η διαδικασία είναι παρόμοια με τις πληροφορίες που αποστέλλονται από ένα μέρος και λαμβάνονται από ένα άλλο, το οποίο μπορεί να στείλει απάντηση. Οι άνθρωποι, τα ζώα και τα φυτά έχουν όλοι τη δυνατότητα να αναλαμβάνουν συγκεκριμένες ενέργειες ως απάντηση στο περιβάλλον τους. Με τον ίδιο τρόπο, οι μηχανές διαθέτουν συστήματα ανατροφοδότησης ώστε οι επιδόσεις τους να τροποποιηθούν ή να αξιολογηθούν μετά από αποτελέσματα. Στο πλαίσιο της κοινωνίας ανθρώπων / μηχανών, η Wiener ορίζει το μήνυμα ως:

“μια ακολουθία γεγονότων στο χρόνο που, αν και από μόνη της έχει κάποια έκτακτη ανάγκη, προσπαθεί να συγκρατήσει την τάση της φύσης προς τη διαταραχή προσαρμόζοντας τα μέρη της σε διάφορα σκοπιμότερα άκρα”.

Όταν επικοινωνούμε με ένα άλλο άτομο, του μεταδίδουμε ένα μήνυμα και όταν επικοινωνεί μαζί μας, επιστρέφει ένα σχετικό μήνυμα το οποίο περιέχει πληροφορίες που είναι κυρίως προσιτές σε αυτόν και όχι σε εμάς. Το μήνυμα μέχρι να διαμορφωθεί και να επιστραφεί περνάει από κάποια ενδιάμεσα βήματα μέσα από κάποιον ανθρώπινο νου, κάποια μηχανή.

Τα μηνύματα αυτά είναι εντολές μέσω των οποίων ασκούμε έλεγχο στο περιβάλλον μας και είναι ένα είδος πληροφοριών που του προσδίδουμε. Όπως κάθε μορφή πληροφοριών, αυτές οι εντολές υπόκεινται σε “αποδιοργάνωση” κατά τη μεταφορά. Γενικά έρχονται λιγότερο συνεκτικά και σίγουρα όχι τόσο συνεκτικά από ό,τι στάλθηκαν. Στον έλεγχο και την επικοινωνία, παλεύουμε πάντα την τάση της φύσης να υποβαθμίζει το οργανωμένο και να καταστρέφει το νόημα, την τάση δηλαδή για αύξηση της εντροπίας. Οι πληροφορίες είναι ένα όνομα για το περιεχόμενο αυτού που ανταλλάσσεται με τον εξωτερικό κόσμο καθώς

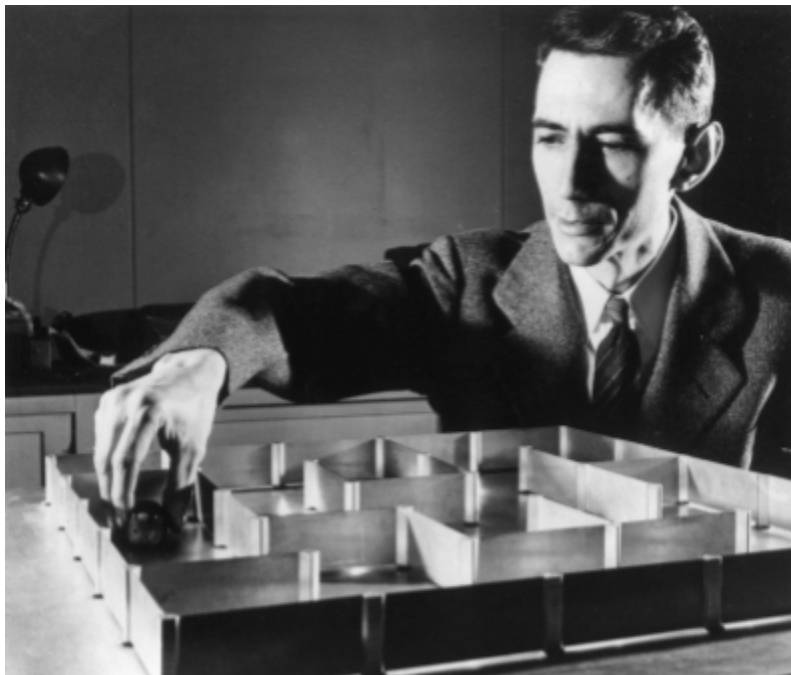


προσαρμόζουμε σε αυτόν και κάνουμε την προσαρμογή μας αισθητή. Οι ανάγκες και η πολυπλοκότητα της σύγχρονης ζωής κάνουν υψηλότερες απαιτήσεις σε αυτήν τη διαδικασία πληροφόρησης από ποτέ. Το να ζεις αποτελεσματικά είναι να ζεις με επαρκείς πληροφορίες.

Τα μηνύματα είναι από μόνα τους μια μορφή μοτίβου και οργάνωσης. Πράγματι, - είναι δυνατόν να αντιμετωπίσουμε τα σύνολα μηνυμάτων ως εντροπία σαν σύνολα του εξωτερικού κόσμου. Ακριβώς όπως η εντροπία είναι ένα μέτρο αποδιοργάνωσης, οι πληροφορίες που μεταφέρονται από ένα σύνολο μηνυμάτων είναι ένα μέτρο οργάνωσης. Είναι δυνατόν να ερμηνεύσουμε τις πληροφορίες που μεταφέρει ένα μήνυμα ως κυρίως την αρνητική της εντροπίας του και τον αρνητικό λογάριθμο της πιθανότητάς του. Δηλαδή, όσο πιο πιθανό είναι το μήνυμα, τόσο λιγότερες πληροφορίες δίνει.

2.2.2 Θεωρία Πληροφορίας & Στοιχεία από το Shannon

Ο Claude Shannon αποφοίτησε το 1936, από το Πανεπιστήμιο του Michigan, παίρνοντας δύο πτυχία του Μαθηματικού και του Ηλεκτρολόγου Μηχανικού. Τις μεταπτυχιακές του σπουδές τις έκανε στο MIT, όπου είχε σαν καθηγητή και τον Norbert Wiener. Ο πρόεδρος



Ο Claude Shannon στις 10 Μαΐου του 1952, με ένα ηλεκτρονικό ποντίκι με μνήμη, το οποίο μπορεί να μάθει την λύση ενός λαβύρινθου χωρίς λάθος, μετά από μία μόνο εκπαίδευση

του τμήματος Μηχανολογίας την εποχή εκείνη, όρισε τον Shannon υπεύθυνο για τη λειτουργία μιας δύσχρηστης υπολογιστικής συσκευής, που είχε κατασκευάσει ο ίδιος και είχε ονομάσει “διαφορικό αναλυτή”. Άρχισε να σκέφτεται τρόπους βελτίωσής του, ενδεχομένως με τη χρήση ηλεκτρικών κυκλωμάτων στη θέση των δυσχρηστων μηχανικών μερών όπου και τελικά συνδίασε την άλγεβρα του Boole που είχε διδαχθεί ως προπτυχιακός φοιτητής, αφού είχε πολλά κοινά στοιχεία με ένα ηλεκτρικό κύκλωμα. Σε μια δημοσίευση, με τίτλο “A Symbolic Analysis of Relay and Switching Circuits”, ο Shannon περιέγραψε με ποιο τρόπο η λογική του Boole,

σύμφωνα με την οποία όλα τα προβλήματα μπορούν να λυθούν με τη χρήση μόλις δύο συμβόλων, του 1 και του 0, μπορούσε να εφαρμοστεί στα ηλεκτρικά διακοπτόμενα κυκλώματα, όπου είναι κλειστός ή ανοιχτός ο διακόπτης αντίστοιχα, καθώς και με αντίστοιχες συνδεσμολογίες οι διακόπτες αυτοί θα μπορούσαν να εκτελούν και πιο πολύπλοκες πράξεις, “and”, “or”, “not”, σύμφωνα με την άλγεβρα Boole.

Η παραπάνω διατριβή του Shannon, ο οποίος οραματίστηκε όλες τις μορφές επικοινωνίας σε δυαδικό κώδικα και υποστήριξε την άποψη ότι τα δυαδικά ψηφία μπορούν να συμβολίσουν ακόμα και λέξεις, ήχους, εικόνες, ίσως και ιδέες, χαρακτηρίστηκε μία από τις σημαντικότερες του 20ού αιώνα, ξεπερνώντας σε αξία ακόμα και τη προηγούμενη διατριβή του ίδιου για το διδακτορικό του, με θέμα “Μια άλγεβρα για τη θεωρητική γενετική”. Σε αυτήν την διατριβή θεωρούσε πως η διπλή έλικα του DNA αποτελεί ένα πληροφοριακό σύστημα. Ο Shannon πίστευε ότι η πληροφορία δεν διέφερε από οποιοδήποτε άλλο μέγεθος και κατά συνέπεια ήταν δυνατός ο χειρισμός της από μηχανές. Εφαρμόζοντας τα αποτελέσματα των προηγούμενων ερευνών του στο πρόβλημα που είχε να αντιμετωπίσει, χρησιμοποίησε και πάλι τη λογική του Boole, καθώς και την εμπειρία του στην κρυπτο / αποκρυπτογράφηση κατά τη διάρκεια του πολέμου, για να αναπτύξει ένα μοντέλο που απλοποιούσε όσο το δυνατόν περισσότερο την πληροφορία.

Πρώτος ο Hartley όρισε έμμεσα την ποσότητα της πληροφορίας, το 1928 κι ένα χρόνο αργότερα ο Szilard συνέδεσε την πληροφορία και την Θερμοδυναμική Εντροπία. Το 1940 ο Shannon ξαναβρήκε τα ίδια αποτελέσματα με αφορμή τις τηλεπικοινωνίες και την εργασία του. Η καθαυτό όμως θεωρία της πληροφορίας αναπτύχθηκε από τους Wiener το 1948 και ακολούθως από τον Shannon. Το 1948 ο Shannon δημοσίευσε την αξιέπαιστη εργασία του, με τίτλο “Η μαθηματική θεωρία της πληροφορίας”. Ήταν ο πρώτος που έκανε την πρώτη ολοκληρωτική μαθηματική απόπειρα θεμελίωσης της Θεωρίας της Πληροφορίας. Στις σελίδες αυτής της εργασίας, την οποία συνυπογράφει ο μαθηματικός Warren Weaver, γίνεται λόγος για πρώτη φορά για μια μονάδα μέτρησης της πληροφορίας, το δυαδικό ψηφίο, το binary digit, που συντημήθηκε αργότερα από επιστήμονες του χώρου αρχικά σε binit και στη συνέχεια στο γνωστό μας bit.

Έχουμε συνηθίσει να σκεφτόμαστε την “πληροφορία” ή το “μήνυμα” ως γεγονότα, δεδομένα, μαρτυρίες. Σύμφωνα όμως με τη θεωρία της πληροφορίας, πληροφορία είναι αυτό που δεν γνωρίζει κάποιος. Αν κάποιος ακούσει πως: “*Αύριο θα βρέχει στο κέντρο της Αθήνας*”, το μήνυμα αυτό, έχει μεγάλη πληροφορία, γιατί είναι ένα αβέβαιο γεγονός. Αν όμως ακούσει κάποιος πως: “*στην Ευρώπη αύριο θα βρέχει*”, τότε το κείμενο αυτό έχει πολύ μικρή πληροφορία. Γιατί στο μήνυμα αυτό η πιθανότητα να βρέχει κάπου στην Ευρώπη είναι πολύ μεγάλη, ίσως αγγίζει και το 100%.

Επομένως η πληροφορία συνδέεται με την αβεβαιότητα. Όσο μικρότερη είναι η πιθανότητα P να γίνει ένα γεγονός, τόσο περισσότερη ποσότητα πληροφορίας I συνοδεύει την πραγματοποίησή του. Και αντίστροφα, αν η πιθανότητα P πραγματοποίησης ενός γεγονότος

είναι μεγάλη, τότε η πληροφορία I που “κουβαλάει” το γεγονός αυτό είναι μικρή. Αν λοιπόν ο παραλήπτης έχει ήδη την πληροφορία, δεν μπορεί να πει κανείς ότι έλαβε χώρα μετάδοση μηνύματος.

Η πληροφορία I συνδέεται με την πιθανότητα P με την σχέση: $I = -\log_2 P$

Σύμφωνα με τη θεωρία του Shannon, περισσότερα bits πληροφορίας παίρνει κανείς από ένα μήνυμα, αν είναι μεγαλύτερη και η αβεβαιότητα που κουβαλάει το μήνυμα, γιατί αυτή η αβεβαιότητα είναι αντιστρόφως ανάλογη με την πιθανότητα P να συμβεί όπως αναφέραμε.

Ένα από τα σημαντικότερα στοιχεία όμως του έργου του Shannon είναι ότι παρέχει στους μηχανικούς τα μαθηματικά εργαλεία που απαιτούνται για τη μέτρηση της απόδοσης ενός καναλιού επικοινωνίας, πόση πληροφορία δηλαδή μπορεί να ξεκινήσει από το σημείο A και να φτάσει στο σημείο B χωρίς σφάλματα. Η επιθυμητή πληροφορία είναι το “σήμα”. Η ανεπιθύμητη είναι τα “παράσιτα” ή ο “θόρυβος”. Ο Shannon είδε πως όσο λιγότερο θόρυβος έχει ένα σύστημα τόσο περισσότερη πληροφορία μεταδίδει. Και αντιστρόφως, όσο αυξάνεται η αταξία (θόρυβος) ενός συστήματος τόσο λιγότερη πληροφορία μεταδίδει. Θα μπορούσαμε να πούμε πως η πληροφορία του συστήματος αποτελεί μέτρο της εσωτερικής του τάξης (δηλ. αντιστρόφως ανάλογη με την αταξία, αλλά η εντροπία είναι το μέτρο της αταξίας ενός συστήματος, άρα η πληροφορία είναι αντιστρόφως ανάλογη της εντροπίας). Ο αριθμός των πιθανών μηνυμάτων που μπορεί κανείς να δημιουργήσει με S αριθμό bits είναι 2^S στη δύναμη του S , δεδομένου ότι έχουμε δύο bits, το 1 και το 0. Αντιστρέφοντας την ιδέα, ο αριθμός των bits που χρειάζεται κανείς για να μεταδώσει ένα μήνυμα είναι ο λογάριθμος με βάση το 2 του αριθμού των πιθανών μηνυμάτων: $S = \log_2 P$

Όπου στο S ο Shannon έδωσε το όνομα “εντροπία”. Η εντροπία αναφέρεται σε μια κατάσταση ενός φυσικού συστήματος, ενώ συγχρόνως αποτελεί μέτρο της αταξίας του συγκεκριμένου συστήματος. Η αταξία δεν είναι τελείως αντικειμενική ιδιότητα. Ο ανθρώπινος παράγοντας δεν μπορεί να την αποκλείσει τελείως, γιατί η ιδέα της τάξης είναι άρρηκτα συνδεδεμένη με τη λειτουργία του νου.

Η θεωρία πληροφορίας είναι τμήμα των εφαρμοσμένων μαθηματικών και ασχολείται με την ποσοτικοποίηση της πληροφορίας. Η θεωρία της πληροφορίας αναπτύχθηκε από τον Claude Shannon κατά την δημιουργική δουλειά “A Mathematical Theory of Communication”, για να βρει τα θεμελιώδη όρια της επεξεργασίας σήματος σε εφαρμογές όπως η συμπίεση δεδομένων και η αξιόπιστη αποθήκευση και μεταφορά δεδομένων. Από τη θεμελίωση της έχει διευρυνθεί ώστε να βρει εφαρμογές σε πολλούς άλλους τομείς, όπως στη στατιστική, στην επεξεργασία φυσικής γλώσσας, στην κρυπτογραφία, στα δίκτυα εκτός των δικτύων επικοινωνίας - όπως και στη νευροβιολογία, στην εξέλιξη, στους κβαντικούς υπολογιστές, στην ανίχνευση λογοκλοπής και σε άλλες μορφές ανάλυσης δεδομένων.

Τα πιο θεμελιώδη αποτελέσματα από την μελέτη του Shannon είναι το θεώρημα του Shannon για τον πηγαίο κώδικα, που καθιερώνει ότι, κατά μέσο όρο, ο αριθμός των δυαδικών ψηφίων (bits) που χρειάζονται για να αναπαρασταθεί το αποτέλεσμα ενός αβέβαιου γεγονότος δίνεται από την εντροπία πληροφοριών του, και από το θεώρημα του Shannon για το θορυβώδες κανάλι μεταφοράς, το οποίο δηλώνει ότι η αξιόπιστη επικοινωνία είναι δυνατή σε θορυβώδη κανάλια με την προϋπόθεση ότι ο συντελεστής της επικοινωνίας είναι κάτω από ένα ορισμένο όριο, το οποίο ονομάζεται χωρητικότητα καναλιού. Στην πράξη μπορούμε να προσεγγίσουμε την χωρητικότητα καναλιού χρησιμοποιώντας κατάλληλα συστήματα κωδικοποίησης και αποκωδικοποίησης.

2.2.3 Εξέλιξη και Εντροπία

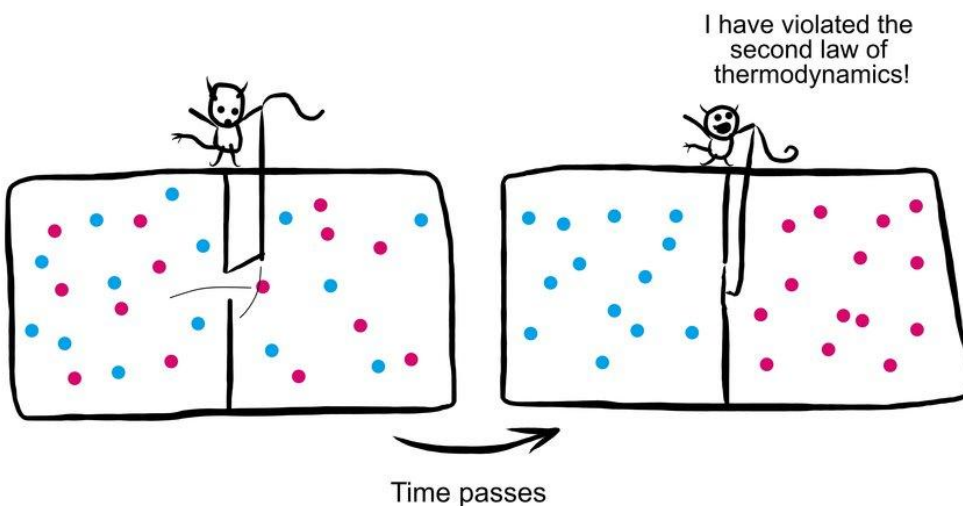
Η στατιστική τάση για διαταραχή της φύσης, η τάση αύξησης της εντροπίας σε απομονωμένα συστήματα εκφράζεται από τον δεύτερο νόμο της θερμοδυναμικής. Εμείς, ως ανθρώπινα όντα, δεν είμαστε απομονωμένα συστήματα. Καταναλώνουμε τρόφιμα, τα οποία παράγουν ενέργεια και ως εκ τούτου, είναι μέρη αυτού του μεγαλύτερου κόσμου που περιέχουν αυτές τις πηγές ζωτικότητας μας. Αλλά ακόμη πιο σημαντικό είναι το γεγονός ότι λαμβάνουμε πληροφορίες μέσω των αισθήσεών μας και ενεργούμε βάσει των πληροφοριών που λαμβάνονται.

Μια λαμπρή έκφραση του ρόλου των πληροφοριών με την βοήθεια της αναλυτικής φιλοσοφίας παρέχεται από τον Clerk Maxwell, με τη μορφή του λεγόμενου “δαίμονα του Maxwell”. Ο δαίμονας του Μάξγουελ είναι ένα νοητό πείραμα, που επινόησε ο Τζέιμς Κλερκ Μάξγουελ και συνδέει τις έννοιες της πληροφορίας και της εντροπίας με την ενέργεια. Το υποθετικό αυτό πείραμα σχεδιάστηκε με σκοπό την καλύτερη κατανόηση και ενδεχομένως κατάρριψη του δεύτερου θερμοδυναμικού νόμου, ο οποίος απαγορεύει την παραγωγή ενέργειας εκ του μηδενός σε ένα κλειστό σύστημα.

Ας υποθέσουμε ότι έχουμε ένα δοχείο αερίου, του οποίου η θερμοκρασία είναι η ίδια παντού. Μερικά μόρια αυτού του αερίου θα κινούνται γρηγορότερα από άλλα. Στη διάταξη του πειράματος ένα ον με ικανότητα να γνωρίζει κάθε στιγμή την ταχύτητα και τη θέση του κάθε μορίου ενός αερίου σε ένα δοχείο, που χωρίζεται σε δύο μέρη, τραβά και ανοίγει μια πόρτα κατά βούληση και αφήνει να περνούν κατά προτίμηση προς τη μία μεριά τα “ψυχρά” (χαμηλής ταχύτητας) και προς την άλλη τα “θερμά” (υψηλής ταχύτητας) μόρια. Η πόρτα αφήνεται να κλείσει αμέσως μετά την επιλεκτική διέλευση του κάθε μορίου και το ελατήριο επιστρέφει πίσω στο δαίμονα την ενέργεια που δαπάνησε για να την ανοίξει. Στο τέλος, χωρίς να έχει δοθεί ενέργεια στο σύστημα, το οποίο θεωρείται απομονωμένο από το περιβάλλον, εμφανίζεται η μία μεριά του κουτιού με θερμό αέριο και η άλλη με ψυχρό και, πιθανόν, και διαφορετική πίεση στο ένα δοχείο από το άλλο. Αυτό ισοδυναμεί με αύξηση της

ενέργειας του συστήματος, αφού μπορούμε να χρησιμοποιήσουμε τις δύο δεξαμενές θερμότητας, αριστερή και δεξιά, που βρίσκονται σε διαφορετική θερμοκρασία, με μια θερμική μηχανή για να παράγουμε έργο.

Εδώ εμφανίζεται μια πολύ ενδιαφέρουσα διάκριση και εξέλιξη της κλασικής φυσικής. Στην κλασική φυσική του Maxwell δεν υπάρχει τίποτα που να εμποδίζει έναν από τους δαίμονες του να δίνει τη δική του πηγή ισχύος. Η σύγχρονη φυσική, ωστόσο, αναγνωρίζει ότι ο δαίμονας μπορεί να κερδίσει μόνο τις πληροφορίες με τις οποίες ανοίγει ή κλείνει την πόρτα από κάτι σαν ένα λογικό όργανο που για τους σκοπούς αυτούς είναι ένα μάτι. Το φως που χτυπά το μάτι του δαίμονα δεν είναι ένα ενεργειακό συμπλήρωμα μηχανικής κίνησης, αλλά μοιράζεται τις κύριες ιδιότητες της μηχανικής κίνησης.



Έναν αιώνα μετά την εμφάνιση του δαίμονα του Maxwell, ο αμερικανός φυσικός Charles Bennett, βασισμένος στη δουλειά των Leo Szilard και Rolf Landauer, έλυσε το παράδοξο συνδέοντας τη θερμοδυναμική με τη νέα επιστήμη της πληροφορίας. Ο Bennett ισχυρίστηκε ότι η γνώση του δαίμονα είναι αποθηκευμένη στη μνήμη του και η μνήμη πρέπει να καθαριστεί, κάτι για το οποίο απαιτείται έργο. Με άλλα λόγια, καθώς ο δαίμονας οργανώνει το αέριο σε θερμό και κρύο μειώνοντας την εντροπία του, ο εγκέφαλός του ξοδεύει ενέργεια και το ισοζύγιο της εντροπίας είναι πάλι θετικό. Η συνολική εντροπία του συστήματος αερίου-δαίμονα αυξάνεται, ικανοποιώντας τον δεύτερο νόμο της θερμοδυναμικής. Τα συμπεράσματα αυτά αποκάλυψαν, όπως το διατύπωσε ο Landauer, ότι η “πληροφορία είναι φυσικό χαρακτηριστικό”.

2.3 Ηθική και Τεχνητή Νοημοσύνη

Οι ηθικές επιπτώσεις της Τεχνητής Νοημοσύνης (AI) έχουν παρουσιαστεί στην επιστημονική φαντασία για δεκαετίες. Οι συγγραφείς μυθιστορημάτων φαντασίας επέτρεψαν να θέσουν πολλά από τα ηθικά ερωτήματα που αντιμετωπίζουμε στον σημερινό κόσμο. Κυρίως σε ποιο βαθμό και πώς θα πρέπει να συμμετέχουν η τεχνητή νοημοσύνη και τα τεχνητά μηχανήματα στην κοινωνία μας και πώς πρέπει να τα αντιμετωπίζουμε ως αντάλλαγμα. Τα διεπιστημονικά πεδία της Μηχανικής Ηθικής και της Ρομποτικής, προσπαθούν τώρα να αντιμετωπίσουν αυτά τα ερωτήματα από επιστημονική άποψη.

Αν και αυτά τα δύο πεδία από μόνα τους είναι αρκετά νέα, το επιχείρημα στην επιστημονική κοινότητα είναι πολύ παλαιότερο. Ο Joseph Weizenbaum, δημιουργός του προγράμματος ELIZA [116] ήταν υποστηρικτής της περιορισμένης χρήσης της AI. Ισχυρίστηκε ότι ορισμένα καθήκοντα, όπως η ιατρική ή η νομική, δεν πρέπει ποτέ να γίνονται από την AI. Υποστήριξε ότι αυτά τα επαγγέλματα απαιτούν συμπόνια και διαίσθηση, ποιότητες που πίστευε ότι δεν μπορούν να επιτευχθούν από μηχανές [118]. Ο Weizenbaum επικρίνει την τεχνητή νοημοσύνη ως παράδειγμα ενός επιχειρήματος από την ομάδα “Anti AI”, καθώς και αντεπιχειρήματα στην άποψη του Weizenbaum.

2.3.1 ELIZA

Ο Joseph Weizenbaum ανέπτυξε ένα πρόγραμμα υπολογιστή που εξασφάλισε τη θέση του στα χρονικά του ιστορικού υπολογιστών - αυτό το πρόγραμμα ήταν η ELIZA. Το πρόγραμμα επέτρεψε σε ένα άτομο να επικοινωνήσει με έναν υπολογιστή χρησιμοποιώντας φυσική γλώσσα. Αυτό είχε ως αποτέλεσμα αποκρίσεις από τον υπολογιστή που θα μπορούσαν να δώσουν την εντύπωση ότι ο υπολογιστής κατανοούσε τι του έλεγαν, και ότι ο υπολογιστής μιλούσε. Ένα άτομο που “συνομιλεί” με το ELIZA θα πληκτρολογούσε ένα μήνυμα σε φυσική γλώσσα σε μια γραφομηχανή συνδεδεμένη σε έναν υπολογιστή που εκτελεί το πρόγραμμα. Αφού πληκτρολογήσει το μήνυμά, ο υπολογιστής θα δημιουργούσε μια απάντηση που θα εμφανίζεται στο ίδιο μηχάνημα.

Ο Weizenbaum αναγνωρίζει ότι οι υπολογιστές μπορούν να φανούν λες και κάνουν μαγεία. Ωστόσο, *“όταν ένα συγκεκριμένο πρόγραμμα αποκαλυφθεί, όταν η εσωτερική του λειτουργία εξηγηθεί σε γλώσσα αρκετά απλή για να προκαλέσει κατανόηση, η μαγεία του καταρρέει. Αποκαλύπτεται ως απλή συλλογή διαδικασιών, καθεμιά αρκετά κατανοητή”* [20]. Στο άρθρο που ακολούθησε αυτήν την αναφορά, ο Weizenbaum εξήγησε με σαφήνεια ακριβώς τον τρόπο με τον οποίο λειτουργούσε η ELIZA - πήγε βήμα προς βήμα για να δείξει ότι δεν ήταν αποτέλεσμα μαγείας ή γνήσιας κατανόησης εκ μέρους του προγράμματος, αλλά αποδόθηκε στον έξυπνο προγραμματισμό. Ακόμη και αν οι άνθρωποι γνώριζαν ότι απλώς ανταλλάσσουν μηνύματα με τον υπολογιστή, ακόμη και αν γνώριζαν το σενάριο και τον προγραμματισμό που είχαν ως αποτέλεσμα τις συγκεκριμένες απαντήσεις, η μαγεία του ELIZA έδειξε *“πόσο εύκολο είναι να δημιουργήσετε και να διατηρήσετε την ψευδαίσθηση της κατανόησης”* [118].

Πρωταρχικός στόχος του προγράμματος ELIZA ήταν, μόλις ξεκίνησε η “συνομιλία”, να συνεχίσει τη συζήτηση με οποιοδήποτε τρόπο. Το πέτυχε τόσο αποκρύπτοντας “οποιαδήποτε παρανόηση” όσο και βασιζόμενοι στην καλή πίστη του ανθρώπινου συνομιλητή, ώστε να δυσκολευτεί να αποδεχτεί ακόμα και στην περίπτωση που του παρουσίαζαν αποδείξεις ότι το πρόγραμμα δεν καταλάβαινε πραγματικά τα μηνύματα που πληκτρολογούσε [118]. Οι επικοινωνίες μεταξύ δύο ανθρώπων βασίζονται, κατά την εκτίμηση του Weizenbaum, σε μια παραδοχή εκ μέρους του καθενός ότι ο άλλος άνθρωπος κατανοεί τι του έλεγαν. Μια τέτοια κατανόηση θα μπορούσε να είναι δύσκολη καθώς τα ποικίλα πλαίσια από τα οποία ήρθαν οι άνθρωποι σημαίνουν ότι δύο άτομα θα μπορούσαν να έχουν πολύ διαφορετικά πλαίσια αναφοράς. Όπως δήλωσε ο Weizenbaum, “δεν μπορεί να υπάρξει απόλυτη κατανόηση και καμία απολύτως αξιόπιστη δοκιμασία κατανόησης” [118]. Για τον Weizenbaum το κλειδί ήταν ότι οι άνθρωποι καταλαβαίνουν ο ένας τον άλλον “σε επαρκή βαθμό”, αλλά αυτό που μπορούσε να επιτύχει προγράμματα όπως το ELIZA ήταν μόνο “να αντιμετωπίσουμε τέτοιες ιδέες συμβολικά” [118]. Το πρόγραμμα θα μπορούσε, μέσω έξυπνου προγραμματισμού, να μιμείται την ανθρώπινη κατανόηση. Αυτό δεν ήταν από μόνο του απόδειξη κατανόησης, αλλά απλώς απόδειξη της επιτυχούς υλοποίησης και εκτέλεσης της παραπάνω εικασίας του Weizenbaum σχετικά με την επικοινωνία.

```
Welcome to

      EEEEE  LL      IIII  ZZZZZZ  AAAAA
      EE     LL      II    ZZ     AA  AA
      EEEEE  LL      II    ZZZ    AAAAAAA
      EE     LL      II    ZZ     AA  AA
      EEEEE  LLLLLL IIII  ZZZZZZ  AA  AA

Eliza is a mock Rogerian psychotherapist.
The original program was described by Joseph Weizenbaum in 1966.
This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?
YOU:   Men are all alike.
ELIZA: What is the connection, do you suppose ?
YOU:   They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?
YOU:   Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here ?
YOU:   He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU:   It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy ?
YOU:
```

Παράδειγμα του προγράμματος “ELIZA”

Ακόμα κι αν ο Weizenbaum ήταν σκεπτικός για το βαθμό στον οποίο δύο άνθρωποι μπορούσαν πραγματικά να καταλάβουν ο ένας τον άλλον, ως δημιουργός του έμεινε έκπληκτος μάλλον περισσότερο από τους τρόπους με τους οποίους οι άλλοι γύρω του φαινόταν να αντιλαμβάνονται το ELIZA. Όπως έγραψε ο Weizenbaum *“άνθρωποι που γνώριζαν πολύ καλά ότι συνομιλούσαν με ένα μηχανήμα ξέχασαν σύντομα αυτό το γεγονός”*, ορισμένοι άνθρωποι θα απαιτούσαν ακόμη και *“να τους επιτραπεί να συνομιλήσουν με το σύστημα ιδιωτικά και, μετά από τη συνομιλία με το ELIZA, επέμεναν, παρά τις εξηγήσεις μου, ότι η μηχανή τους κατάλαβε πραγματικά”* [118]. Επιπλέον, ο βαθμός στον οποίο το ELIZA μπόρεσε να μιμηθεί επιτυχώς το έργο των Rogerian ψυχαναλυτών εντυπωσίασε πολύ πολλούς ψυχιατρους, μερικοί από τους οποίους πρότειναν ότι το πρόγραμμα θα μπορούσε να χρησιμοποιηθεί για θεραπεία με πραγματικούς ασθενείς. Μπερδεμένος από τις απαντήσεις του ELIZA, ο Weizenbaum βρέθηκε απογοητευμένος από ορισμένες τάσεις στον τομέα της επιστήμης των υπολογιστών, όπως τάσεις που προσπάθησαν να εξισώσουν τον άνθρωπο ως παρόμοιο με τους υπολογιστές και ισχυρισμούς που έκαναν τον ανθρώπινο εγκέφαλο ως *“απλώς μια μηχανή κρέατος”* [118].

Αυτό που έγινε σαφές για τον Weizenbaum ήταν ότι ο υπολογιστής είχε γίνει όχι μόνο ένα ισχυρό εργαλείο στη ζωή των ανθρώπων, αλλά ότι *“επιτρέψαμε στις ίδιες τις τεχνολογικές μεταφορές να διεισδύσουν τόσο διεξοδικά στις διαδικασίες σκέψης μας που τελικά παραδώσαμε στην τεχνολογία το ίδιο καθήκον να διατυπώνουμε ερωτήσεις”* [118]. Μέρος αυτής της τεχνολογικής μεταφοράς ήταν το αποτέλεσμα των επιστημόνων των υπολογιστών να παρακάμπτουν την ευθύνη για αυτό που είχαν δημιουργήσει.

Το βιβλίο *Computer Power and Human Reason* ξεκινάει με τον Weizenbaum να εξηγεί πώς οι εμπειρίες του με τους υπολογιστές τον οδήγησαν να γράψει ένα βιβλίο τόσο επικριτικό για αυτά τα μηχανήματα. Το βιβλίο ξεκινάει με το ELIZA όπου αφηγείται την ιστορία του προγράμματος με τρόπο που τονίζει τη δική του έκπληξη στις αντιδράσεις που προκάλεσε το πρόγραμμα όπως όταν ότι οι ασκούμενοι ψυχίατροι πίστευαν ότι το πρόγραμμα είχε θεραπευτικές δυνατότητες. Παραδέχεται ότι βρέθηκε προ εκπλήξεως από την ευκολία με την οποία οι άνθρωποι επένδυσαν συναισθηματικά στην επικοινωνία τους με τον υπολογιστή και τονίζει την έκπληξή του στον αριθμό των ατόμων στο δικό του επιστημονικό πεδίο που πίστεψαν ότι το ELIZA αντιπροσώπευε ένα πρόγραμμα που θα μπορούσε πραγματικά να κατανοήσει αυτά που λαμβανε στη φυσική γλώσσα [19].

Ο Weizenbaum τόνισε ότι ο υπολογιστής δεν ήταν το πρόβλημα, αλλά επαληθεύει κάτι που ήταν από καιρό μια επικίνδυνη κοινωνική τάση, να βλέπουμε τα ανθρώπινα όντα να λειτουργούν με έναν ολοένα και πιο μηχανικό τρόπο. Ξεκίνησε μια συζήτηση στην εκτίμηση του Weizenbaum και *“από τη μία πλευρά είναι εκείνοι που, εν συντομία, πιστεύουν ότι οι υπολογιστές μπορούν, πρέπει και θα κάνουν τα πάντα, και από την άλλη πλευρά εκείνοι που πιστεύουν ότι υπάρχουν όρια σε αυτό που πρέπει να έχουν οι υπολογιστές”* [118]. Η παρουσία του *“θα έπρεπε”* είναι ιδιαίτερης σημασίας για το επιχείρημα του Weizenbaum, καθώς αλλάζει τη συζήτηση από μια επικεντρωμένη στο ποια λειτουργικότητα μπορούν να

έχουν οι υπολογιστές για το αν πρέπει να κατασκευάζονται ή όχι να εκτελούν τέτοιες λειτουργίες.

Αυτό που ξεχωρίζει τον υπολογιστή από άλλα εργαλεία ήταν ο βαθμός στον οποίο αυτά τα μηχανήματα ήταν αυτόνομα - πράγμα που σημαίνει ότι μόλις ενεργοποιηθούν δεν χρειάζεται περαιτέρω ανθρώπινος έλεγχος. Ενώ τα ρολόγια ήταν σημαντικά πρώιμα παραδείγματα τέτοιων αυτόνομων μηχανημάτων, οι υπολογιστές ήταν αυτόνομοι αλλά ήταν ικανοί για πολύ πιο σημαντικές λειτουργίες από το να παρακολουθούν απλώς τον χρόνο. Η σημασία τέτοιων μηχανών ήταν ότι λειτουργούσαν σύμφωνα με ένα μοντέλο κάποιου μέρους του πραγματικού κόσμου - όπως οριοθέτηση μιας ημέρας σε ποσότητα 24 ωρών με κάθε ώρα να αποτελείται από εξήντα λεπτά και με κάθε λεπτό να αποτελείται από εξήντα δευτερόλεπτα. Σταδιακά αυτό που προέκυψε ήταν ότι μιμούμενοι κάποια πτυχή της πραγματικότητας, αυτές οι αυτόνομες μηχανές θα ερχόταν να ενσταλάξουν αυτό το μοντέλο πραγματικότητας στους ανθρώπους που είχαν αρχικά κατασκευάσει τη μηχανή και το μοντέλο ήρθε να αντικαταστήσει αυτό που ήταν μοντελοποίηση. Έτσι, υπό την αιγίδα της τεχνολογίας, οι εμπειρίες της πραγματικότητας έπρεπε να είναι αντιπροσωπευτικές ως αριθμοί για να φανούν νόμιμοι στα μάτια της κοινής σοφίας” [118].

Ο υπολογιστής εμφανίστηκε ως απαραίτητο στοιχείο της αναδυόμενης σύγχρονης κοινωνίας. Το τελικό αποτέλεσμα αυτής της τάσης ήταν ότι έγινε σχεδόν αδιανόητο να επιστρέψουμε στον τρόπο ζωής που υπήρχε πριν. Ωστόσο, μόνο και μόνο επειδή ο υπολογιστής είχε φανεί απαραίτητος δεν σημαίνει ότι ήταν πραγματικά. Μάλλον αυτό που συνέβη ήταν ότι ο υπολογιστής είχε απλώς γίνει *“απαραίτητος για την επιβίωση της κοινωνίας με τη μορφή που ο ίδιος ο υπολογιστής είχε καθοριστικό ρόλο στη διαμόρφωση με αυτή τη μορφή”* [118]. Ακόμη και το 1976, ο Weizenbaum γνώριζε ότι ο υπολογιστής είχε ήδη συνδεθεί πολύπλοκα με την κοινωνία, που ως εκ τούτου είχε σημασία ήταν να αναγνωρίζει πώς “οι νέοι τρόποι δράσης της κοινωνίας συχνά εξαλείφουν τη δυνατότητα δράσης με παλαιότερους τρόπους” [118]. Ο υπολογιστής βασίστηκε σε συγκεκριμένους τύπους πληροφοριών, υπερέχει σε συγκεκριμένους τύπους εργασιών και καθοδηγείται από συγκεκριμένες κοινωνικές, πολιτικές και οικονομικές δυνάμεις - και παρόλο που ο υπολογιστής θα μπορούσε να απεικονίζεται ως άνοιγμα πόρτας - ο Weizenbaum τόνισε ότι είχε επίσης κλείσει πολλές.

Στο *Computer Power and Human Reason*, ο Weizenbaum διερεύνησε τον τρόπο με τον οποίο δουλεύουν οι υπολογιστές, παρουσιάζοντας τα περίπλοκα υπολογιστικά σενάρια ως επεξήγηση σχετικά απλών παιχνιδιών, επισημαίνοντας ότι πρέπει να θυμόμαστε ότι ο υπολογιστής είναι εντελώς δεσμευμένος να ακολουθήσει τους κανόνες του παιχνιδιού. Ο Weizenbaum εξηγεί τις μηχανές Turing και εξηγεί τους τρόπους με τους οποίους ένας υπολογιστής μπορεί να συσσωρεύει πολλές ρουτίνες τη μια πάνω στην άλλη, δηλώνοντας, “το αλφάβητο των γλωσσών μηχανών όλων των σύγχρονων υπολογιστών είναι το σύνολο που αποτελείται από τα δύο σύμβολα ‘0’ και ‘1’” [118]. Παρά την αποτελεσματικότητα του υπολογιστή, παραμένει υποχρεωμένο να ακολουθεί τους κανόνες του προγραμματισμού του

και να βασίζεται στη δική του συγκεκριμένη γλώσσα. Το γεγονός ότι ένα πρόγραμμα μπορεί να εκτελέσει με επιτυχία το σενάριό του δεν σημαίνει ότι έχει πραγματική κατανόηση του κόσμου. Πράγματι, όπως το θέτει ο Weizenbaum, “ενας πραγματικός λόγος για τον οποίο ο προγραμματισμός είναι πολύ δύσκολος είναι ότι, στις περισσότερες περιπτώσεις, ο υπολογιστής δεν γνωρίζει τίποτα από αυτές τις πτυχές του πραγματικού κόσμου με τις οποίες προορίζεται να αντιμετωπίσει το πρόγραμμά του”.

Οι υπολογιστές και τα προγράμματα δεν εμφανίζονται οργανικά στη φύση. Αντ' αυτού, ο υπολογιστής και τα προγράμματά του είναι οι φυσικές εκδηλώσεις των συνόλων επιλογών που έχουν γίνει από τα ανθρώπινα όντα. Ως επιστήμονας υπολογιστών και καθηγητής στο MIT, ο Weizenbaum γνώριζε καλά τους τύπους ανθρώπων που ήταν υπεύθυνοι για τη λήψη των αποφάσεων που οδήγησαν στα συστήματα υπολογιστών που τελικά θα χρησιμοποιούσε το ευρύτερο κοινό. Για τον Weizenbaum υπάρχει μια διαφορά μεταξύ του “επαγγελματία” σαν τον εαυτό του που “θεωρεί τον προγραμματισμό ως μέσο προς το τέλος, όχι ως αυτοσκοπό” και αυτό που ο Weizenbaum αναγνώρισε ως “ο υποχρεωτικός προγραμματιστής”. Ο Weizenbaum περιγράφει “τον καταναγκαστικό προγραμματιστή” ως άτομο που βλέπει την αλληλεπίδραση με τον υπολογιστή ως αυτοσκοπό και ο κύριος στόχος του είναι να εργάζεται στον υπολογιστή.

Αυτό που τραβάει τον “υποχρεωτικό προγραμματιστή” στον υπολογιστή, κατά την εκτίμηση του Weizenbaum, είναι μια γοητεία και λατρεία για τη δύναμη που εκδηλώνεται στα συστήματα υπολογιστών. Για τον “καταναγκαστικό προγραμματιστή, η ζωή δεν είναι τίποτα άλλο από ένα πρόγραμμα που τρέχει σε έναν τεράστιο υπολογιστή” και έτσι “κάθε πτυχή της ζωής μπορεί τελικά να εξηγηθεί με όρους προγραμματισμού”. Σύμφωνα με τον Weizenbaum, ο κίνδυνος των υπολογιστών θα μπορούσε να βρεθεί με τον τρόπο που όλο και περισσότερα άτομα που εργάζονταν σε αυτά είχαν έρθει να αντιπροσωπεύουν αυτόν τον τύπο προγραμματιστών των οποίων η πίστη στον υπολογιστή, ως τέτοια, είχε ξεπεράσει οποιαδήποτε άλλη αξία. Παρ' όλα αυτά, δεν το αποδίδει σε κακία, αλλά σε κάποια κενή απροσεξία κατά την οποία ορισμένοι προγραμματιστές μαζί με άλλους σύγχρονους επιστήμονες και τεχνολόγους - είχαν μπερδέψει τα τεχνο-επιστημονικά τους μέσα με το να είναι αυτοσκοπός. Και παρόλο που οι “καταναγκαστικοί προγραμματιστές” μπορεί να μην έχουν έλλειψη δεξιοτήτων, αυτή η “ικανότητα” είναι ουσιαστικά χωρίς κάποιο στόχο.

Η προσωπική εμπειρία του Weizenbaum με τις προβληματικές παρανοήσεις αφορούσε σε μεγάλο βαθμό τις εμπειρίες του με το ELIZA. Το ELIZA ήταν πολλά πράγματα, αλλά η εξυπνάδα δεν ήταν ένα από αυτά. Για τον Weizenbaum, το ELIZA ήταν μια επίδειξη του τρόπου με τον οποίο οι άνθρωποι ήταν πρόθυμοι να αποδώσουν νοημοσύνη σε μηχανήματα όταν αυτό δεν ήταν δικαιολογημένο - και αυτή ήταν μια τάση που είδε έντονα ανάμεσα σε ορισμένους από τους συναδέλφους του που εργάζονται στον τομέα της τεχνητής νοημοσύνης. Όσοι δεν είναι εξοικειωμένοι με τους υπολογιστές είχαν λόγο να γοητεύονται από την απόδοση του ELIZA, αλλά ποια ήταν η δικαιολογία των επιστημόνων στην Τεχνητή Νοημοσύνη;

Η κοινωνική κριτική του Weizenbaum είναι η μνήμη του να εγκαταλείπει τη Ναζιστική Γερμανία ως παιδί. Δεν εστιάζει στον Χίτλερ και την ηγεσία του ναζιστικού κόμματος, αλλά η εστίασή του είναι στο πρόσωπο του “καλού Γερμανού” - εκείνου που επικαλείται άγνοια για τις φρίκες που συμβαίνουν γύρω του, εκείνων των ανθρώπων που έχασαν την ευθύνη έναντι της υπόλοιπης ανθρωπότητας. Η δικαιολογία ότι η προσεκτική οργάνωση του ναζιστικού καθεστώτος κράτησε “τον καλό Γερμανό” σε κατάσταση άγνοιας δεν πείθει τον Weizenbaum, ο οποίος αντίθετα προτείνει “ο πραγματικός λόγος που δεν γνώριζε ο καλός Γερμανός είναι ότι δεν ένιωσε ποτέ ότι είναι δική του ευθύνη να ρωτήσει τι είχε συμβεί στον Εβραίο γείτονα του, του οποίου το διαμέρισμα ξαφνικά έγινε διαθέσιμο”. Όταν ο Weizenbaum επέστρεψε για πρώτη φορά στη Γερμανία τη δεκαετία του 1950, βρέθηκε να κοιτάζει γύρω τους Γερμανούς που είχαν ζήσει τα ναζιστικά χρόνια και αναρωτιόταν, εάν όλους αυτούς τους απασχολούσε τι έκαναν, έκλεισαν το στομα τους; Αντιστάθηκαν; Συμμετείχαν;

2.4 Η Κυβερνοηθική στις μέρες μας

Ο Weizenbaum απεβίωσε το 2008. Κατά τη διάρκεια της ζωής του βρισκόταν στην πρώτη γραμμή των σημαντικών αλλαγών που υποβλήθηκαν οι υπολογιστές από μηχανικά μεγαθήρια που απαιτούσαν ολόκληρα δωμάτια, σε προσωπικούς υπολογιστές, και στις πρώτες ενσαρκώσεις του smartphone. Καθώς έβλεπε τους υπολογιστές να γίνονται όλο και μικρότεροι, ολοένα και πιο δυνατοί και όλο και πιο δεσμευμένοι στις δραστηριότητες της καθημερινής ζωής, η κριτική του παραμένει σταθερή. Για αυτόν, η στρατιωτική προέλευση του υπολογιστή δεν μπορούσε απλώς να ξεχαστεί ως μια άβολη ιστορική λεπτομέρεια.

Τα άρθρα του Weizenbaum και το βιβλίο του *Computer Power and Human Reason* συνεχίζουν να έχουν σημαντική επιρροή στους μελετητές που γράφουν σχετικά με τον αντίκτυπο των υπολογιστών στην κοινωνία. Στο βιβλίο “Ο κλειστός κόσμος: Υπολογιστές και η πολιτική του λόγου στον Ψυχρό Πόλεμο της Αμερικής” [28] ο Edwards υποστηρίζει ότι “τα εργαλεία και οι μεταφορές συνδέονται μέσω του λόγου” και βασίζεται στην εκτίμηση του Weizenbaum για την τεχνολογική μεταφορά καθώς συζητά πώς “τα εργαλεία και οι χρήσεις τους αποτελούν αναπόσπαστο μέρος του ανθρώπινου λόγου και, μέσω του λόγου, όχι μόνο διαμορφώνουν την υλική πραγματικότητα άμεσα αλλά επίσης διαμορφώνουν τα διανοητικά μοντέλα, τις έννοιες και τις θεωρίες που καθοδήγηση που διαμορφώνουν ” [81]. Η περιγραφή του Weizenbaum σχετικά με τον καταναγκαστικό προγραμματιστή αντηχεί με την πάροδο των ετών, ακόμη και όταν μια τέτοια μορφή έχει μεταφερθεί από τις σκοτεινές γωνίες των εργαστηρίων υπολογιστών των πανεπιστημίων, στις αίθουσες συνεδριάσεων των μεγάλων εταιρειών. “Ο Weizenbaum ισχυρίστηκε ότι ο προγραμματισμός δημιουργεί μια νέα ψυχική διαταραχή στον καταναγκασμό για προγραμματισμό” [120] όπως γράφει ο Wendy Hui Kyong Chung.

Οι πτυχές της σκέψης του Weizenbaum έχουν αμφισβητηθεί με την πάροδο των ετών, και ορισμένες από τις προβλέψεις του έχουν αποδειχθεί απλώς λανθασμένες, όπως το σχόλιό του, που τέθηκε το 1978: “Ο οικιακός υπολογιστής θα είναι τόσο διαδεδομένος όσο οι σημερινές τηλεοράσεις;” [81]. Η απάντηση του Weizenbaum ήταν “όχι” και αποδεικνύεται ψευδής σε έναν κόσμο με smartphones, tablet, φορητών υπολογιστών και τηλεοράσεων, που συνδέονται με το Διαδίκτυο. Ωστόσο, ακόμη και αν ορισμένα από τα σχόλια του Weizenbaum σχετικά με τους υπολογιστές έχουν ξεπεραστεί, τα επιχειρήματά του δεν έχουν χάσει τίποτα από το ηθικό τους βάρος. Ο David Golumbia, στην *The Cultural Logic of Computation* (Πολιτιστική Λογική του Υπολογισμού), επικρίνει την άνοδο του “υπολογισμού”, την οποία ορίζει “ως δέσμευση για την άποψη ότι πολλά, ή ίσως και όλες οι ανθρώπινες και κοινωνικές εμπειρίες μπορούν να εξηγηθούν μέσω υπολογιστικών διαδικασιών” [117] και σημειώνει ότι “ο Weizenbaum διαφώνησε δημόσια από την υπολογιστική άποψη και συνέχισε να γράφει έναν συναρπαστικό τόμο για τα προβλήματά του” [117].

Όταν ο Golumbia συντάσσει τη δική του λίστα “καθιερωμένων μελετητών” που είχαν επικρίνει τον “υπολογισμό” - όχι σε αντίθεση με μια λίστα με εκείνους που αμφισβητούν την “απεριόριστη πορεία της επιστήμης και της τεχνολογίας”- τοποθετεί τον Weizenbaum σε μια λίστα μαζί με πολλά από τα ίδια άτομα που είχαν εμφανιστεί στη λίστα του Weizenbaum.

Ο Luciano Floridi είναι καθηγητής στο Πανεπιστήμιο της Οξφόρδης όπου είναι επίσης Διευθυντής του Εργαστηρίου Ψηφιακής Ηθικής του Ινστιτούτου Διαδικτύου της Οξφόρδης. Από το 2008 μέχρι και το 2013, κατείχε την Έδρα Έρευνας στη φιλοσοφία της πληροφορίας και την Πρόεδρος της UNESCO στην Πληροφορική και Ηθική των Υπολογιστών στο Πανεπιστήμιο του Hertfordshire.

Είναι κυρίως γνωστός για το έργο του σε δύο τομείς της φιλοσοφικής έρευνας: τη φιλοσοφία της πληροφορίας και της ηθικής της πληροφορίας. Η έρευνά του αφορά κυρίως την Ηθική των Πληροφοριών και των Υπολογιστών (γνωστή και ως Ψηφιακή Ηθική), τη Φιλοσοφία της Πληροφορίας και τη Φιλοσοφία της Τεχνολογίας. Άλλα ερευνητικά ενδιαφέροντα περιλαμβάνουν την Επιστημολογία, τη Φιλοσοφία της Λογικής και την Ιστορία και τη Φιλοσοφία του Σκεπτικισμού. Το έργο του είναι μια τετραλογία στη βάση της φιλοσοφίας της πληροφορίας, που ονομάζεται *Principia Philosophiae Informationis*.

Σύμφωνα με τον Floridi, είναι απαραίτητο να αναπτυχθεί μια κατασκευαστική φιλοσοφία, όπου ο σχεδιασμός, η μοντελοποίηση και η εφαρμογή αντικαθιστούν την ανάλυση. Με τη μετάβαση από το ένα σύνολο των καθηκόντων στο άλλο, η φιλοσοφία θα μπορούσε στη συνέχεια να σταματήσει να υποχωρεί στην ολοένα και πιο μικρή γωνία των αυτοσυντηρούμενων ερευνών της και, συνεπώς, να αποκτήσει μια ευρύτερη άποψη για το τι έχει σημασία. Ο Floridi έχει καταφέρει να χαρακτηρίσει την φιλοσοφία ως ένα καινοτόμο πεδίο, πλέον γνωστό ως η φιλοσοφία της πληροφορίας, ένας νέος τομέας έρευνας που προέκυψε από την υπολογιστική / ενημερωτική στροφή.

Προσεγγίζει τη φιλοσοφία της πληροφορίας από τη σκοπιά της λογικής της επιστήμης των υπολογιστών, της θεωρητικής πληροφορικής. Ενδεικτικά αναφέρουμε ότι στην εισαγωγή του βιβλίου του *Philosophy and Computing*, αναφέρει ότι το βιβλίο προορίζεται για φοιτητές φιλοσοφίας που χρειάζονταν γνώσεις πληροφορικής για να χρησιμοποιούν υπολογιστές αποτελεσματικά ή απαραίτητες βασικές γνώσεις για την κριτική κατανόηση της ψηφιακής μας εποχής. Το ερευνητικό ενδιαφέρον για τον κλάδο της Πληροφορικής Φιλοσοφίας και Φιλοσοφίας Πληροφοριών άρχισε να αναπτύσσεται περισσότερο.

Επί του παρόντος, ο Floridi εργάζεται σε δύο τομείς έρευνας: ηθική των υπολογιστών (δείτε την ηθική των πληροφοριών εισόδου) και την έννοια της πληροφορίας. Το κλειδί σε αυτόν τον τομέα εργασίας είναι ο ισχυρισμός ότι οι Τεχνολογίες Πληροφοριών και Επικοινωνιών επανεξετάζουν ριζικά ή επαναπροσδιορίζουν την σφαίρα πληροφοριών, δεδομένων, γνώσεων και επικοινωνίας

Η Τεχνητή Νοημοσύνη έχει ήδη σημαντικό αντίκτυπο στην κοινωνία. Τα βασικά ερωτήματα είναι πώς, πού, πότε και από ποιον θα γίνει αισθητός ο αντίκτυπος της τεχνητής νοημοσύνης. Ως αποτέλεσμα, πολλοί οργανισμοί έχουν ξεκινήσει ένα ευρύ φάσμα πρωτοβουλιών για τη θέσπιση ηθικών αρχών για την υιοθέτηση κοινωνικά ωφέλιμης τεχνητής νοημοσύνης. Δυστυχώς, ο τεράστιος όγκος των προτεινόμενων αρχών απειλεί να γίνει συντριπτικός και σύγχυση, θέτοντας δύο πιθανά προβλήματα. Είτε τα διάφορα σύνολα ηθικών αρχών για το AI είναι παρόμοια, οδηγώντας σε περιττή επανάληψη και πλεονασμό ή, εάν διαφέρουν σημαντικά, θα προκύψει σύγχυση και ασάφεια.

2.5 Συμπεράσματα

Το παρόν κεφάλαιο, λοιπόν, παρέχει μια επισκόπηση της ηθικής από τους Σοφιστές, τον Πλάτωνα, τον Αριστοτέλη και τον Καντ και στη συνέχεια την εφαρμοσμένη ηθική και κυβερνητική με κύρια σημεία τον Wiener και τον Weizenbaum καταλείγοντας στον Luciano Floridi και τις τάσεις και προσεγγίσεις της ηθικής στις μέρες μας.

Όπως γράφει ο Weizenbaum, με λίγο χιούμορ: *“Το Διαδίκτυο είναι μια μεγάλη χωματερή, φυσικά με μερικά μαργαριτάρια μέσα σε αυτό, αλλά χρειάζεται πρώτα να τα βρείτε”* [119]. Ως κοινωνικός κριτικός βρήκε, και σε αυτόν τον τόμο συνεχίζει να αποδεικνύει τη δέσμευσή του να δηλώνει δυνατά άβολες απόψεις: *“ο υπολογιστής είναι ενσωματωμένος στην τρελή κοινωνία μας, όπως και η τηλεόραση. Όλα είναι ενσωματωμένα σε αυτήν την κοινωνία, και αυτή η κοινωνία είναι προφανώς παράλογη”* [119]. Για τον Weizenbaum ήταν απαραίτητο για τους ανθρώπους να γίνουν “Islands of Reason”, γιατί το δυναμικό παρέμεινε ότι τέτοια νησιά θα μπορούσαν να προσελκύσουν παρόμοιους ανθρώπους πρόθυμους να αναγνωρίσουν ότι *“ίσως είμαστε τώρα εθισμένοι στη σύγχρονη επιστήμη και τεχνολογία και πρέπει να εξασκηθούν θεραπείες”* [119].

Είναι σημαντικό ότι αυτό δεν είναι μια έκκληση για μια μοναστική ζωή και αποχώρηση από τα κοινά, αλλά μια έκκληση για μια βαθύτερη επένδυση στα ηθικά ζητήματα. Για να αναγνωρίζουν οι άνθρωποι την τρέλα στον κόσμο γύρω τους και όταν προκύπτει μια τέτοια αναγνώριση *“πρέπει να μιλήσουμε, πρέπει να μοιραστούμε αυτό που έχουμε συνειδητοποιήσει με άλλους”* [119].

Κεφάλαιο 3

Εφαρμογές και Επιπτώσεις

Σε αυτό το κεφάλαιο, θα επικεντρωθούμε σε τρέχουσες εφαρμογές αλγορίθμων και πώς μπορούμε να αναπτύξουμε μια ηθική κατανόηση αυτών. Γι' αυτό, η προσοχή θα δοθεί πρώτα στην κανονιστική ηθική. Μια σύντομη περιγραφή, ορισμό, επισκόπηση και αιτιολόγηση για την υιοθέτηση μιας τέτοιας θέσης. Στη συνέχεια, θα εισαχθεί η έννοια της τεχνολογίας διπλής χρήσης, για να πλαισιώσει αργότερα τους αλγόριθμους, όπου πέρα από την προφανή λειτουργία και χρήση της έχουν και μια δεύτερη φύση, αποτέλεσμα. Αυτό σημαίνει ότι υπάρχουν τόσο θετικές όσο και αρνητικές πτυχές στην χρήση αλγορίθμων. Αυτή είναι μια αρκετά συνήθης πρακτική στην αξιολόγηση των τεχνολογιών από την επακόλουθη προοπτική [\[52\]](#).

3.1 Η Συνεπειοκρατική Κανονιστική Ηθική

Ο συνεπειοκρατισμός είναι η φιλοσοφική παράδοση που υποστηρίζει ότι οι κανονιστικές ιδιότητες εξαρτώνται μόνο από τις συνέπειες των ενεργειών. Υποστηρίζει ότι πρέπει να “βασίσουμε τις ενέργειές μας στην προώθηση καλών συνεπειών και στην αποφυγή κακών” [\[71\]](#), δηλαδή, κατ'επέκταση, για τον εντοπισμό και την επιδίωξη αυτού που θεωρείται προτιμότερο ενώ ελαχιστοποιεί τυχόν αρνητικές επιπτώσεις. Είναι αυθαίρετο αυτό που θεωρείται προτιμότερο, για ποιον και σε ποιο βαθμό αυτές οι αυθαίρετες αξιολογήσεις πρέπει να αντισταθμίζουν η μία την άλλη, με αποτέλεσμα διαφωνίες μεταξύ των ίδιων των συνεπειών [\[54\]](#).

Οι συνεπειοκρατικές θεωρίες στηρίζονται τόσο στους σκοπούς όσο και στα αποτελέσματα των πράξεων. Οι πράξεις, λοιπόν, χαρακτηρίζονται ορθές ή λανθασμένες αποκλειστικά δυνάμει των συνεπειών τους. Σύμφωνα με την ωφελιμιστική τοποθέτηση, η αρχή, η οποία πρεσβεύεται, είναι: “για την παραγωγή του μεγαλύτερου δυνατού καλού για το μεγαλύτερο δυνατό αριθμό ανθρώπων” [\[55\]](#). Συνεπώς, στην κανονιστική συνεπειοκρατία ανήκουν οι ωφελιμιστές, οι οποίοι πέρα από την ευχαρίστηση, την ευτυχία ή τις προσωπικές προτιμήσεις, οι οποίες μπορούν να κρίνουν τις συνέπειες ποικίλων πράξεων μας προτείνοντας ηθικούς κανόνες, θεωρούν ότι υπάρχουν και άλλα στοιχεία, τα οποία πρέπει να ληφθούν υπόψη.

Το κεφάλαιο αυτό, θα δούμε πώς η τεχνολογία “μπορεί να προκαλέσει βλάβη σε ορισμένες περιπτώσεις”. Με χρήση ισχυρών επιχειρημάτων, γιατί οι αλγόριθμοι δεν μπορούν να θεωρηθούν ουδέτερα μέσα για την επίτευξη ορισμένων στόχων, αλλά πρέπει να θεωρηθούν

ότι έχουν ηθικές επιπτώσεις, απαιτώντας συνεχή σκέψη. Ο συνεπειοκρατισμός εν γένει θεωρείται ευρέως ως ένα καλό σημείο εκκίνησης για ηθικές αναλύσεις.

3.2 Η Τεχνολογία Διπλής Χρήσης

Αρχικά θα εισαγουμε την έννοια της τεχνολογίας διπλής χρήσης. Η δυνατότητα αλγορίθμων για τη βελτίωση της ατομικής και κοινωνικής ευημερίας συνοδεύεται από σημαντικούς ηθικούς κινδύνους, κάτι που αμέσως μας κάνει να σκεφτούμε ότι οι αλγόριθμοι δεν είναι ηθικά ουδέτεροι. Για να κάνουμε αυτή την σκέψη περισσότερο επίσημη και όχι απλώς ως μια αίσθηση, της υποκειμενικότητας των αλγορίθμων, θα δούμε ότι οι αλγόριθμοι είναι τεχνολογία διπλής χρήσης. Θα επικεντρωθούμε στη “διπλή” φύση που έχουν, που δημιουργεί ηθικά θέματα.

Η έννοια των τεχνολογιών διπλής χρήσης αναφέρεται ευρέως στην ιδέα ότι υπάρχουν νόμιμες και παράνομες πτυχές σε μια τεχνολογία, όπου ο αυθαίρετος διαχωρισμός μεταξύ των δύο κατηγοριών εξαρτάται από το πλαίσιο και μπορεί να αλλάζει με την πάροδο του χρόνου. Για να το περιγραφεί με έναν τρόπο περισσότερο πρακτικό, ο Forge [\[39\]](#) προσπάθησε να δώσει έναν πιο καταλληλο ορισμό. Ωστόσο, με την προσπάθεια εύρεσης μιας μέσης οδού, ο τελικός ορισμός φαίνεται να εστιάζεται πολύ σε συγκεκριμένους τύπους δυνητικά επιβλαβών τεχνολογιών:

“Ένα αντικείμενο (τεχνολογία γνώσης, τεχνικό αντικείμενο) είναι διπλής χρήσης εάν υπάρχει (αρκετά υψηλός) κίνδυνος να χρησιμοποιηθεί για το σχεδιασμό ή την παραγωγή όπλου, ή εάν υπάρχει (αρκετά μεγάλη) απειλή ότι μπορεί να χρησιμοποιηθεί σε αυτοσχέδιο όπλο, όπου σε καμία περίπτωση δεν είναι όπλο, όπου σε καμία περίπτωση υπάρχει ανάπτυξη όπλων στον επιδιωκόμενο ή πρωταρχικό σκοπό”.

Ωστόσο, άλλοι ορισμοί είναι λιγότερο περιοριστικοί στη χρήση τεχνολογιών διπλής χρήσης και επισημαίνουν χρήσεις που δεν προκαλούν άμεση σωματική βλάβη. Για παράδειγμα, η απόκρυψη κώδικα χρησιμοποιείται για νόμιμη προστασία λογισμικού, αλλά και από κακόβουλο λογισμικό, επισημαίνοντας τη φύση της διπλής χρήσης. Αυτό τονίζει την τεχνολογία διπλής φύσης ως έννοια, και εξηγείται από τους Van Wynsberghe και Nagenborg [\[114\]](#) να αναφέρονται τόσο σε μια τεχνολογία που χρησιμοποιείται από διπλές ομάδες ανθρώπων, ιδίως στρατιωτικές και πολιτικές ομάδες, καθώς “χρησιμοποιείται για καλούς ή κακούς σκοπούς”.

Η έννοια της τεχνολογίας διπλής χρήσης είναι ευρετική για την επισήμανση των δυσμενών πλευρών των αλγορίθμων. Έχει ήδη επισημανθεί από τον Miller στο πλαίσιο των βιοεπιστημών: [\[69\]](#)

“οι περισσότεροι επιστήμονες και μηχανικοί δεν ξοδεύουν πολύ χρόνο για να σκεφτούν τις ανεπιθύμητες ή απροσδόκητες παρενέργειες που μπορεί να προκύψουν όταν

χρησιμοποιούνται τα προϊόντα τους. Σκέφτονται ακόμη λιγότερο για σκόπιμη κατάχρηση. Η ευαισθητοποίηση επιστημόνων, μηχανικών και άλλων σχεδιαστών για την πιθανή κατάχρηση του πνευματικού τους παιδιού είναι ο κύριος στόχος της πολιτικής διπλής χρήσης που έχει αναπτυχθεί τα τελευταία χρόνια”.

3.3 Μερικές Ηθικές Ανησυχίες Σχετικά με Αλγόριθμους

Οι αλγόριθμοι έχουν καταστεί βασικό στοιχείο που υποστηρίζει κρίσιμες υπηρεσίες και υποδομές των κοινωνιών της πληροφορίας. Τα άτομα αλληλεπιδρούν με συστήματα σύστασης - αλγοριθμικά συστήματα που κάνουν προτάσεις σχετικά με το τι μπορεί να θέλει ένας χρήστης - σε καθημερινή βάση, είτε να επιλέγουν ένα τραγούδι, μια ταινία, ένα προϊόν ή ακόμη και έναν φίλο [85]. Ταυτόχρονα, σχολεία και νοσοκομεία [101], χρηματοπιστωτικά ιδρύματα [36], δικαστήρια [38], τοπικοί κυβερνητικοί φορείς [68] και συνέχεια βασίζονται όλο και περισσότερο σε αλγόριθμους για τη λήψη σημαντικών αποφάσεων.

Η δυνατότητα αλγορίθμων για τη βελτίωση της ατομικής και κοινωνικής ευημερίας συνοδεύεται από σημαντικούς ηθικούς κινδύνους [37]. Στο επιχείρημα “Οι αλγόριθμοι δεν είναι ηθικά ουδέτεροι”, ας εξετάσουμε για παράδειγμα, πώς τα αποτελέσματα των αλγορίθμων μετάφρασης και μηχανών αναζήτησης θεωρούνται σε μεγάλο βαθμό ως αντικειμενικά, αλλά συχνά κωδικοποιούν τη γλώσσα με διαφορετικούς τρόπους [72]. Η προκατάληψη έχει επίσης αναφερθεί σε αλγοριθμική διαφήμιση, με ευκαιρίες για υψηλότερες αμοιβές και θέσεις εργασίας στον τομέα της επιστήμης και της τεχνολογίας που διαφημίζονται στους άνδρες συχνότερα από ό, τι στις γυναίκες [22]. Ενώ συζητούνται και σχεδιάζονται λύσεις σε αυτά τα ζητήματα, ο αριθμός των αλγοριθμικών συστημάτων που παρουσιάζουν ηθικά προβλήματα συνεχίζει να αυξάνεται.

Όμως αντί να δούμε παραδείγματα και να ισχυριζόμαστε ότι υπάρχουν πρότυπα μεταξύ τους που απαιτούν ηθική αξιολόγηση, θα ξεκινήσουμε με ένα θεωρητικό πλαίσιο πώς και γιατί οι αλγόριθμοι έχουν ηθικές επιπτώσεις και προεκτάσεις και στην συνέχεια με παραδείγματα να το κάνουμε όλο αυτό πιο κατανοητό. Θα χρησιμοποιήσουμε τις έξι διαστάσεις του Gillespie [43] για τους “αλγορίθμους δημόσιας συνάφειας”, έναν όρο που υποδηλώνει ότι οι αλγόριθμοι είναι κοινωνικής και επομένως ηθικής σημασίας. Αυτή η άποψη ενός υποσυνόλου αλγορίθμων δημοσιεύθηκε σε μια έκθεση του Global Conference of Cyberspace (GCCS) το 2015 [40].

3.3.1 Μοτίβα των ηθικών πτυχών των αλγορίθμων

Κατά κάποιον τρόπο, λοιπόν, οι υποδομές όπου λειτουργούν και χρησιμοποιούνται οι αλγόριθμοι φαίνεται να είναι σε αναντιστοιχία με τη λειτουργία των αλγορίθμων. Αυτές οι

υποδιαιρέσεις δείχνουν τον τρόπο με τον οποίο οι αλγόριθμοι είναι ένα συστατικό μιας μεγαλύτερης δομής, και πρέπει να διαχωριστούν και να μελετηθούν. Τα τρία σημεία που οδηγούν σε αδιαφάνεια:

1. Σκόπιμη απόκρυψη, από εταιρείες ή κράτος, έτσι ώστε οι διαδικασίες λήψης αποφάσεων να μην είναι ανοιχτές σε έλεγχο.
2. Τεχνικός αναφαβητισμός εκ μέρους όσων βρίσκονται εκτός της κοινότητας παραγωγής αλγορίθμων. Το αποτέλεσμα είναι ότι υπάρχει έλλειψη κατανόησης και γνώσης.
3. Μια αναντιστοιχία μεταξύ της πολυπλοκότητας των λειτουργιών και της ανθρώπινης κατανόησης, με την έννοια ότι η μαθηματική βελτιστοποίηση της ανάλυσης δεδομένων είναι δύσκολο να μεταφραστεί σε ανθρώπινο επίπεδο, αφήνοντας όλα αυτά μπερδεμένα.

3.3.2 Πτυχές του Gillespie

Για μια ολοκληρωμένη επισκόπηση των ηθικών πτυχών των αλγορίθμων, στραφούμε στις έξι αλληλένδετες διαστάσεις που όρισε ο Gillespie [\[43\]](#). Ο Gillespie ασχολείται πρωτίστως με τον τρόπο με τον οποίο οι αλγόριθμοι ήρθαν να μας ενημερώνουν και να διαμορφώνουν με πολλούς τρόπους. Αυτό αφού οι αλγόριθμοι:

- διαδραματίζουν ολοένα και σημαντικότερο ρόλο στην επιλογή των πληροφοριών που θεωρούνται πιο σχετικές για εμάς,
- οι μηχανές αναζήτησης μας βοηθούν να περιηγηθούμε σε τεράστιες βάσεις δεδομένων πληροφοριών ή σε ολόκληρο το web.
- χαρτογραφούν τις προτιμήσεις μας έναντι άλλων.
- διαχειρίζονται τις αλληλεπιδράσεις μας σε ιστότοπους κοινωνικής δικτύωσης, επισημαίνοντας τις ειδήσεις ενός φίλου, ή εξαιρούν ειδήσεις και πληροφορίες για άλλους.
- που έχουν σχεδιαστεί για να υπολογίζουν τι είναι "καυτό" ή "τάση" ή "πιο συζητημένο".

Έτσι λοιπόν, αυτοί οι αλγόριθμοι όχι μόνο μας βοηθούν να βρούμε πληροφορίες, αλλά παρέχουν ένα μέσο για να γνωρίζουμε **τι** πρέπει να γνωρίζουμε και **πώς** να γνωρίζουμε.

Επίσης σημειώνει ότι οι αλγόριθμοι δεν χρειάζεται να είναι λογισμικό “με την ευρύτερη έννοια, είναι κωδικοποιημένες διαδικασίες για τη μετατροπή των δεδομένων εισόδου σε μια επιθυμητή έξοδο, βάσει συγκεκριμένων υπολογισμών” [\[43\]](#). Θα μπορούσαμε να σκεφτούμε τους υπολογιστές ως μηχανές αλγορίθμων - σχεδιασμένοι να αποθηκεύουν και να διαβάζουν δεδομένα, να εφαρμόζουν μαθηματικές διαδικασίες με ελεγχόμενο τρόπο και να προσφέρουν νέες πληροφορίες ως έξοδο. Αλλά αυτές είναι διαδικασίες που θα μπορούσαν

να γίνουν χειροκίνητα - και στην πραγματικότητα μέχρι πριν κάποια χρόνια ήταν [\[67\]](#). Ωστόσο, καθώς έχουμε αγκαλιάσει τα υπολογιστικά εργαλεία ως τα κύρια μέσα έκφρασης, και δεν έχουμε κάνει μόνο τα μαθηματικά αλλά και όλες τις ψηφιακές πληροφορίες, “υποβάλλουμε ανθρώπινο λόγο και γνώση σε αυτές τις διαδικαστικές λογικές που διέπουν όλους τους υπολογισμούς” [\[43\]](#). Και υπάρχουν συγκεκριμένες επιπτώσεις όταν χρησιμοποιούμε αλγόριθμους για να επιλέξουμε τι είναι πιο σχετικό από ένα σύνολο δεδομένων που αποτελείται από ίχνη δραστηριοτήτων, προτιμήσεων και εκφράσεων.

Θα δοθεί μια σύντομη περιγραφή των έξι πτυχών του Gillespie για τους αλγόριθμους. Στη συνέχεια θα παρουσιάσουμε κάποια παραδείγματα για να διευκρινιστούν αυτά τα έξι σημεία σε βάθος.

1. Μοτίβα ένταξης:

Οι αλγόριθμοι είναι αδρανείς, χωρίς νόημα μηχανισμοί έως ότου συνδυαστούν με βάσεις δεδομένων στις οποίες θα λειτουργούν. Μια κοινωνιολογική έρευνα για έναν αλγόριθμο πρέπει πάντα να αντιμετωπίζει τις βάσεις δεδομένων με τις οποίες είναι συνδεδεμένος. Για τους χρήστες, οι αλγόριθμοι και οι βάσεις δεδομένων συνδέονται εννοιολογικά: οι χρήστες συνήθως τους αντιμετωπίζουν ως μία οντότητα. “Μαζί, οι δομές δεδομένων και οι αλγόριθμοι είναι δύο μισά της οντολογίας του κόσμου του υπολογιστή” [\[76\]](#). Ωστόσο, μπορούμε να εξετάσουμε αυτά τα δύο διακριτά.

Σε ένα δίκτυο επικοινωνίας σχεδιασμένο έτσι ώστε με κάθε σύνδεση, κάθε προβολή σελίδας και κάθε κλικ αφήνουμε ένα ψηφιακό ίχνος. “Τα ακατέργαστα δεδομένα είναι κάτι το οξυμωρο” [\[64\]](#), επισημαίνοντας ότι οι πληροφορίες που περιλαμβάνονται στη βάση δεδομένων πρέπει να μετατραπούν σε δεδομένα, να τροποποιηθούν έτσι ώστε οι αλγόριθμοι να μπορούν να ενεργήσουν αυτόματα.

Σημαντικό ρόλο παίζει η οργάνωση των βάσεων δεδομένων. Όσο λιγότερο άκαμπτη είναι η οργάνωση της βάσης δεδομένων καθώς χρησιμοποιούνται περισσότερο σχεσιακές και αντικειμενοστρεφείς αρχιτεκτονικές βάσεων δεδομένων, οι πληροφορίες μπορούν να οργανωθούν με πιο ευέλικτους τρόπους, όπου κομμάτια δεδομένων μπορούν να έχουν πολλαπλούς συσχετισμούς με άλλα κομμάτια δεδομένων. Όπως σημειώνει ο Rieder, με την ευρεία χρήση των σχεσιακών βάσεων δεδομένων έρχεται μια “σχεσιακή οντολογία” [\[107\]](#) που κατανοεί τα δεδομένα ως “χαλαρά συνδεδεμένα αντικείμενα που μπορούν να ταξινομηθούν με δυναμικά απεριόριστο αριθμό τρόπων εκείνη τη στιγμή ανάκτησης, μετατοπίζοντας έτσι την εκφραστική ισχύ από τη σχεδίαση της δομής της βάσης δεδομένων στο query που θα τεθεί στη βάση”. Η κατηγοριοποίηση είναι μια ισχυρή σημασιολογική παρέμβαση: ποιες είναι οι κατηγορίες, τι ανήκει σε μια κατηγορία και ποιος αποφασίζει πώς για να εφαρμόσουν αυτές τις κατηγορίες στην πράξη, είναι όλοι ισχυροί ισχυρισμοί για το πώς είναι τα πράγματα και πρέπει να είναι.

Μια άλλη πτυχή είναι ότι ο αποκλεισμός των “suggested videos” του YouTube εκτός από τις λίστες με τα βίντεο με τις περισσότερες προβολές [43]. Τα συγκεκριμένα μοτίβα με τα οποία οι πληροφορίες είτε εξαιρούνται είτε από μια βάση δεδομένων, ή επανειλημμένα παραλείπεται θα μπορούσε να θεωρηθεί λογοκρισία. Είτε λοιπόν οι επιλογές έγιναν από έναν συντάκτη εφημερίδων ή από εργαλεία ευρετηρίου μιας μηχανής αναζήτησης, αυτές οι επιλογές συμβάλλουν στον καθορισμό του περιεχομένου. Παρατηρούμε λοιπόν την αντίθεση ότι ενώ οι αλγόριθμοι χαρακτηρίζονται ως αυτόματοι, ενώ τα πρότυπα συμπερίληψης “προκαθορίζουν τι θα εμφανίζεται ή δεν θα εμφανίζεται μεταξύ των αποτελεσμάτων τους” [43].

2. Κύκλοι πρόβλεψης:

Οι αλγόριθμοι αναζήτησης καθορίζουν τι πρέπει να προβάλλεται με βάση την είσοδο του χρήστη. Ωστόσο, οι περισσότερες πλατφόρμες καθιστούν πλέον καθήκον τους να γνωρίζουν πολύ περισσότερα για τον χρήστη από το ερώτημα που μόλις εισήγαγαν. Σχετικά με τη συλλογή δεδομένων και τις πρακτικές παρακολούθησης των σύγχρονων παρόχων πληροφοριών έχει επικεντρωθεί το ενδιαφέρον στις σημαντικές ανησυχίες περι απορρήτου που προκαλούν. Ο Zimmer [123] υποστηρίζει ότι οι μηχανές αναζήτησης φιλοδοξούν τώρα όχι μόνο να δημιουργούν ευρετήρια στο Διαδίκτυο, αλλά και να αναπτύσσουν “τέλεια ανάκληση” των χρηστών τους. Οι πάροχοι αλγορίθμων ενδιαφέρονται πολύ για τη συλλογή όσο το δυνατόν περισσότερων δεδομένων σχετικά με τους χρήστες και να “εκπαιδεύσουν” τους αλγόριθμους για καλύτερα αποτελέσματα. Ένα παράδειγμα αυτού είναι το πανταχού παρόν κουμπί “like” του Facebook, το οποίο εμφανίζεται σε φαινομενικά άσχετα άρθρα, βίντεο κ.λπ. Οι πάροχοι εκμεταλλεύονται επίσης το όλο και περισσότερο ότι οι χρήστες ενθαρρύνονται να προσφέρουν εθελοντικά κάθε είδους πληροφορίες για τον εαυτό τους. Οι πιο γνωστές πληροφορίες που γνωρίζει το Facebook για τους χρήστες του (γεωγραφική τοποθεσία, πληροφορίες προφίλ, φίλοι, ενημερώσεις κατάστασης, σύνδεσμοι που ακολουθούνται στον ιστότοπο, χρόνος στον ιστότοπο, δραστηριότητα σε άλλους ιστότοπους που φιλοξενούν κουμπιά ή cookies “likes”) είναι ουσιαστικά ένας “ψηφιακός φάκελος” [108] ή “αλγοριθμική ταυτότητα” [15] που να μεν είναι ατελής αλλά συνέχεια συμπληρώνεται και γίνεται επαρκής. [43].

Ένα δεύτερο μέρος που καλύπτεται από αυτήν τη διάσταση είναι ότι οι αλγόριθμοι εστιάζουν απαραίτητα σε αυτό που είναι πιο εύκολα ευανάγνωστο σε αυτούς. Το αλγοριθμικό προφίλ κρίνεται από την πλευρά του παρόχου - από την πλευρά των διαφημίσεων, όπου τα υπόλοιπα προσεγγίζονται ή αγνοούνται. Μια επικίνδυνη πλευρά είναι ότι το αποδεκτό επίπεδο προσέγγισης αρχίζει επίσης να διαμορφώνει το προφίλ του χρήστη, μέσα από κάποιο μενού ζητώντας από τους χρήστες να καθορίσουν σε ποια κατηγορία πιστεύουν ότι εμπίπτουν.

3. Η αξιολόγηση της συνάφειας:

Όταν οι χρήστες επιλέγουν "Αναζήτηση" στο Google, ή ανανεώνουν τη ροή ειδήσεων στο Facebook ή επιλέγουν από τα "προτεινόμενα" του Netflix, οι αλγόριθμοι πρέπει να εντοπίζουν άμεσα και αυτόματα ποια από τα τρισεκατομμύρια πληροφοριών πληρεί καλύτερα τα κριτήρια και θα ικανοποιήσει καλύτερα έναν συγκεκριμένο χρήστη.

Σύμφωνα με την Google, οι αλγόριθμοι αναζήτησης της εξετάζουν περισσότερα από 200 κριτήρια για κάθε ερώτημα που τίθεται. Αυτά τα σήματα είναι τα μέσα με τα οποία ο αλγόριθμος προσεγγίζει την "συνάφεια". Η συνάφεια ενός περιεχομένου δεν είναι κάτι σταθερό ή αντικειμενικό, και ως εκ τούτου είναι ανοιχτή σε κριτική. Δεδομένου ότι δεν υπάρχει ανεξάρτητη μέτρηση για το ποια είναι πραγματικά τα πιο σχετικά αποτελέσματα αναζήτησης για οποιοδήποτε δεδομένο ερώτημα, οι μηχανικοί πρέπει να αποφασίσουν ποια αποτελέσματα φαίνονται "σωστά" και να τροποποιήσουν τον αλγόριθμο τους για να πετύχουν αυτό το αποτέλεσμα ή να κάνουν αλλαγές με βάση το προφίλ των χρηστών, ώστε να ανταπεξέλθουν στα γρήγορα κλικ σε αναζητήσεις όχι σχετικής σημασίας, αλλά ικανοποίησης του χρήστη. Το να κατηγορείται ένας αλγόριθμος για προκατάληψη σημαίνει ότι υπάρχει μια αμερόληπτη κρίση σχετικότητας.

Αυτό που είναι ασταθές είναι ο τρόπος μέτρησης αυτών των κριτηρίων, ο τρόπος με τον οποίο σταθμίζονται το ένα το άλλο, ποια άλλα κριτήρια έχουν επίσης ενσωματωθεί και πότε, αν ποτέ, αυτά τα κριτήρια θα παρακαμφθούν. Ένας πάροχος πληροφοριών όπως το Twitter δεν μπορεί να είναι πολύ πιο σαφής ή ακριβής σχετικά με τη λειτουργία του αλγορίθμου. Αυτό γιατί θα έδινε στους ανταγωνιστές έναν εύκολο τρόπο αναπαραγωγής και υπέρβασης της υπηρεσίας τους. Θα απαιτούσε επίσης μια πιο τεχνική εξήγηση κάτι που οι περισσότεροι χρήστες δεν είναι προετοιμασμένοι.

Τέλος, πρέπει να εξετάσουμε εάν τα κριτήρια αξιολόγησης του αλγορίθμου είναι δομημένα από συγκεκριμένες πολιτικές ή οργανωτικές αρχές που οι ίδιες έχουν πολιτικές επιπτώσεις. Όταν μια "πρωτόγονη" μηχανή αναζήτησης μέτρησε τον αριθμό των εμφανίσεων ενός όρου αναζήτησης στις ιστοσελίδες που είχε προσθέσει στο ευρετήριο, επανέλαβε μια συγκεκριμένη λογική, που υποθέτει ότι οι σελίδες που περιλαμβάνουν τον ερωτηθέντα όρο ήταν πιθανό να σχετίζονται με κάποιον που ενδιαφέρεται για αυτόν τον όρο. Όταν η Google ανέπτυξε το PageRank, συνυπολογίζοντας τους εισερχόμενους συνδέσμους προς μια σελίδα ως απόδειξη της αξίας του, δημιούργησε μια διαφορετική λογική: μια σελίδα με πολλούς εισερχόμενους συνδέσμους, από ιστότοπους υψηλής ποιότητας, θεωρείται "επικυρωμένη" από άλλους χρήστες και είναι πιο πιθανό να είναι συναφής και με αυτόν τον χρήστη. Προτιμώντας τους εισερχόμενους συνδέσμους από τους ίδιους τους ιστότοπους που θεωρούνται υψηλής ποιότητας, σημειώνει ο Finkelstein, ότι η

Google έχει μετατοπιστεί από μια πιο λαϊκιστική προσέγγιση σε μια “δημοκρατία των συμμετεχόντων: Ένας σύνδεσμος δεν είναι μία ψήφος, αλλά έχει επιρροή ανάλογη με τη σχετική δύναμη (όσον αφορά τη δημοτικότητα) του ψηφοφόρου. Επειδή μπλοκ κοινών ενδιαφερόντων, ή κοινωνικές φατρίες, μπορούν να επηρεάσουν τα αποτελέσματα μιας αναζήτησης σε κάποιο βαθμό ανάλογα με το σχετικό βάρος τους στο δίκτυο, τα αποτελέσματα του αλγοριθμικού υπολογισμού από μια μηχανή αναζήτησης έρχονται να αντικατοπτρίζει τους πολιτικούς αγώνες στην κοινωνία” [35].

Οι περισσότεροι χρήστες τείνουν να τα αντιμετωπίζουν ως μη προβληματικά εργαλεία για την εύρεση μιας απάντησης, επίλυση ενός προβλήματος, ψυχαγωγία. Ωστόσο, παρόλο που η τεχνολογία αυτή μπορεί να είναι “ένα μαύρο κουτί” [73] από σχεδιαστές και χρήστες, αυτό δεν πρέπει να μας οδηγήσει να πιστέψουμε ότι παραμένει σταθερή. Στην πραγματικότητα, οι αλγόριθμοι μπορούν να αλλάξουν εύκολα, άμεσα, ριζικά και αόρατα. Ενώ οι σημαντικές αναβαθμίσεις ενδέχεται να πραγματοποιούνται μόνο περιστασιακά, οι αλλαγές σε αλγορίθμους “τροποποιούνται τακτικά” [87]. Οι μηχανές αναζήτησης όπως το Google συμμετέχουν τακτικά σε δοκιμές “A / B” [17], παρουσιάζοντας διαφορετικές κατατάξεις σε διαφορετικά υποσύνολα χρηστών για να αποκτήσουν on-the-fly δεδομένα σχετικά με την ταχύτητα και την ικανοποίηση των πελατών,

4. Υπόσχεση αλγοριθμικής αντικειμενικότητας:

Πάνω από απλά εργαλεία, οι αλγόριθμοι είναι επίσης σταθεροποιητές εμπιστοσύνης, πρακτικές και συμβολικές εγγυήσεις ότι οι αξιολογήσεις τους είναι δίκαιες και ακριβείς, απαλλαγμένες από υποκειμενικότητα, λάθη ή απόπειρα επιρροής. Στην πραγματικότητα, καμία υπηρεσία πληροφοριών δεν μπορεί να είναι εντελώς απρόσκοπτη κατά την παράδοση των πληροφοριών. Ωστόσο, είναι πολύ σημαντικό για τους παρόχους αυτών των αλγορίθμων να φαίνονται ότι είναι αμερόληπτοι.

Η Google έχοντας καταστήσει ως θεμελιώδη στόχο της την αλγοριθμική αντικειμενικότητα και την ουδετερότητα των αλγορίθμων αναζήτησης, απατώντας τακτικά σε αιτήματα για αλλαγή των αποτελεσμάτων αναζήτησης με τον ισχυρισμό ότι ο αλγόριθμος δεν πρέπει να αλλοιωθεί. Η Google αποχώρησε από την Κίνα το 2010 διότι αρνήθηκε να λογοκρίνει τα αποτελέσματά της, παρόλο που στο παρελθόν είχε συμμορφωθεί με τους κανόνες της Κίνας. Αρκετοί υποστηρίζουν ότι αποσύρθηκε αντί να παραδεχτεί ότι έχανε από τους κινέζους ανταγωνιστές της. Αντίστοιχα, άλλαξε τα αποτελέσματα της αναζήτησής της όταν προέκυψαν παράπονα σχετικά με μια εικόνα της Michelle Obama στην κορυφή των αποτελεσμάτων αναζήτησης εικόνων και κατηγορήθηκε για ρατσιστικά αποτελέσματα. Παρόλα αυτά η Google ισχυρίζεται τακτικά ότι δεν αλλοιώνει το ευρετήριό της ούτε ότι χειρίζεται τα αποτελέσματά της.

5. Δέσμευση με πρακτικές:

Αν και θα μπορούσαν να μελετηθούν ως αφηρημένα υπολογιστικά εργαλεία, οι αλγόριθμοι έχουν δημιουργηθεί για να ενσωματωθούν στην πράξη στον πραγματικό κόσμο που παράγει τις πληροφορίες που επεξεργάζονται [21].

Οι χρήστες αναδιαμορφώνουν τις πρακτικές τους ώστε να ταιριάζουν με τους αλγόριθμους τους οποίους αντίστοιχα χρησιμοποιούν. Άλλες πρακτικές θα εφαρμόσουν στο instagram ή στο twitter χρησιμοποιώντας hashtags όχι μόνο για να συμμετέχουμε σε μια συνομιλία, αλλά ελπίζοντας να διαβαστούν από περισσότερους χρήστες. Αντίστροφα, για υλικό (ταινίες, μουσική) που προστατεύονται από πνευματικά δικαιώματα προσπαθούμε να ‘καλύψουμε’ τα ίχνη μας.

Αλλά ποιος είναι στην καλύτερη θέση για να κατανοήσει και να χειριστεί τους δημόσιους αλγόριθμους; Όπως αναφέρθηκε προηγουμένως, τα κριτήρια και ο κώδικας των αλγορίθμων γενικά αποκρύπτονται - αλλά όχι εξίσου ή από όλους. Για τους περισσότερους χρήστες, η κατανόηση αυτών των αλγορίθμων μπορεί να είναι ασαφής, απλοϊκή, μερικές φορές λανθασμένη. Πολλές φορές παρανοούν τα κριτήρια του αλγορίθμου όπως με την επανειλημμένη επανάληψη του ίδιου μηνύματος με την ελπίδα του Trending στο Twitter. Πλέον αρκετός κόσμος τον απασχολεί να αποκαλύψει τη λειτουργία των συστημάτων που χρησιμοποιούν. Οι νομοθέτες όταν άρχισαν να κάνουν ερωτήσεις σχετικά με τις συνέπειες των αλγορίθμων για το εμπορικό δίκαιο ή στον πολιτικό λόγο, μέχρι στιγμής τους έχουν δοθεί μόνο γενικές εξηγήσεις, αφού οι πάροχοι πληροφοριών συχνά υποστηρίζουν ότι οι αλγόριθμοί τους είναι εμπορικά μυστικά που δεν πρέπει να αποκαλυφθούν.

Ο Silverstone [92] πρότεινε ότι όταν οι τεχνολογίες προσφέρονται στο κοινό, υποβάλλονται σε μια διαδικασία «εξημέρωσης», αυτές οι τεχνολογίες εισέρχονται στο σπίτι, αλλά οι χρήστες τις κάνουν δικές τους έστω και εικονικά, ενσωματώνοντάς τις στον τρόπο ζωής τους. Είναι ουσιαστικά πιο δύσκολο να τεκμηριωθούν, πώς οι χρήστες μπορούν να μετατοπίσουν τις κοσμοθεωρίες τους για να προσαρμόσουν τις υποκείμενες λογικές και υπονοούμενες υποθέσεις των αλγορίθμων που χρησιμοποιούν τακτικά. Οι λογικές λειτουργίας αυτών των αλγορίθμων όχι μόνο διαμορφώνουν τις πρακτικές των χρηστών, αλλά οδηγούν τους χρήστες να εσωτερικεύσουν τους κανόνες και τις προτεραιότητές τους. Ο Bucher [13] υποστηρίζει ότι ο αλγόριθμος EdgeRank, που χρησιμοποιείται από το Facebook για να προσδιορίσει ποιες ενημερώσεις κατάστασης παίρνουν προτεραιότητα ώστε να εμφανίζονται στη ροή ειδήσεων των χρηστών, ενθαρρύνει μια “συμμετοχική υποκειμενικότητα” σε χρήστες, οι οποίοι αναγνωρίζουν ότι οι χειρονομίες συγγένειας (όπως σχολιασμός μιας φωτογραφίας φίλου) είναι βασικά κριτήρια στον αλγόριθμο του Facebook. Ο Longford (2005) [75] υποστηρίζει ότι ο κώδικας της εμπορικής πλατφόρμας μας “συνηθίζει”, μέσω προσεκτικά σχεδιασμένων προεπιλεγμένων ρυθμίσεων, σε συνδυασμό με τα προσωπικά μας στοιχεία και το προφίλ μας. Οι

Mager [77] και van Couvering [110] υποστηρίζουν και οι δύο ότι οι αρχές του καπιταλισμού είναι ενσωματωμένες στη λειτουργία των μηχανών αναζήτησης.

6. Παραγωγή υπολογισμένης πολιτικής:

Η ικανότητα εξατομίκευσης των αποτελεσμάτων αναζήτησης και των ειδήσεων στο Internet ήταν η πρώτη και ίσως καλύτερα διατυπωμένη από τις ανησυχίες για τον τρόπο που πλοηγούμαστε στις πληροφορίες και αν υπονομεύονται άλλες. Τα αποτελέσματα που έχουν δύο χρήστες στο ίδιο ερώτημα μπορεί να είναι αρκετά διαφορετικά. Σε κάποιο κοινωνικό δίκτυο, τα αποτελέσματα πληροφοριών μπορούν να προσαρμοστούν με ακρίβεια στις προτιμήσεις του χρήστη όπου τελικά οι ιστορίες που παρουσιάζονται να είναι τόσο διαφορετικές από χρήστη σε χρήστη. Οι Sunstein [111] και, πιο πρόσφατα, ο Pariser [102] ισχυρίστηκαν ότι, όταν οι αλγοριθμικές υπηρεσίες πληροφοριών μπορούν να εξατομικευτούν σε αυτόν τον βαθμό, μπορεί να υπονομευθεί η ποικιλομορφία της δημόσιας γνώσης και του πολιτικού διαλόγου. Μας οδηγούν, με αλγόριθμους και τις δικές μας προτιμήσεις, σε "φυσάλιδες φίλτρου", όπου βρίσκουμε μόνο τις ειδήσεις που περιμένουμε και τις πολιτικές προτιμήσεις.

Όμως οι αλγόριθμοι όχι μόνο δομούν τις αλληλεπιδράσεις μας με άλλους ως μέλη των κοινωνικών δικτύων. Όταν η Amazon προτείνει ένα βιβλίο που αγόρασαν "πελάτες σαν κι εσένα", καλεί και ισχυρίζεται ότι γνωρίζει ένα κοινό με το οποίο μπορούμε να αισθανθούμε οικεία - αν και ο πληθυσμός στον οποίο βασίζει αυτές τις προτάσεις δεν είναι συναφείς με ολόκληρη την πελατειακή της βάση. Όταν το Facebook προσφέρει ως ρύθμιση απορρήτου ότι οι πληροφορίες πρέπει να εμφανίζονται από "φίλους και φίλους φίλων", μετατρέπει ένα ξεχωριστό σύνολο χρηστών σε κοινό - είναι μια ομάδα που δεν υπήρχε μέχρι εκείνη τη στιγμή και μόνο το Facebook γνωρίζει ακριβώς τα μέλη. Αυτές οι αλγοριθμικά παραγόμενες ομάδες μπορεί να αλληλεπικαλύπτονται, να είναι μια ανακριβής προσέγγιση ή να μην έχουν καμία απολύτως σχέση με τις δημοσιεύσεις που ο χρήστης αναζήτησε.

Οι αλγόριθμοι που σκοπεύουν να προσδιορίσουν τι είναι "καυτό" εμπλέκονται σε μια υπολογισμένη προσέγγιση του κοινού μέσω της ανιχνεύσιμης δραστηριότητάς του και, στη συνέχεια, τους αναφέρουν τι έχει συζητηθεί περισσότερο. Αλλά πίσω από αυτό, δημιουργείται το ερώτημα, ποιο είναι το κέρδος για τους παρόχους να κάνουν κάτι τέτοιο και πώς διαμορφώνει αυτό που ψάχνουν; Ποιος επιλέγεται να μετρηθεί και ποιος μένει εκτός υπολογισμού; Και ίσως το πιο σημαντικό, πώς αυτές οι τεχνολογίες συμβάλλουν στη σύσταση και την κωδικοποίηση του κοινού που ισχυρίζονται ότι μετράει. Αυτές οι ερωτήσεις έχουν μεγάλη σημασία, και θα έχουν μεγαλύτερη σημασία, στο βαθμό που αυτά που παράγονται από αλγόριθμους πληροφοριών λαμβάνονται από τους χρήστες ή από τις αρχές, ως νόμιμες, και ενσωματώνονται στο ευρύτερο μοντερνιστικό σχέδιο αντανάκλαστικότητας [42]. "Η κοινωνία ασχολείται να παρακολουθεί τον εαυτό της, να εξετάζει τον εαυτό της, να απεικονίζεται με διάφορους τρόπους, και να τροφοδοτεί τις προκύπτουσες αντιλήψεις για να οργανώσει τις δραστηριότητές της" [9]. Αυτό που ισχυρίζεται το Twitter έχει σημασία

για τους “Αμερικανούς” ή οι προτάσεις της Amazon βιβλίων για έφηβους είναι μορφές έγκυρης γνώσης που μπορούν και θα επικαλεστούν τα όσους έχουν ως στόχο τη ρύθμιση τέτοιων πληθυσμών.

Η πεποίθηση ότι αυτοί οι αλγόριθμοι, σε συνδυασμό με τεράστια δεδομένα χρηστών, είναι καλύτερα να μας πουν πράγματα για τη φύση του κοινού ή τη κοινωνία, έχει αποδειχθεί δελεαστικό και για τους μελετητές. Η κοινωνική επιστήμη στράφηκε με ανυπομονησία σε υπολογιστικές τεχνικές, ή στη μελέτη της ανθρώπινης κοινωνικότητας μέσω “big data” [66] με την ελπίδα να απολαύσει το είδος των πληροφοριών που οι βιολογικές επιστήμες έχουν επιτύχει. Η προσέγγιση είναι σαγηνευτική: η ύπαρξη εκατομμυρίων δεδομένων παρέχει μεγάλη νομιμότητα και ο τρόπος με τον οποίο οι αλγόριθμοι φαίνεται να εντοπίζουν μοτίβα που οι ερευνητές δεν μπορούσαν να δουν διαφορετικά είναι συναρπαστικό. Αλλά αυτή η μεθοδολογική προσέγγιση πρέπει να προσέχει τις επιπλοκές που έχουν περιγραφεί μέχρι στιγμής, ιδιαίτερα όταν τα δεδομένα τους δημιουργούνται από τους ίδιους τους εμπορικούς αλγόριθμους. Οι υπολογιστικές τεχνικές έρευνας δεν είναι βαρόμετρα της κοινωνίας.

Οι σχετικές αντισταθμίσεις αυτού του είδους γενικά δεν ενδιαφέρουν τους παρόχους και τους σχεδιαστές τέτοιων σημαντικών αλγορίθμων, και ως εκ τούτου πρέπει να επικριθούν από έξω από τη δική της κοινωνικοτεχνική κουλτούρα ή το δικό τους σύστημα. Ο Gillespie καταλήγει συνεπώς στο συμπέρασμα ότι “μπορεί να δούμε αλγόριθμους όχι μόνο ως κώδικες με συνέπειες, αλλά ως τον πιο πρόσφατα κοινωνικά κατασκευασμένο και θεσμικά διαχειριζόμενο μηχανισμό για τη διασφάλιση του δημόσιου πνεύματος” [30].

3.4 Οι Περιπτώσεις των Ηθικών Πτυχών των Αλγορίθμων

Στα παραδείγματα θα δούμε νέες δυναμικές που χρειάζονται περαιτέρω έρευνα ώστε να δούμε τι σημαίνει από ηθικής πλευράς. Σύμφωνα με τον Sandvig [91] η περαιτέρω έρευνα και τα ερωτήματα που θα προκύψουν θα είναι “για ένα συγκεκριμένο είδος διαδικασίας ή στρατηγικής και όχι για τον ίδιο τον στόχο”. Το επίκεντρο θα είναι πρώτα στις εμπορικές εφαρμογές αλγορίθμων, και στην συνέχεια δημόσιες και κυβερνητικές εφαρμογές, αλλά και τα δυο σε συνδυασμό..

3.4.1 Συνάφεια: Μηχανές Αναζήτησης

Σε αυτό το σύνολο παραδειγμάτων θα δοθεί έμφαση στον τρόπο με τον οποίο οι μηχανές αναζήτησης και τα κοινωνικά δίκτυα έχουν φτάσει να “κυβερνούν” τις πληροφορίες στο Διαδίκτυο και να εκτελούν έναν τύπο “κατασκευής πραγματικότητας” [58]. Για να γίνει αυτό σωστά, η εστίαση θα δοθεί στα δύο μεγαλύτερα παραδείγματα κάθε κατηγορίας, το Google

και το Facebook. Πέρα από το ότι είναι οι μεγαλύτεροι στην κατηγορία τους, είναι επίσης, σύμφωνα με τον Just και τον Latzer [58], οι μεγαλύτεροι δύο ιστότοποι που χρησιμοποιούν αλγοριθμική επιλογή. Σε αυτή την ενότητα, θα δώσουν πληροφορίες σχετικά με τη λειτουργία αρκετά εμπορικών αλγοριθμικών εφαρμογών.

Οι μηχανές αναζήτησης αποτελούν πύλη για πληροφορίες στο Διαδίκτυο. Δεδομένης της σπουδαιότητος τους, δεν αποτελεί έκπληξη το γεγονός ότι το έργο σχετικά με τις επιπτώσεις των μηχανών αναζήτησης γενικά, πέρα από τα τεχνικά επιτεύγματά τους, έχει μια ιστορία που ξεκινάει από τα τέλη της δεκαετίας του '90 σε ένα έργο των Introna και Nissenbaum [97]. Στο έργο τους αυτό, διερευνούν, μεταξύ άλλων, τα πολιτικά και κοινωνικά στοιχεία που εμπλέκονται στον τρόπο με τον οποίο οι αλγόριθμοι που χρησιμοποιούνται από τις μηχανές αναζήτησης δημιουργώντας “μια πολύπλοκη ανθρώπινη αξία” [97]. Ένα από τα σημαντικότερα συμπεράσματα είναι ότι οι μηχανές αναζήτησης αποκλείουν συστηματικά ορισμένους τύπους ιστότοπων και πληροφοριών, με επιπτώσεις τόσο για τους παρόχους περιεχομένου όσο και για τους χρήστες περιεχομένου. Ενώ στα τέλη της δεκαετίας του 1990 υπήρχε ακόμη πληθώρα μηχανών αναζήτησης, αυτό το πλήθος επιλογών αντικαταστάθηκε σταδιακά από τη Google, αντιπροσωπεύοντας πλέον περίπου το 72% όλων των αναζητήσεων σε επιτραπέζιους υπολογιστές και το 95% των αναζητήσεων για κινητά παγκοσμίως - συμπεριλαμβανομένης της επισκεψιμότητας από την Κίνα, όπου η Google ήταν αποκλεισμένη για ένα διάστημα, αλλά πλέον κάνει προσπάθειες για να επιστρέψει στην αγορά της Κίνας [84]. Επομένως, ενώ άλλες μηχανές αναζήτησης ενδέχεται να διαφέρουν ως προς την προσέγγισή τους. Με απλοποιημένο τρόπο, η Google χρησιμοποιεί ένα ερώτημα αναζήτησης ως εισαγωγή, στο οποίο οι αλγόριθμοι χρησιμοποιούν πάνω από 200 παράγοντες για τον προσδιορισμό των αποτελεσμάτων και τη σειρά αυτών. Η ακριβής εξισορρόπηση των παραγόντων διαφέρει ανά είσοδο και ανά χρήστη, αν και οι ακριβείς μηχανισμοί πίσω από αυτό προστατεύονται από “συμφωνίες μη αποκάλυψης” [94]. Στην πραγματικότητα, η απόπειρα καταγραφής των αλγορίθμων με ακρίβεια θα αποδείξει ότι είναι μια μάταιη προσπάθεια, όπως ο Sandvig, Hamilton και Karahalios [91] εξηγούν, ο επαναπρογραμματισμός των αλγορίθμων χρησιμοποιείται για να κάνει “ένα βασικό χαρακτηριστικό του συστήματός τους αρθρωτό και υπόκειται σε συνεχή αναθεώρηση”.

Ένα διάσημο παράδειγμα είναι το έργο του Pariser [86], που ανέλυσε τις μηχανές αναζήτησης, και ειδικότερα τη Google, διαπιστώνοντας ότι παρουσιάζει εξατομικευμένα αποτελέσματα αναζήτησης στους χρήστες του, με βάση τις γνώσεις τους για αυτόν τον χρήστη. Δηλαδή, μέσω της χρήσης των υπηρεσιών τους - στην περίπτωση της Google που συχνά περιλαμβάνει κάτι παραπάνω από τη μηχανή αναζήτησης, π.χ. Gmail, Youtube - είναι σε θέση να δημιουργήσουν ένα προφίλ, το οποίο χρησιμοποιεί ένα σύνολο αλγορίθμων για να παρουσιάσει συγκεκριμένα αποτελέσματα αναζήτησης, ενδεχομένως οδηγώντας σε μεγάλες διαφορές μεταξύ των χρηστών. Ένα σημαντικό στοιχείο σε αυτό είναι η λειτουργία της μεροληψίας για επιβεβαίωση [86][91], όπου οι άνθρωποι αισθάνονται καλά για να επιβεβαιωθούν σε αυτά που ήδη γνωρίζουν ή σκέφτονται.

Μέσω αυτού, είναι θεωρητικό ότι οι άνθρωποι:

- Πρώτον, αναζητούν πληροφορίες που επιβεβαιώνουν τις απόψεις τους,
- Δεύτερον, βιώνουν ευχαρίστηση όταν οι μηχανές αναζήτησης προσφέρουν τέτοια αποτελέσματα, και
- Τρίτον, αφιερώνουν περισσότερο “ευχάριστο” χρόνο χρησιμοποιώντας αυτές τις μηχανές αναζήτησης και αυτο οδηγεί τις εταιρείες αυτές σε υψηλότερα έσοδα λόγω διαφημιστικής προβολής και κερδών.

Στη συνέχεια, τα εμπορικά οφέλη είναι ξεκάθαρα, καθώς “Οι προτεινόμενοι μπορούν να έχουν θετική επίδραση στις πωλήσεις και στις εμφανίσεις ιστού” [94], και ως εκ τούτου θα χρησιμοποιηθεί σε ανταγωνιστικές αγορές, χωρίς τους κανονιστικούς ελέγχους. Αυτοί οι μηχανισμοί, επομένως, με περιορισμένο τρόπο, εξυπηρετούν τόσο τον χρήστη όσο και την εταιρεία της μηχανής αναζήτησης.

Ένα άλλο παράδειγμα του τρόπου με τον οποίο η Google προσαρμόζει τα αποτελέσματά της, είναι η περιοχή του Κασμίρ που βρίσκεται σε διένεξη, η οποία φαίνεται να είναι μέρος του Πακιστάν ή της Ινδίας, ανάλογα με την τοποθεσία όποιου κάνει την αναζήτηση [40]. Αν και αυτό μπορεί να θεωρηθεί ως απόπειρα να είναι “πολιτικά σωστό” ή ουδέτερο, θα πρέπει να σημειωθεί ότι δεν υπάρχει τέτοια επιλογή, και η παροχή διαφορετικών πληροφοριών στους ανθρώπους ανάλογα με τη θέση τους επηρεάζει αναπόφευκτα την κοσμοθεωρία τους.

Ενώ ορισμένα από τα συμπεράσματα του βιβλίου του Pariser “The Filter Bubble” έχουν αποχρωματιστεί μετά από έρευνα, σε ένα άρθρο του 2015 των Epstein και Robertson [52] όπου λαμβάνουν πολύ σοβαρά την επίδραση των αλγοριθμικά ελεγχόμενων αποτελεσμάτων αναζήτησης και προειδοποιεί συγκεκριμένα για την πιθανή επιρροή των κατατάξεων αναζήτησης της Google στο διαδίκτυο για το αποτέλεσμα των προεδρικών εκλογών του 2016 στις ΗΠΑ. Λόγω του αποτελέσματος χειραγώγησης της μηχανής αναζήτησης (SEME), το οποίο έχει αποδειχθεί ότι μεταβάλλει σοβαρά την άποψη των ανθρώπων χωρίς να είναι παρατηρήσιμο, η έκθεση συγκεκριμένου υλικού θα μπορούσε να επηρεάσει σημαντικά την κοινή γνώμη [52]:

1. Οι προκατειλημμένες κατατάξεις αναζήτησης μπορούν να αλλάξουν τις προτιμήσεις ψήφου των αναποφάσιστων ψηφοφόρων κατά 20% ή περισσότερο
2. Η αλλαγή μπορεί να είναι πολύ υψηλότερη σε ορισμένες δημογραφικές ομάδες
3. Τέτοιες κατατάξεις μπορούν να καλυφθούν, έτσι ώστε οι άνθρωποι να μην γνωρίζουν τη χειραγώγηση.

Ο κίνδυνος επιδεινώνεται από το γεγονός ότι πολλοί άνθρωποι δεν γνωρίζουν [55][52] το γεγονός ότι οι μηχανές αναζήτησης, των οποίων η Google χρησιμεύει εδώ ως πρωταρχικό παράδειγμα, στην πραγματικότητα λειτουργούν αλγοριθμικά και μέσω αυτού αναπτύσσουν ένα “προφίλ” που επιλέγει αποτελέσματα αναζήτησης ανάλογα με τη συνάφεια για

μεμονωμένους χρήστες. Ο κίνδυνος σε αυτό, λοιπόν, είναι ότι δεν υπάρχει πλέον ουδέτερη αναζήτηση: ανεξάρτητα από τη διατύπωση, τα αποτελέσματα θα προκαταλάβουν ως συνέπεια της αλγοριθμικής δομής. Επιπλέον, όπως διαπίστωσαν οι Epstein και Robertson [30], ανησυχεί το γεγονός ότι όταν οι άνθρωποι δεν γνωρίζουν τα αποτελέσματα της χειραγώγησης, και υπάρχει μια τάση να πιστεύουν ότι έχουν διαμορφώσει τη γνώμη τους ανεξάρτητα. Αυτές οι δυναμικές είναι επίσης σημαντικές για τους παρόχους περιεχομένου, οι οποίοι προσπαθούν να μεγιστοποιήσουν την κυκλοφορία και την έκθεση του περιεχομένου τους. Προκειμένου να επιτευχθεί αυτό, είναι σημαντικό το περιεχόμενο να οδηγείται και να θεωρείται σχετικό με μια σημαντική ομάδα ατόμων σύμφωνα με τους αλγόριθμους της Google. Αυτό, στην πραγματικότητα, θέτει περιορισμούς στον τύπο του περιεχομένου που θα δημιουργηθεί και στον τρόπο με τον οποίο διατίθεται στην αγορά.

Εν ολίγοις, οι συνέπειες της συγκεκριμένης αλγοριθμικής προσέγγισης των μηχανών αναζήτησης και ειδικότερα του Google είναι ότι οι χρήστες παρουσιάζονται ένα σύνολο προεπιλεγμένων αποτελεσμάτων, δεν γνωρίζουν αυτήν την επιλογή και με τη σειρά τους βασίζουν τις γνώσεις και τη συμπεριφορά τους σε αυτήν τη δυναμική, με ενισχυμένη την προϋπάρχουσα άποψη με τη μορφή προκατάληψης και επιβεβαίωσης καθιστώντας πιο δύσκολο να ξεφύγουν από αυτό. Για τους παρόχους περιεχομένου παραμένει ασαφές πώς λειτουργούν πραγματικά οι διαδικασίες [87]. Αυτό αφήνει το ζήτημα της αξιολόγησης ακόμη πιο προβληματικό, καθώς χωρίς να γνωρίζουμε τι πραγματικά κάνει η Google όταν κατατάσσει τους ιστότοπους, δεν μπορούμε να εκτιμήσουμε πότε ενεργεί με καλή πίστη για να βοηθήσει τους χρήστες και πότε προκαλούνται προκαταλήψεις για να ευνοήσει τα δικά της εμπορικά συμφέροντα.

3.4.2 Κοινωνικά Δίκτυα

Παρόμοια με τις μηχανές αναζήτησης, τα κοινωνικά δίκτυα, των οποίων το Facebook χρησιμεύει ως παράδειγμα εδώ, έχουν ως μέρος της λειτουργικότητάς τους τη δυνατότητα για τους χρήστες και να ενημερώνονται. Ωστόσο, αντί να λειτουργεί με ερωτήματα αναζήτησης, λειτουργεί με κοινωνικές συνδέσεις και "συνδρομές" (subscriptions), οι οποίες απαιτούν εθελοντικές και συνειδητές πράξεις έκφρασης ενδιαφέροντος από το χρήστη προς άλλο χρήστη ή σελίδα. Αυτό αποδείχθηκε ένα πολύ επιτυχημένο μοντέλο που εκτείνεται πέρα από τον αρχικό στόχο της κοινωνικής επαφής. Μελέτες έχουν δείξει ότι, ήδη από το 2014, το Facebook ήταν πηγή πολιτικών ειδήσεων για το 48% των Αμερικανών [88], ενώ το 61% των νέων που είναι γεννημένα κοντά στο 2000 λαμβάνουν τα νέα τους από το λεγόμενο News Feed [25], και πλέον η επισκεψιμότητα σε σελίδες από κορυφαίους εκδότες είναι μεγαλύτερη μέσω του Facebook απ' ό,τι μέσω της Google, για πρώτη φορά από τον Ιούλιο του 2015 [89].

Ωστόσο, παρόμοια με την Google, οι αλγόριθμοι είναι υπεύθυνοι για την επιλογή περιεχομένου με βάση τη συνάφεια, εξατομικεύονται με τρόπους που δεν βασίζονται μόνο στον εαυτό τους (δηλαδή μια θετική ανατροφοδότηση, όπου οι άνθρωποι ενισχύονται στις πεποιθήσεις και τις απόψεις που ήδη έχουν), αλλά είναι επίσης αδιαφανείς στη λειτουργία τους. Μελέτες έχουν δείξει ότι, για παράδειγμα, “το 62% των μαθητών σε ένα σχολείο υψηλού επιπέδου δεν γνώριζαν καν την ύπαρξη αυτού του αλγορίθμου”, αλλά όπως και με τη Google, ότι η λειτουργία του αλγορίθμου είναι εμπορικό μυστικό. Μια βασική διαφορά των social media με τις μηχανές αναζήτησης είναι ότι ένας μεγάλος αριθμός παραγόντων που εμπλέκονται στην επιλογή περιεχομένου βασίζεται στον τύπο και το βάθος της σχέσης με άλλους χρήστες. Αυτή η ιδέα έχει συνδεθεί με έναν “θάλαμο ηχούς” [89] όπου οι συσχετισμοί που σχηματίζονται ερμηνεύονται αλγοριθμικά. Το ίδιο το Facebook έχει πραγματοποιήσει μια μελέτη, που δημοσιεύθηκε στο περιοδικό Science, σχετικά με το μέγεθος των πιθανών καθοριστικών παραγόντων για το τι είδους αποτελέσματα παίρνουν προτεραιότητα στα News Feed [8], το οποίο περιείχε το συμπέρασμα ότι οι κοινωνικές συνδέσεις παίζουν μεγαλύτερο ρόλο στα κριτήρια που θα εφαρμόσουν οι αλγόριθμοι. Ωστόσο, ο Pariser σε απάντηση είπε ότι “Σίγουρα, ποιοι είναι οι φίλοι σας έχουν μεγάλη σημασία στα μέσα κοινωνικής δικτύωσης. Όμως το γεγονός ότι το αποτέλεσμα μείωσης του αλγορίθμου είναι σχεδόν τόσο ισχυρό όσο η αποφυγή απόψεων με τις οποίες διαφωνούμε δείχνει ότι είναι στην πραγματικότητα μια πολύ μεγάλη υπόθεση” [86]. Επιπλέον, φαίνεται διαισθητικό, οι άνθρωποι γνωρίζουν περισσότερο ποιοι είναι οι φίλοι τους και τι πιστεύουν, από την εμφανώς χαμηλή επίγνωση της αλγοριθμικής επιλογής.

Παραδοσιακά, ειδησεογραφικά πρακτορεία, όπως εφημερίδες και τηλεοπτικοί σταθμοί, έχουν σαφώς επιλέξει ποια γεγονότα θα καλύψουν και από ποια γωνία. Δηλαδή, η επιλογή και η μεροληψία δεν είναι ένα νέο φαινόμενο που εισάγεται από αλγόριθμους και οποιαδήποτε κοσμοθεωρία θα διαμορφωθεί από μερικές και επιλεκτικές πληροφορίες. Οι μηχανισμοί που περιγράφονται στις προηγούμενες παραγράφους, ωστόσο, είναι καινοτόμοι, ότι οι διαδικασίες λήψης αποφάσεων είναι αυτοματοποιημένες και η ευθύνη για όσα περιλαμβάνονται ή αποκλείονται, ή έχουν προτεραιότητα κ.λπ. ανατίθενται στον αλγόριθμο.

Ένα άλλο βασικό σημείο που πρέπει να παρατηρήσουμε εδώ είναι ο τρόπος με τον οποίο οι ειδήσεις και οι πληροφορίες, μέσω αλγοριθμικής επιλογής, καταναλώνονται βάσει ζήτησης, με μεμονωμένα άρθρα να εμφανίζονται στο προσκήνιο. Αυτό είναι άλλος ένας τρόπος όπου τα πιο παραδοσιακά μέσα ενημέρωσης, καθώς τα νέα ή οι ειδήσεις από τις εφημερίδες ή ότι μεταδίδεται στην τηλεόραση ή το ραδιόφωνο περιέχει γενικά πληροφορίες για μια σειρά θεμάτων, όπου η διαδικασία επιλογής είναι συγκεντρωτική. Όπως συνοψίζουν οι Just και Latzer, “σε σύγκριση με την κατασκευή πραγματικότητας από τα παραδοσιακά μέσα μαζικής ενημέρωσης, η κατασκευή αλγοριθμικής πραγματικότητας τείνει να αυξάνει την εξατομίκευση (προσαρμογή), την εμπορευματοποίηση, τις ανισότητες και την αποτρεπτικοποίηση και να μειώσει τη διαφάνεια” [58].

3.4.3 Η Αλγοριθμική Ιθαγένεια

Σε αυτή την ενότητα θα εξετάσουμε την “αλγοριθμική ιθαγένεια”, μια ορολογία που επινοήθηκε από τον Cheney-Lippold [15] για να περιγράψει το ρόλο των αλγορίθμων στην οικοδόμηση της ιθαγένειας στις Ηνωμένες Πολιτείες.

Η ιθαγένεια ορίζεται παραδοσιακά με τους όρους:

- *jus sanguinis* που αναφέρεται στην ιθαγένεια που προέρχεται “εξ αίματος”, που σημαίνει ότι γεννιούνται από γονείς που έχουν συγκεκριμένη ιθαγένεια
- *jus soli*, που αναφέρεται στην ιθαγένεια που σχετίζεται από τον τόπο γέννησης.

Επιπλέον, η αναγνώριση των ατόμων παρεμβάλλεται από τεχνολογίες και υποδομές που αποκρύπτουν ταυτότητες. Διευθύνσεις πρωτοκόλλου διαδικτύου (IP) και network hardware identifiers, δύο συνηθισμένοι τρόποι αναγνώρισης συσκευών που έχουν πρόσβαση στο Διαδίκτυο, μπορεί να ανήκουν σε μηχανήματα που λειτουργούν από πολλούς χρήστες ή να είναι ουσιαστικά δημόσια όπως στην περίπτωση υπολογιστών σε βιβλιοθήκες, σχολεία και άλλα μέρη με σχετικά εύκολη πρόσβαση σε ομάδες ανθρώπων. Κατά συνέπεια, όταν η NSA προσπάθησε να συλλέξει όσο το δυνατόν περισσότερα δεδομένα για τους ανθρώπους, διασφαλίζοντας παράλληλα ότι οι πολίτες των ΗΠΑ θα απαλλάσσονταν από την παρακολούθηση - μια νομική απαίτηση των επιχειρήσεων- αντιμετώπισε το πρόβλημα της μη δυνατότητας αναγνώρισης πολιτών των ΗΠΑ. Για να ξεπεράσει αυτήν τη δυσκολία, εισήγαγε έναν νέο τύπο ιθαγένειας, με βάση την αλγοριθμική καταταξη σε μια λίστα, για να προσδιορίσει πόσο ξένος είναι ένας χρήστης. Ο Cheney-Lippold περιγράφει τις απότομες συνέπειες αυτής της βαθμολογίας ως εξής: “Εάν η ξένη κατάσταση ενός ατόμου βρίσκεται σε ποσοστό 51% ή μεγαλύτερο, τότε αυτό το άτομο γίνεται νόμιμα αλλοδαπός και έτσι χάνει το δικαίωμα στην ιδιωτική ζωή” [15]. Ονομάζει αυτή τη νέα κατηγοριοποίηση της ιθαγένειας ως αλγόριθμο ή δικαίωμα υπηκοότητας σύμφωνα με την αλγοριθμική λογική. Αυτή η ιθαγένεια βασίζεται σε τεχνικές απαιτήσεις και δεν μπορεί να διεκδικήσει ή να ασκήσει έφεση. Λόγω των περιορισμών στη διερεύνηση της συγκεκριμένης διαδικασίας με την οποία καθορίζεται το “51%”, δεν είναι δυνατόν να αιτιολογηθεί ή να περιγραφεί πώς επιλύει τα προβλήματα αφάνειας λόγω τεχνικής διαμεσολάβησης όπως περιγράφεται παραπάνω.

Η εικασία του Cheney Lippold είναι ότι υπάρχουν συγκεκριμένες περιπτώσεις κατά τις οποίες ένας αλγόριθμος ανέλυσε τα διαθέσιμα δεδομένα μιας συγκεκριμένης σύνδεσης, το ποσοτικοποιεί και παράγει την έξοδο του εάν αυτό το άτομο είναι, σύμφωνα με την ταξινόμηση *jus algorithmi* πολίτης των ΗΠΑ ή όχι. Αυτό δίνει επίσης τη δυνατότητα για το ίδιο άτομο να μετακινηθεί μεταξύ της νόμιμης ιθαγένειας απλά αλλάζοντας μια συσκευή ή παράγοντας διαφορετικά μετα-δεδομένα, για παράδειγμα εισάγοντας ένα συγκεκριμένο ερώτημα αναζήτησης ή επικοινωνώντας με κάποιον αλγόριθμο που αποφάσισε να είναι *jus algorithmi* πολίτης ή όχι. Ενδεικτικά, τα έγγραφα που ο Edward Snowden αποκάλυψε, έχουν δείξει μεταξύ άλλων ότι, ένα αναγνωριστικό σε αυτόν τον αλγόριθμο είναι η χρήση γλώσσας σε email και τηλεφωνικές κλήσεις. Ως εκ τούτου, η εφαρμογή μιας αλγοριθμικά

“βαθμολογημένης” ιθαγένειας, παράγει αυθαίρετες και διφορούμενες συνέπειες μέσω μιας διαδικασίας που αποκρύπτεται ενεργά και βασίζεται σε μετρήσεις για τις οποίες δεν υπάρχει αποδεδειγμένο δημοκρατικό θεμέλιο.

Εκτός από τεχνικές λεπτομέρειες, οι νομικές πολιτικές και ηθικές συνέπειες αυτών των πρακτικών είναι σοβαρές. Αυτή η αλγοριθμική προσέγγιση της ιθαγένειας υπερισχύει των συνταγματικά προστατευόμενων δικαιωμάτων των πολιτών, ενσωματωμένων σε ένα σύστημα που επιδιώκει ενεργά την αφάνεια, με αποτέλεσμα μια πλήρη ασυμμετρία εξουσίας υπέρ των κυβερνητικών θεσμών. Επομένως, οι πολίτες δεν γνωρίζουν ιδιαίτερα αυτήν την πρακτική για να ξεκινήσουν, οι χρήστες του Διαδικτύου δεν ενημερώνονται ιδίως για τις διαδικασίες, δεν έχουν δικαίωμα να γνωρίζουν ή να ασκήσουν έφεση κατά των συμπερασμάτων, και ουσιαστικά δεν ενημερώνονται για την πρακτική να ξεκινήσει με. Επιπλέον, το συμπέρασμα και κατά συνέπεια αυτού του είδους η ιθαγένεια και η προστασία των συνταγματικά προστατευόμενων δικαιωμάτων, εξαρτώνται από το χρόνο, είναι εύπλαστο και εφήμερο. Για να επισημάνετε το ρόλο των αλγορίθμων σε αυτήν τη διαδικασία, είναι σημαντικό να εξετάσουμε την εγγενώς ποσοτικοποιημένη φύση της διαδικασίας, όπως στην ακόλουθη παράθεση [\[15\]](#):

3.4.4 Η Αναγνώριση Προσώπων

Η αναγνώριση προσώπου είναι η διαδικασία όπου μια εικόνα ή ένα βίντεο υποβάλλονται σε επεξεργασία από αλγόριθμους με τρόπο που τα pixel (είσοδος) αναγνωρίζονται ότι αποτελούν το πρόσωπο ενός ατόμου. Η αναγνώριση προσώπου μέσω αλγορίθμων μπορεί να πραγματοποιηθεί με πολλούς διαφορετικούς τρόπους, με τον καθένα να έχει οφέλη και προκλήσεις όσον αφορά την ακρίβεια και την αποτελεσματικότητα σε ορισμένες συνθήκες [\[91\]](#). Γενικά, τα συστήματα αναγνώρισης προσώπου λειτουργούν σύμφωνα με τα παρακάτω βήματα [\[115\]](#):

1. Ανίχνευση προσώπου: ανίχνευση αντικειμένου και προβλήματα εντοπισμού.
2. Εκτίμηση θέσης: προσδιορισμός βασικών σημείων όπως τα μάτια, η μύτη και η δυνατότητα ενσωμάτωσης αυτών ανάλογα με τις γωνίες και το φωτισμό.
3. Κανονικοποίηση (frontalization / normalization): σε αυτό το βήμα γίνεται η χαρτογράφηση της εικόνας σε μια κανονικοποιημένη εικόνα ενός προσώπου.
4. Εξαγωγή χαρακτηριστικών: μόλις ολοκληρωθεί το προηγούμενο βήμα, παρουσιάζεται ως διανυσματική εικόνα που μπορεί να συγκριθεί με άλλες εικόνες που προκύπτουν.
5. Αναγνώριση: σύγκριση των ληφθέντων δεδομένων εικόνας με άλλα τέτοια δεδομένα στο υπάρχον σύνολο δεδομένων. Παρά τις τεχνικές λεπτομέρειες αυτών των διαφορετικών προσεγγίσεων, η εφαρμογή της αναγνώρισης προσώπου μπορεί να οδηγήσει σε σημαντικά ακούσια αποτελέσματα.

Ένα παράδειγμα αυτού είναι ο αλγόριθμος για την ανίχνευση προσώπου σε κάμερες για υπολογιστές από την Hewlett-Packard αρχικά δεν κατάφερε να αναγνωρίσει και να εντοπίσει τα πρόσωπα των μαύρων ανθρώπων σε κοινές συνθήκες φωτισμού. Αυτό οδήγησε στο άβολο αποτέλεσμα ότι τα πρόσωπά τους δεν επικυρώθηκαν ως τέτοια [91], με τον αλγόριθμο να είναι αυτός που θα το κρίνει. Αυτό το λάθος, λόγω του ότι ήταν άμεσα διαθέσιμο στον εμπορικό τομέα, ανακαλύφθηκε γρήγορα από τους χρήστες και αντιμετωπίστηκε από την εταιρεία μέσω ενημερώσεων λογισμικού. Αυτό που είναι σημαντικό σε αυτήν την περίπτωση, ωστόσο, είναι ότι ακριβώς τα κριτήρια επιλογής που αρχικά πήγαν στο σχεδιασμό του αλγορίθμου και του υλικού της κάμερας δημιούργησαν αυτό το αποτέλεσμα. Ενώ η διόρθωση ήταν δυνατή και εκτελέστηκε γρήγορα, αυτό ήταν δυνατό κυρίως επειδή οι χρήστες επεσήμαναν το λάθος, το οποίο απαιτεί γνώση των αποτελεσμάτων που παράγουν αυτοί οι αλγόριθμοι, μια συγκεκριμένη περίπτωση που συχνά δεν συμβαίνει για μη εμπορικές χρήσεις. Λαμβάνοντας υπόψη τις συνέπειες αυτού του αλγορίθμου θα μας οδηγήσει στο συμπέρασμα ότι σε ορισμένες περιπτώσεις, ο αλγόριθμος θα μπορούσε να παράγει ρατσιστικά αποτελέσματα. Επιπλέον, οι αρνητικές συνέπειες της λειτουργίας του αλγορίθμου παρατηρήθηκαν μόνο μετά την εφαρμογή, υποδηλώνοντας ότι η Hewlett-Packard δεν είχε εξετάσει τη δυνατότητα, ούτε έκανε εκτεταμένες δοκιμές του προϊόντος τους. Ενώ αυτό μπορεί να φαίνεται να είναι ένα απλό ακούσιο αποτέλεσμα του αλγορίθμου και θα μπορούσε να διατυπωθεί με πολύ τεχνικό τρόπο όπου ορισμένες συνθήκες οδηγούν στην αποτυχία του αλγορίθμου να εκτελέσει την επιδιωκόμενη λειτουργία του, μια τέτοια προσέγγιση αγνοεί την κοινωνική πραγματικότητα των συναισθημάτων αποκλεισμού ή ρατσισμού.

Ένα άλλο, πιο πρόσφατο, παράδειγμα αλγοριθμικής επεξεργασίας προσώπων δεν περιλαμβάνει μόνο την ανίχνευση, αλλά και την αναγνώριση προσώπων. Ενώ οι υπηρεσίες πληροφοριών και ασφάλειας χρησιμοποιούν παγκοσμίως αυτές τις τεχνικές σε μεγάλο βαθμό - όπως το Ομοσπονδιακό Γραφείο Ερευνών (FBI) στις Ηνωμένες Πολιτείες - σε μια έρευνα που διενήργησε το Γραφείο Ευθύνης της Κυβέρνησης των Ηνωμένων Πολιτειών (GAO), διαπιστώθηκε ότι τα ποσοστά σφάλματος της αναγνώρισης κυμαίνονταν μεταξύ "λίγο τοις εκατό έως και πενήντα τοις εκατό, ανάλογα με την τεχνολογία". Αυτό οφείλεται σε μεγάλο βαθμό σε ξεπερασμένες και υπο δοκιμασμένες εκδόσεις, οι οποίες θέτουν περιορισμούς στις δυνατότητες επίλυσης του προβλήματος. Το βασικό σημείο, ωστόσο, είναι ότι τα περισσότερα από τα ποσοστά σφάλματος δεν έχουν δοκιμαστεί σωστά πριν από την αγορά και οι επακόλουθες κριτικές περιείχαν τα ίδια ελαττώματα, συμπεριλαμβανομένων ψευδών θετικών. Αυτό δείχνει πώς η λειτουργία του αλγορίθμου που στηρίζει τις τεχνικές αναγνώρισης προσώπου θεωρήθηκε δεδομένη, ενσωματώθηκε σε ένα μεγαλύτερο κοινωνικο-πολιτισμικό σύστημα ως ουδέτερο και ορθολογικό όργανο και έχει διαδραματίσει ενεργό ρόλο στην έρευνα και την πεποίθηση των ανθρώπων, ενώ η ισχύς του παραμένει αμφισβητήσιμη. Αυτή η αλληλεπίδραση της χρήσης ως φαινομενικά ουδέτερου οργάνου, που χρησιμοποιείται για τη σάρωση περισσότερων από 400 εκατομμυρίων διαφορετικών εικόνων και της εγγενώς περιορισμένης λειτουργίας του, έχει σοβαρές συνέπειες για την επισήμανση των ανθρώπων, καθώς και την παρεμπόδιση ή την επιπλοκή (εγκληματικών)

ερευνών λόγω της ανάγκης να παρακολουθήσουμε “δυναμικούς πελάτες” που στην ουσία ήταν αλγοριθμικό λάθος.

Ένα άλλο σημείο ενδιαφέροντος εδώ είναι ότι, ενώ τα ποσοστά ακρίβειας είναι γενικά πολύ υψηλά σε εκπαιδευτικά σετ, σύμφωνα με μια μελέτη του 2016 [6] οι άνθρωποι εξακολουθούν να ξεπερνούν τα καλύτερα αυτοματοποιημένα συστήματα στην επαλήθευση προσώπου όταν πρόκειται για μη περιορισμένα, δηλαδή τυχαία, σύνολα εικόνας. Το να στηριχθούμε σε έναν αλγόριθμο φαίνεται να ευνοεί την ταχύτητα και την ευκολία έναντι της ακρίβειας, η οποία σε ορισμένες περιπτώσεις μοιάζει με μια αμυντική επιλογή, αλλά δεν πρέπει να αμφισβητείται ως τυπική διαδικασία δεδομένων των τρεχουσών παραμέτρων.

Πρόσφατες μελέτες έχουν δείξει ότι η αναγνώριση προσώπου, πέρα από τα προβλήματα με τους ίδιους τους αλγόριθμους, αρχίζει να θέτει έναν πιο διαδεδομένο κίνδυνο σχετικά με τα ζητήματα απορρήτου. Παρά τις αδυναμίες που αναφέρονται στα προηγούμενα δύο παραδείγματα, τα ποσοστά ακρίβειας ανεβαίνουν σταθερά, εν μέρει μέσω νέων τεχνικών, στο σημείο που γίνεται πολύ δύσκολο να ξεφύγουμε από την αναγνώριση από τις εικόνες λόγω της ποσότητας της έκθεσης, ακόμη και όταν χρησιμοποιούνται συμβατικές τεχνικές για να το επιτύχουν [115]. Οι επιπτώσεις για ζητήματα απορρήτου είναι σοβαρές και μελετημένες όταν οποιαδήποτε ροή κάμερας ή εικόνα θα μπορούσε (αυτόματα) να χρησιμοποιηθεί για την αναγνώριση ατόμων, τόσο στον κυβερνητικό όσο και στον εμπορικό τομέα. Η συζήτηση μεταξύ της ιδιωτικής ζωής και της ασφάλειας συνεχίζεται σαφώς, με την πρόοδο του αλγοριθμικού κοσκινίσματος (ζωντανών) δεδομένων που επιδεινώνει ορισμένα από τα υπάρχοντα πεδία έντασης.

Το πρόσχημα του ορθολογισμού και του ουδέτερου αυτοματισμού, ειδικά όταν ενσωματώνεται σε πολύπλοκα συστήματα, έχει αρνητικές συνέπειες που δεν είναι πάντα προφανείς και κάποιες φορές απρόσμενες. Οι Goodman και Flaxman [12] ορίζουν το *profiling* ως:

“κάθε μορφή αυτοματοποιημένης επεξεργασίας προσωπικών δεδομένων που αποτελείται από τη χρήση προσωπικών δεδομένων για την αξιολόγηση ορισμένων προσωπικών πτυχών που σχετίζονται με ένα φυσικό πρόσωπο”,

με βάση άσχετες μετρήσεις γενικά, αλλά ο αποκλεισμός, ο ρατσισμός και η σπατάλη χρόνου και χρήματος με πρακτικές έννοιες, είναι μειονεκτήματα αυτής της προσέγγισης που πρέπει να αντιμετωπιστούν ουσιαστικά και τουλάχιστον να ληφθούν υπόψη κατά την αξιολόγηση πρακτικών αναγνώρισης προσώπου.

3.5 Αντιμετώπιση των συνεπειών

Στην προηγούμενη ενότητα εξετάσαμε μια ποικιλία παραδειγμάτων που δείχνουν τις διαστάσεις του Gillespie και άλλων που δημιουργούν αρνητικές συνέπειες. Ενώ οι προσπάθειες για την εξέταση αλγορίθμων από ηθική σκοπιά άρχισαν να αναπτύσσονται μάλλον πρόσφατα, έχουν προταθεί μερικές ελπιδοφόρες εναλλακτικές λύσεις και προσθήκες. Επομένως, αυτή η ενότητα θα ασχοληθεί με διάφορες δυνατότητες, με στόχο τη μείωση των δυσμενών συνεπειών που περιγράφονται στην προηγούμενη ενότητα.

Μια προσέγγιση “χαμηλού επιπέδου” για την αντιμετώπιση των προβλημάτων που σχετίζονται με τους αλγόριθμους έρχεται με τη μορφή διαφάνειας μέσω της διάδοσης γνώσεων που ξεκινά με την εκπαίδευση δημοσιογράφων και υπευθύνων χάραξης πολιτικής, αυξάνοντας την επίγνωση και την ευαισθησία τους στις διάφορες αυτές λειτουργίες. Η συλλογιστική σε αυτό είναι ότι μόλις αυξηθεί η ενημερότητα αυτών των τομέων, αυτό θα επεκταθεί στην κοινωνία ευρύτερα και με δημοκρατικές διαδικασίες, έτσι ώστε να πραγματοποιηθεί μια ενημερωμένη συζήτηση. Έχει ήδη πραγματοποιηθεί κάποια εργασία [25] σχετικά με τον τρόπο με τον οποίο οι δημοσιογράφοι μπορούν να αναγνωρίσουν και να κατανοήσουν τις σχέσεις ισχύος που μεσολαβούν μέσω αλγορίθμων. Αν και αυτός είναι ένας καλός τρόπος αντιμετώπισης ορισμένων ζητημάτων αδιαφάνειας, έχει κι αυτός τα όριά του.

1. Πρώτον, ενώ η αυξημένη ευαισθητοποίηση εκ μέρους των “υπευθύνων για τη ενημέρωση” θα οδηγήσει σε πιο διορατική κάλυψη και συνεπώς κατανόηση, οι αλγοριθμικές εφαρμογές από τη φύση τους είναι ασαφείς και ότι επιπλέον οι εμπορικοί παράγοντες και λόγοι ασφάλειας θέτουν περιορισμούς στο τι μπορεί και πόσο μπορεί να αναλυθεί.
2. Δεύτερον, εάν οι προσεγγίσεις περιορίζονται σε μεθοδολογικές προσεγγίσεις εκ μέρους δημοσιογράφων και άλλων αφοσιωμένων επαγγελματιών, η λειτουργία των αλγορίθμων εμποδίζει τη διάδοση, επειδή οι ειδήσεις αποκτώνται περισσότερο βάσει ζήτησης, δηλαδή εξατομικευμένες.

Αυτή η προσέγγιση προσανατολίζεται στην αλλαγή της δομής στο πλαίσιο που οι αλγόριθμοι λειτουργούν, αντί να αλλάζουν οτιδήποτε σχετικά με τον αλγόριθμο ως τεχνολογία ή τη λογική που τον στηρίζει. Δηλαδή, η κοινωνική κατασκευή οποιουδήποτε αλγορίθμου θα υπόκειται σε αλλαγή, αντί για τα χαρακτηριστικά που συνέβαλαν σε αυτή την κατασκευή. Ωστόσο, από μια επακόλουθη προοπτική, ο μετριασμός των αρνητικών επιπτώσεων της αδιαφάνειας των αλγορίθμων, είναι μια αξιόλογη προσπάθεια.

Μια πιο αυστηρή προσέγγιση έρχεται με τη μορφή εφαρμογής πολιτικών που αντιμετωπίζουν άμεσα ορισμένες από τις αρνητικές συνέπειες της αλγοριθμικής λειτουργίας. Αυτό παραλείπει το βήμα της ενημέρωσης ενός ευρύτερου αριθμού ατόμων. Πρόσφατα, το Συμβούλιο της Ευρωπαϊκής Ένωσης επιβεβαίωσε ότι πρότεινε τον Γενικό Κανονισμό Προστασίας Δεδομένων, που τέθηκε σε ισχύ το 2018 [12]. Ο πιο αξιοσημείωτος στόχος είναι η ενημέρωση όσων επηρεάζονται από αλγόριθμους και ταυτόχρονη δημιουργία προφίλ,

ζητώντας κυριολεκτικά ένα “δικαίωμα στην εξήγηση” [12]. Αυτό παρουσιάζει προκλήσεις για τις τρέχουσες εμπορικές και κυβερνητικές δραστηριότητες στην ΕΕ που πρέπει να αναθεωρηθούν λόγω αυτών των απαιτήσεων. Αυτό είναι μια ευκαιρία ακριβώς επειδή επιστά την προσοχή σε πρακτικές που εισάγουν διακρίσεις και μπορεί να οδηγήσει σε ένα νέο παράδειγμα αλγοριθμικού σχεδιασμού. Με τη σειρά του προκύπτουν παρόμοιες σκέψεις όπως ο Διακόπουλος υποστήριξε [26] πως για τους σχεδιαστές είναι ευκαιρία να προσπαθούν να δημιουργήσουν μια κουλτούρα διαφάνειας σε όλα τα επίπεδα αλγοριθμικής ανάπτυξης.

3.6 Συμπεράσματα

Στο κεφάλαιο αυτό είδαμε ότι οι αλγόριθμοι δεν αποτελούν ουδέτερη και αντικειμενική τεχνολογία καθώς και μοτίβα των ηθικών πτυχών των αλγορίθμων. Είδαμε τις έξι αλληλένδετες διαστάσεις που όρισε ο Gillespie, και είδαμε τρία παραδείγματα αλγοριθμικών εφαρμογών. Τέλος είδαμε τρόπους αντιμετώπισης των συνεπειών, των ηθικών επιπτώσεων. Ωστόσο, οι προτάσεις παραμένουν ασαφείς, αποτελούμενες από κατευθυντήριες αρχές και συχνά δεν έχουν τη δύναμη να ξεπεράσουν έναν μη δεσμευτικό χαρακτήρα. Ακόμη και στην περίπτωση των κανονισμών της ΕΕ, είναι περιοριστικό - τι δεν επιτρέπεται - αφήνοντας τις λύσεις σε εμπορικά μέρη που είναι εν μέρει υπεύθυνα για την τρέχουσα κατάσταση. Παρόλα αυτά, αυτές οι δυσκολίες, η αυξημένη προσοχή στις ηθικές πλευρές των αλγοριθμικών εφαρμογών, επισημαίνουν την ύπαρξη ενός προβληματισμού, είναι λόγος για συγκρατημένη αισιοδοξία. Ενώ οι αρνητικές συνέπειες της αλγοριθμικής λειτουργίας μπορεί να αλλάξουν, το γεγονός ότι τα προβλήματα λαμβάνουν την προσοχή είναι ένα σημάδι ότι η αναμφισβήτητη ουδετερότητα των αλγορίθμων έχει αποβληθεί.

Εξετάσαμε τις συνέπειες των αλγοριθμικών εφαρμογών. Τα επιχειρήματα υπέρ και τα μειονεκτήματα επικεντρώθηκαν στο κατά πόσον αυτές οι συνέπειες ήταν επιθυμητές ή όχι και σε τι συνίσταται η εξισορρόπηση των αρνητικών πτυχών σε διάφορες περιπτώσεις. Τα επιχειρήματα αυτά μπορεί να θεωρηθούν ότι ανήκουν στο φάσμα της αξιολόγησης συγκεκριμένων περιπτώσεων, όπου απαιτούν στενότερη επιθεώρηση με ηθικούς όρους. Το θεωρητικό πλαίσιο όπου οι αλγόριθμοι επιφέρουν αυτές τις δυναμικές, καθώς και τα παραδείγματα που τις απεικονίζουν, εστιάζουν σε δύο στοιχεία:

1. Τι είδους ακούσιες συνέπειες παράγουν οι ίδιοι οι αλγόριθμοι
2. Πώς μια αναντιστοιχία μεταξύ της λειτουργίας των αλγορίθμων και των κοινωνικοπολιτισμικών δομών που ενσωματώνονται ή χρησιμοποιούνται έχουν αρνητικά αποτελέσματα.

Αυτή ήταν μια πολύτιμη άσκηση, αποδίδοντας μια κριτική που προσπαθεί να αφαιρέσει τη μάσκα της “ορθολογικότητας”, της “ουδετερότητας” ή της αυστηρής λειτουργικότητας που συχνά συνοδεύει τους αλγόριθμους. Ωστόσο, υπάρχει μια διάσταση στους αλγόριθμους που δεν καταγράφεται από μια τέτοια ανάλυση, λόγω των εγγενών ιδιοτήτων των αλγορίθμων.

Κεφάλαιο 4

Αλγοριθμική Επιλογή

Στο προηγούμενο κεφάλαιο, τα ζητήματα που τέθηκαν θα μπορούσαν να αντιμετωπιστούν πρακτικά, δηλαδή με την εξασφάλιση ότι ο αλγόριθμος δεν παράγει ανεπιθύμητα αποτελέσματα, ή ενημερώνοντας τους ενδιαφερόμενους σε επαρκή βαθμό έτσι ώστε η αδιαφάνεια να μην είναι πλέον προβληματική πτυχή. Εάν οι προγραμματιστές αλγορίθμων, νομοθέτες κ.λπ. μπορούσαν να αντιμετωπιστούν αυτά τα προβλήματα επαρκώς, θα ήταν δεοντολογικά δικαιολογημένο να χρησιμοποιούνται αλγόριθμοι για τέτοιους σκοπούς. Αυτά τα προβλήματα και τα συναφή αντίμετρα ή στρατηγικές υπολείπονται, ωστόσο, για την αντιμετώπιση του ζητήματος εάν η επιλογή για την υιοθέτηση μιας αλγοριθμικής προσέγγισης για οποιαδήποτε κατάσταση ή φαινόμενο μπορεί από μόνη της να είναι ηθικά αμφισβητούμενη. Αυτή δεν είναι απλώς μια θεωρητική διάσταση των αλγορίθμων, καθώς οι απαιτήσεις για την αλγοριθμική λειτουργία είναι συγκεκριμένες και εγγενώς περιοριστικές, καθώς και ότι η ανάθεση ή η αυτοματοποίηση της ευθύνης μπορεί να είναι αμφιλεγόμενη.

Σε αυτό το κεφάλαιο θα εξετάσει πώς η επιλογή, η δημιουργία και η αποδοχή των απαιτούμενων προϋποθέσεων για μια αλγοριθμική προσέγγιση μπορεί να είναι προβληματική ανεξάρτητα από τις συνέπειες, υιοθετώντας μια δεοντολογική ηθική προσέγγιση. Η δεοντολογία επιλέγεται επειδή δεν βασίζεται ρητά στην αξιολόγηση των συνεπειών ή των περιστάσεων των ενεργειών, αλλά στην αρχή στην οποία αναλαμβάνεται. Αυτό αντιστοιχεί στην επιλογή των απαιτούμενων συνθηκών για αλγοριθμική λειτουργία και αιτιολόγηση της εφαρμογής του με ευρύτερη έννοια, επισημαίνοντας τις ερωτήσεις που δεν έχουν αντιμετωπιστεί στο προηγούμενο κεφάλαιο. Εν ολίγοις, ποιες ηθικά προβληματικές πτυχές της αλγοριθμικής λειτουργίας δεν μπορούν να βελτιωθούν ή να επιλυθούν χωρίς να αμφισβητηθεί η χρήση αλγορίθμων.

4.1 Δεοντολογία

Η δεοντολογία είναι η κανονιστική ηθική παράδοση που προέρχεται από την αρχαία ελληνική λέξη για το καθήκον, δέον. Η δεοντολογία επικεντρώνεται στην αξιολόγηση των δράσεων με βάση την τήρηση των αρχών, την τήρηση των καθηκόντων και τον σεβασμό των δικαιωμάτων και όχι τα αποτελέσματα. Σε γενικές γραμμές, εκτιμά την πρόθεση του κινήτρου των δράσεων έναντι των αποτελεσμάτων [61], καθώς και η τήρηση των αρχών και ο σεβασμός των δικαιωμάτων, που αποτελούν καθήκοντα. Οι δεοντολογικές αξιολογήσεις εστιάζουν επομένως σε “βαθιά αισθητές πεπαιθώσεις και υπαρξιακά ενδιαφέροντα” [97] που καθοδηγούν αυτές τις αρχές, τα καθήκοντα και τα δικαιώματα. Όπως συμβαίνει με όλες τις

γενικές κανονιστικές θεωρίες, υπάρχουν πολλές αποχρώσεις και διαφορές μεταξύ διαφορετικών εκδόσεων δεοντολογίας, αν και τα χαρακτηριστικά που περιγράφονται παραπάνω οριοθετούν το πεδίο εντός του οποίου πραγματοποιείται μια τέτοια διαφοροποίηση. Χρησιμοποιούμε την δεοντολογία για να διερευνήσουμε πώς η χρήση αλγορίθμων δεν είναι ουδέτερη πράξη, αλλά σχετίζεται και εισχωρεί σε σημαντικές ανθρώπινες αξίες [96].

Η δεοντολογική προσέγγιση σε αυτό το κεφάλαιο προέρχεται από το γεγονός ότι επιτρέπει την αμφισβήτηση της επιλογής χρήσης αλγορίθμων, ανεξάρτητα από τα αποτελέσματα. Δεδομένου ότι η δεοντολογία ασχολείται με το εάν τηρούνται ορισμένες αρχές, τα δικαιώματα και τα καθήκοντα, θα δοθεί έμφαση στο ποιες αξίες διακυβεύονται και αν είναι ηθικά επιτρεπτή η αποδοχή καταπάτησης σε αυτές. Μια ηθική προσέγγιση είναι ιδιαίτερα διαφωτιστική όταν οι συνέπειες είναι αβέβαιες, καθώς και όταν η δυσκολία στη μέτρηση και την εκτίμηση σε ποιο βαθμό και από ποιον τα αποτελέσματα είναι θετικά. Η ουδετερότητα της χρήσης αλγορίθμων θα αμφισβητηθεί αφού φέρνουν σε δεύτερη μοίρα σημαντικές ηθικές αξίες που παίζουν ρόλο πέρα από το αποτέλεσμα. Αυτή η προσέγγιση διαφέρει σε σχέση με το προηγούμενο κεφάλαιο, διότι δεν στοχεύουμε σε πιθανές στρατηγικές για την αποφυγή αρνητικών συνεπειών, ο στόχος εδώ είναι να πυροδοτήσει και να ενημερώσει και να προβληματίσει σχετικά με το εάν αυτές οι αξίες είναι αρκετά σημαντικές για σταματήσει μια αλγοριθμική προσέγγιση χωρίς να πάρει μια οριστική στάση για το θέμα.

4.2 Τα χαρακτηριστικά των αλγορίθμων από την σκοπιά της Δεοντολογίας

Όπως έχουμε δει στο πρώτο κεφάλαιο, το οποίο στοχεύει στην περιγραφή των αλγορίθμων γενικά, υπάρχουν θεμελιώδη χαρακτηριστικά που δεν εξαρτώνται από το περιβάλλον. Οι αρχές αυτές, θα χρησιμοποιηθούν σε αυτήν την ενότητα, και θα συνδυαστούν με διάφορες διαστάσεις του θεωρητικού πλαισίου του προηγούμενου κεφαλαίου, το οποίο υπογράμμισε πρότυπα ηθικής ανησυχίας. Αυτή η σύζευξη θα δείξει ότι τα χαρακτηριστικά των αλγορίθμων οδηγούν αναγκαστικά σε αυτά τα μοτίβα, και συνεπώς δεν μπορούν να αντιμετωπιστούν βελτιώνοντας τον αλγόριθμο ή την υποδομή που τον περιβάλλει, με αποτέλεσμα δεοντολογικές ανησυχίες.

Οι τρεις ιδιότητες ή απαιτήσεις αλγορίθμων που σχετίζονται με την τρέχουσα ανάλυση είναι:

1. Καλά καθορισμένη είσοδος από μια επίσημη γλώσσα
2. Η κωδικοποίηση και αυτοματοποίηση εργασιών
3. Η υπόθεση της θετικιστικής λογικής.

Όλες αυτές οι ιδιότητες προέρχονται από τις βάσεις που έχουν τεθεί νωρίτερα, αλλά πρέπει να προσδιοριστούν, επειδή αυτές οι ιδιότητες δημιουργούν, όταν εξετάζονται από μια δεοντολογική προοπτική, σε συγκεκριμένα ζητήματα που δεν αντιμετωπίστηκαν ως τώρα.

4.2.1 Η καλά ορισμένη Είσοδος

Η απαίτηση των αλγορίθμων ότι η είσοδος προέρχεται από μια τυπική γλώσσα, δηλαδή δεν είναι ανοιχτή σε πολλαπλές ερμηνείες, αλλά αντίθετα ορίζεται σαφώς, θέτει περιορισμούς στον τύπο εισόδου που επιτρέπεται και συνεπώς στις πιθανές εφαρμογές των αλγορίθμων. Δίνεται έμφαση, επειδή ενώ ένας μεγάλος αριθμός καθημερινών φαινομένων δεν ανήκουν σε μια κατηγορία επίσημων γλωσσών, αυτά τα φαινόμενα μπορούν να μεταμορφωθούν με τέτοιο τρόπο ώστε να γίνουν ευανάγνωστα ή χρησιμοποιήσιμα για αλγόριθμους. Πριν δείξουμε πώς αυτός ο μετασχηματισμός μπορεί να είναι προβληματικός, είναι σημαντικό να εξηγήσουμε ότι ο λόγος για να επιμείνουμε στη χρήση μη τυπικών, είναι ότι οι αλγόριθμοι 'μπορούν' να επεξεργαστούν είσοδο που είναι αυστηρά ποιοτική. Για παράδειγμα, σκεφτείτε έναν αλγόριθμο ο οποίος για κάθε είσοδο από μια σειρά ακέραιων προσδιορίζει μόνο την ποιοτική ιδιότητα να είναι θετικός, αρνητικός ή μηδέν. Σε πρώτο επίπεδο, αυτή είναι μια ποιοτική ιδιότητα της εισαγωγής, αλλά είναι εύκολο να δούμε ότι μπορεί επίσης να πλαισιωθεί από τρεις πιθανές κατηγοριοποιήσεις, η οποία είναι μια ποσοτική και διακριτή διαδικασία.

Επιστρέφοντας στην αναντιστοιχία μεταξύ φαινομένων που δεν ταιριάζουν άμεσα για μια αλγοριθμική προσέγγιση και την επιθυμία να τα επεξεργαστούμε με τέτοιο ακριβώς τρόπο, πρέπει να σημειωθεί ότι ο μετασχηματισμός που απαιτείται για την ικανοποίηση αυτής της επιθυμίας οδηγεί σε μια διαφορετική μορφή του θέματος. Αυτό είναι ανάλογο με τη γενικότερη ιδέα ότι ένα "μοντέλο" ή άλλος τύπος αναπαράστασης είναι πάντα μια συγκεκριμένη ερμηνεία, και αναπαριστάνεται εύστοχα με τη φράση "ο χάρτης δεν είναι το έδαφος" [1]. Μια δεοντολογική αντίρρηση έρχεται με τη μορφή ότι οι αλγόριθμοι έχουν αναγκαστικά την συνέπεια της σύγχυσης της αναπαράστασης. Η πρακτική του μετασχηματισμού δεν είναι κάτι καινούριο. Για παράδειγμα, δοκιμές IQ που μεταμορφώνουν την αφηρημένη έννοια των γνωστικών ικανοτήτων σε έναν μόνο αριθμό ενσωματωμένο σε ένα μοντέλο με τυποποιημένη μέση και τυπική απόκλιση. Το γεγονός ότι οι αλγόριθμοι χρησιμοποιούνται όλο και περισσότερο στην κοινωνία, όπως φαίνεται και στο προηγούμενο κεφάλαιο, υπογραμμίζει την ανάγκη αιτιολόγησης στη μετατροπή των φαινομένων σε σαφώς καθορισμένες εισόδους, ώστε να προστατεύσει αυτό που "χάνεται στη μετάφραση".

4.2.2 Κωδικοποίηση και αυτοματοποίηση

Οι αλγόριθμοι είναι από τη φύση τους κωδικοποιημένες διαδικασίες για τη μετατροπή της εισόδου σε έξοδο, αν και με ποικίλους βαθμούς πολυπλοκότητας και δυνατότητες για μηχανισμούς ανάδρασης, π.χ. μια έξοδος που χρησιμεύει ως περαιτέρω είσοδος, κ.λπ.

Ωστόσο, ακόμη και στην περίπτωση τέτοιων δυναμικών αλγορίθμων, η ίδια η δυναμική φύση κωδικοποιείται, έτσι ώστε να υπάρχουν ακριβείς και σταθερές αρχές.

Ένα άλλο σημείο τριβής σχετικά με την αυτοματοποιημένη φύση των αλγοριθμικών λειτουργιών είναι ότι η αξιολόγηση του αποτελέσματος είναι δυνατή μόνο μετά την εφαρμογή, η οποία είναι ιδιαίτερα εμφανής σε σύνθετες καταστάσεις που είτε είναι απρόβλεπτες, είτε διαμορφώνονται σε σχέση με τους αλγόριθμους που λειτουργούν. Για παράδειγμα, ακόμη και τα καλύτερα μετεωρολογικά μοντέλα διαθέτουν αβεβαιότητα και περιθώρια σφάλματος, τα οποία είναι αποδεκτά ως “βέλτιστη πρακτική”. Ωστόσο, η εγγενής πολυπλοκότητα άλλων συστημάτων απαιτεί ένα διαφορετικό πρότυπο, όπως κατά τη μοντελοποίηση της συμπεριφοράς (ημι-) αυτόνομων συστημάτων, λόγου χάρη, αυτοκίνητα ή λογισμικό με αντίκτυπο στις χρηματοπιστωτικές αγορές. Αυτά τα επιχειρήματα περιέχουν ένα συνεκτικό αποτέλεσμα, αλλά η εστίαση εδώ είναι μάλλον στην επιλογή που βασίζεται στην αποδοχή της αβεβαιότητας στη μοντελοποίηση σύνθετων καταστάσεων και στην αποδοχή του γεγονότος ότι οι αυτοματοποιημένες διαδικασίες, με κωδικοποιημένες τιμές που συν-διαμορφώνουν την πραγματικότητα.

Η κωδικοποίηση και ο αυτοματισμός καθίστανται ιδιαίτερα προβληματικοί όταν η εργασία που πρόκειται να εκτελεστεί είναι η ίδια ηθική. Δηλαδή, όταν υπάρχει τουλάχιστον μια εγγενής ηθική συνιστώσα στη λειτουργία, όπως στην περίπτωση της συμπεριφοράς των αυτόνομων οχημάτων, ένα παράδειγμα που θα συζητηθεί στην επόμενη ενότητα.

4.2.3 Υπολογιστικός Θετικισμός

Η χρήση αλγορίθμων, λόγω των δύο προηγούμενων σημείων - που απαιτούν σαφώς καθορισμένη είσοδο, καθώς και κωδικοποίηση έχει ως συνέπεια να συνεπάγεται με ευθύνη - σημαίνει την επιλογή εγγραφής σε μια συγκεκριμένη θετικιστική επιστημολογία. Η θετικιστική προσέγγιση στα φαινόμενα μεγάλης κλίμακας είναι στην πραγματικότητα απαίτηση ή δικαιολογία για τη χρήση αλγορίθμων.

Αυτή η αποδοχή μιας θετικιστικής προοπτικής καθίσταται ιδιαίτερα προβληματική σε περιπτώσεις ασάφειας, όπου σαφείς λύσεις είτε σε θέμα προσωπικής ή ηθικής “προτίμησης”. Ο Kraemer [63] έχει δείξει ότι, σε ιατρικό περιβάλλον κατά την ανάπτυξη αλγορίθμων για την ερμηνεία των σαρώσεων, οι ερευνητές μπορούν να έχουν διαφορετικούς λογικούς λόγους για την ανάπτυξη διαφορετικών αλγορίθμων. Ως εκ τούτου, ο “ορθολογισμός” δεν είναι το κρίσιμο κριτήριο που περιγράφει τις επιλογές στη διαδικασία σχεδιασμού των αλγορίθμων, οπότε αντί να ανακαλύψουμε μοτίβα και αφήνοντας τα “δεδομένα” να μιλήσουν από μόνα τους, επιστρέφουμε στην ιδέα ότι “ακατέργαστα” ή μη δομημένα δεδομένα είναι οξύμωρο [44]. Ενώ αυτό γίνεται ως επί το πλείστον πρόβλημα κατά την εξέταση των κοινωνικών φαινομένων, είναι ωστόσο αναπόφευκτη συνέπεια της αλγοριθμικής προσέγγισης, επειδή ο επιστημολογικός θετικισμός υποστηρίζει τους

αλγορίθμους στο ότι απαιτεί σαφώς καθορισμένη, “αντικειμενική” εισαγωγή [41], με αποτέλεσμα υπολογιστικό θετικισμό.

4.3 Περιπτώσεις Δεοντολογικών Ηθικών Θεμάτων

Στην προηγούμενη ενότητα περιγράψαμε σε γενικές γραμμές πώς τα βασικά χαρακτηριστικά των αλγορίθμων κάνουν την επιλογή για τη χρήση ηθικών αλγορίθμων ηθικά αμφισβητούμενων λόγω του μετασχηματισμού των φαινομένων, της ανάθεσης καθηκόντων και ευθυνών και της υπόθεσης του θετικισμού στα κοινωνικά φαινόμενα.

Σε αυτήν την ενότητα θα αναλυθούν ορισμένες περιπτώσεις στις οποίες αυτό είναι πρακτικά ανησυχητικό. Σε αυτήν την ενότητα, ο στόχος είναι να επισημανθούν πεδία έντασης μεταξύ αλγοριθμικών προσεγγίσεων και των αξιών που καταπατούν. Αντί να προσφέρουμε προτάσεις για την επίλυση αυτών των εντάσεων, ο στόχος είναι να πυροδοτήσουμε και να ενημερώσουμε, διότι η φύση της δεοντολογικής ανάλυσης που υιοθετήθηκε εδώ δεν αφορά τόσο το επιχείρημα για την προτίμηση συγκεκριμένων αξιών, αλλά για να δείξουμε με ποιον τρόπο διακυβεύονται αξίες. Δύο παραδείγματα θα επεξηγήσουν αυτό το επιχείρημα

1. Επανεξέταση της περίπτωσης στο παράδειγμα της Google που παρουσιάστηκε επίσης στο προηγούμενο κεφάλαιο, διευκρινίζοντας πώς μια διαφορετική ηθική προσέγγιση αποφέρει διαφορετικά συμπεράσματα.
2. Αυτόνομα οχήματα (Autonomous Vehicles AVs): η συμπεριφορά σε ακραίες καταστάσεις με κωδικοποίηση ηθικής συμπεριφοράς σε (στρατιωτικά) αυτόνομα αεροσκάφη, τα οποία από τη φύση τους ασκούν έλεγχο καθώς και (δυναμικά) προκαλούν σωματική βλάβη.

Και τα δύο παραδείγματα αποτελούν περιπτώσεις ηθικά αμφισβητούμενες. Θα ενισχύσουν την υπόθεση ότι πριν από την επιλογή μιας αλγοριθμικής προσέγγισης, θα πρέπει να εξετάσουμε προσεκτικά τι ακριβώς σημαίνει αυτή η επιλογή όσον αφορά τις εκτιμήσεις αξίας.

4.3.1 Η περίπτωση της Google

Η πρακτική της Google να ταξινομεί αλγοριθμικά τα αποτελέσματα αναζήτησης, ενώ δημιουργεί ή επιτρέπει σε ένα περιβάλλον στο οποίο οι άνθρωποι δεν γνωρίζουν το γεγονός ότι τα αποτελέσματα είναι συνέπειες που αποτελούν συγκεκριμένες επιλογές που έκανε η εταιρεία. Ενώ έχει αποδειχθεί ότι αυτές οι συνέπειες μπορεί να είναι αρνητικές, υπάρχει μια άλλη διάσταση για την εκτέλεση τέτοιων εργασιών. Σε ακραίες περιπτώσεις, αυτό είχε ως αποτέλεσμα προκατειλημμένα ή ακόμη και εντελώς ρατσιστικά αποτελέσματα, όπως ο προσδιορισμός των εικόνων των μαύρων ως “γορίλες” [5], ή συσχετίζοντας έντονα θετικά χαρακτηριστικά με λευκά άτομα [90]. Επιδιορθώνοντας τις συγκεκριμένες περιπτώσεις, η

εταιρεία απάντησε ότι ενώ λυπάται βαθιά για τα βλαβερά αποτελέσματα, τέτοιες ατυχείς συσχετίσεις ήταν αναπόφευκτες, λόγω του γεγονότος ότι οι αλγόριθμοι της Google παράγουν περιγραφικά αποτελέσματα, πράγμα που σημαίνει ότι τυχόν υπάρχουσες προκαταλήψεις στα δεδομένα που ταξινομούνται σχετικά αντανακλάται στο “αποτέλεσμα αναζήτησης”. Αυτό, από την ίδια την παραδοχή της Google, είναι αναπόφευκτο αποτέλεσμα της ανάθεσης σε αλγορίθμους. Αυτό που κάνει αυτό το σημείο να ξεχωρίζει από μια συγκεκριμένη συνέπεια των αλγορίθμων που χρησιμοποιεί η Google, είναι ότι δεν υπάρχει “επίλυση” σε καθαρά αλγοριθμικό επίπεδο, δηλαδή αυτό είναι ένα πρόβλημα που δεν μπορεί να επιλυθεί καθαρά αλγοριθμικά. Θα πρέπει να σημειωθεί ότι η επιμέλεια ή η ταξινόμηση της συνάφειας από τους ανθρώπους θα παράγει επίσης αμφιλεγόμενα αποτελέσματα, δεδομένου του αρκετά μεγάλου αριθμού περιπτώσεων και της ποικιλίας των ερωτημάτων και των ατόμων που τα λαμβάνουν, αλλά ότι σε τέτοιες περιπτώσεις δεν υπάρχει “σκληρή” κωδικοποίηση, καθώς και ότι δεν υπάρχει ανάθεση ευθύνης.

Σε αυτήν την περίπτωση, από δεοντολογική ματιά στη θεμελιώδη επιλογή να προσεγγίσει μια κατάσταση αλγοριθμικά εκτός από μια συνεπειοκρατική προσέγγιση, είναι ότι η δεοντολογία είναι σε θέση να αναλύσει τις αναπόφευκτες “συνθήκες”, επιτρέποντας μια μετα-άποψη των ηθικών επιπτώσεων. Ενώ τα προβλήματα που περιγράφονται στο προηγούμενο κεφάλαιο θα μπορούσαν να αντιμετωπιστούν βελτιώνοντας τον τρόπο λειτουργίας του αλγορίθμου ή αναπτύσσοντας γνώσεις και την κοινωνικο-τεχνική υποδομή, αυτό το ζήτημα με τον αλγόριθμο δεν μπορεί να λυθεί χωρίς να αλλάξει την προσέγγιση σε κάτι διαφορετικό από (απλώς) ένα αλγόριθμο.

4.3.2 Αυτόνομα Οχήματα και Στρατιωτικά Drones.

Οι πρόσφατες εξελίξεις σε αλγόριθμους και αντίστοιχα σε hardware, έχουν ωθήσει την ανάπτυξη αυξανόμενου αυτοματισμού οχημάτων. Το 2000, έξι ομάδες ερευνητών κατάφεραν να ολοκληρώσουν το “Urban Challenge”, το οποίο ορίστηκε από τον Οργανισμό Προχωρημένων Ερευνητικών Έργων Άμυνας (Defense Advanced Research Projects Agency - DARPA), το “πρώτο κριτήριο αναφοράς για αυτόνομη οδήγηση σε ρεαλιστικά αστικά περιβάλλοντα” [8]. Υπάρχουν δύο γενικά συστήματα ταξινόμησης για τα επίπεδα αυτοματισμού, με το “υψηλότερο” επίπεδο να αναφέρεται σε έναν βαθμό αυτοματισμού που ονομάζεται αυτονομία. Για λόγους σαφήνειας, αυτή η ενότητα θα ασχοληθεί μόνο με αυτό το επίπεδο, αν και έχουν ήδη εφαρμοστεί ευρέως διάφορες μορφές αυτοματοποίησης, όπως αυτόματο φρενάρισμα ή αυτόματος πιλότος. Η πρώτη περιγραφή της αυτονομίας στα οχήματα, από την National Highway Traffic Safety Administration ορίζει ως εξής: “Το όχημα έχει σχεδιαστεί για να εκτελεί όλες τις κρίσιμες για την ασφάλεια λειτουργίες οδήγησης και να παρακολουθεί τις οδικές συνθήκες για ένα ολόκληρο ταξίδι. Ένας τέτοιος σχεδιασμός προβλέπει ότι ο οδηγός θα παρέχει είσοδο προορισμού ή πλοήγησης, αλλά δεν αναμένεται να είναι διαθέσιμος για έλεγχο οποιαδήποτε στιγμή κατά τη διάρκεια του ταξιδιού” [83]. Το δεύτερο επικρατέστερο πρότυπο είναι ότι από την Adaptive Consortium και περιγράφει την κατάσταση ως “τα συστήματα μπορούν να ολοκληρώσουν το πλήρες ταξίδι από την αρχή

έως τον προορισμό σε έναν τρόπο αυτοματισμού υψηλής ταχύτητας και μπορούν να το κάνουν οπουδήποτε στο δρόμο ώστε ένας άνθρωπος να μπορεί νόμιμα να οδηγήσει ένα όχημα. Εκτός από την ενεργοποίηση, την απενεργοποίηση και τον καθορισμό σημείων και προορισμών, δεν απαιτείται πλέον ανθρώπινο πρόγραμμα οδήγησης". Τέτοια οχήματα βρίσκονται σε εξέλιξη, με πρώιμα αποτελέσματα σε συγκεκριμένα μοντέλα - τα οποία μπορούν "τεχνικά" να το κάνουν ήδη, αλλά απαιτούν από τον επιβάτη να παραμείνει σε εγρήγορση και ικανό να αναλάβει τον έλεγχο του οχήματος", είναι πολύ θετικό όσον αφορά την ασφάλεια και την αξιοπιστία.

Πέρα από τις πρακτικές ανησυχίες σχετικά με την ασφάλεια, όπως η ευπάθεια του λογισμικού όσον αφορά σφάλματα και την ασφάλεια, οι ερευνητές ανησυχούν για καταστάσεις στις οποίες η βλάβη είτε στις δομές είτε στους ανθρώπους είναι αναπόφευκτη [8]. Αυτές οι καταστάσεις θα συμβούν, δεδομένου του αρκετού οχήματος και του χρόνου, και απαιτούν μοντέλα αποφάσεων προκειμένου να επιτευχθεί το αποτέλεσμα που θεωρείται πιο επιθυμητό. Έχουν διεξαχθεί προκαταρκτικές μελέτες σχετικά με τη στάση απέναντι σε τύπους αποφάσεων που λαμβάνονται από αυτόνομο όχημα σε διάφορες θεωρητικές καταστάσεις που απειλούν τη ζωή [8]. Αυτές οι μελέτες διαπίστωσαν ότι οι άνθρωποι ακολουθούν σε μεγάλο βαθμό τους ακόλουθους τρόπους ηθικής αξιολόγησης, όπως για παράδειγμα, σε ένα ατύχημα η ευνοϊκή επιλογή θα είναι η εξοικονόμηση μεγαλύτερου αριθμού ατόμων, ηλικιακά κριτήρια κ.λπ. Ωστόσο, υπάρχουν διαφορές μεταξύ της δύναμης αυτών των πεποιθήσεων, ειδικά όταν η ζωή του επιβάτη βρίσκεται σε κίνδυνο. Χωρίς να προχωρήσουμε πάρα πολύ στις λεπτομέρειες, η ύπαρξη διαφοράς στην προτιμώμενη συμπεριφορά του αυτόνομου οχήματος είναι θεωρητική ώστε να μεταφραστεί σε διαφορετικά μοντέλα αλγοριθμικών αποφάσεων, με αποτέλεσμα διαφορετικά, πιθανώς ανταγωνιστικά αυτόνομα οχήματα [8]. Αν και όλα αυτά τα στοιχεία δείχνουν μια ανάλυση της δυναμικής, το πιο θεμελιώδες ερώτημα για το αν είναι ηθικά επιτρεπτό να μεταβιβάσουν τέτοιες ηθικές αποφάσεις σε μια αυτόματη διαδικασία για να ξεκινήσει πρέπει να αρχίσει στη συζήτηση. Αυτή η ερώτηση σπάνια τίθεται, και πολύ λιγότερο αντιμετωπίζεται, και παρά τη σχετικά χαμηλή προσοχή που έχει προσελκύσει αυτό το ζήτημα, το να γνωρίζουμε είναι βασικό. Ενώ η εμφάνιση καταστάσεων "ζωής και θανάτου" είναι θεωρητική [8], η απαίτηση κωδικοποίησης αλγορίθμων πριν από την εμφάνιση τέτοιων καταστάσεων, είναι προβληματική επειδή δεν επιτρέπει αυτοσχεδιασμό και υποθέτει ότι το προκαθορισμένο σύνολο κωδικοποιημένων μεταβλητών είναι αρκετό για να κρίνει ικανοποιητικά οποιαδήποτε πιθανή κατάσταση. Ένα δεύτερο σημείο είναι η απαίτηση για σαφώς καθορισμένη είσοδο, η οποία συγκρούεται με κάτι τόσο υπαρξιακό και άυλο όσο η ανθρώπινη ζωή. Το δεοντολογικό ζήτημα επομένως επικεντρώνεται στο κατά πόσον είναι ηθικά επιτρεπτή η ανάθεση τέτοιων αποφάσεων και εάν είμαστε πρόθυμοι να μετατρέψουμε την ανθρώπινη ζωή, την υγεία και την ασφάλεια σε σαφώς καθορισμένη, επίσημη είσοδο σε έναν κώδικα.

Παρόμοια δυναμική σχετίζεται με την ανάπτυξη στρατιωτικών drone. Η ανάπτυξη αλγορίθμων που προορίζονται ειδικά για στρατιωτικούς σκοπούς οδήγησε σε αυτόνομα συστήματα που υπερέχουν σε αξιόπιστα από τους ανθρώπινους χειρισμούς [31]. Ενώ με

αυτόνομα οχήματα, η εμφάνιση βλάβης είναι μια ατυχής, αναπόφευκτη παρενέργεια, για την οποία καταβάλλονται προσπάθειες για την ελαχιστοποίηση των πιθανοτήτων, για στρατιωτικούς σκοπούς η σωματική και ψυχική βλάβη μπορεί να είναι ρητοί στόχοι. Ενώ τα ζητήματα με την κωδικοποιημένη συμπεριφορά και τα καλά καθορισμένα στοιχεία είναι τα ίδια, τότε, υπάρχει ένα επιπλέον επίπεδο προεξοχής στην εκπροσώπηση της αντιμετώπισης της βίας. Μέχρι στιγμής, σε πραγματικές συγκρούσεις-καταστάσεις, υπήρξε απαίτηση για τους ανθρώπους να είναι “στο βρόχο”, έτσι ώστε η αντιπροσωπεία να μην είναι ανταγωνιστική και να μην υπάρχει πόλεμος “χωρίς κίνδυνο” [88]. Ωστόσο, ενώ πλήρως αυτόνομα drones δεν έχουν δει καμία χρήση εκτός προσομοιώσεων, η ιδέα ότι η απαιτούμενη τεχνολογία θα αποκτηθεί από πολλές αντίπαλες στρατιωτικές ομάδες επισημαίνει τα λειτουργικά οφέλη της ανάπτυξης πλήρως αυτόνομων όπλων, καθιστώντας τις ηθικές ανησυχίες που στηρίζουν την τρέχουσα πρακτική της διατηρώντας τους ανθρώπους ανίκανους και αδιάφορους. Αυτό, λοιπόν, απαιτεί μια προσέγγιση παρόμοια με αυτή των πυρηνικών όπλων [82] [88], όπου η χρήση τους θεωρείται αδικαιολόγητη, ανεξάρτητα από τη σύγκρουση, και ως εκ τούτου πρέπει να υπάρχει αυστηρή ρύθμιση.

4.3.3 Τεχνητή Νοημοσύνη και Robots

Η Ηθική των Ρομπότ, Roboethics είναι ο τομέας που αφορά τους κανόνες που πρέπει να δημιουργηθούν για τα ρομπότ ώστε να διασφαλίζεται η ηθική συμπεριφορά τους και πώς να σχεδιάζονται ηθικά ρομπότ. Ο σκοπός της ρομποτικής είναι να διασφαλίσει ότι οι μηχανές με τεχνητή νοημοσύνη (AI) συμπεριφέρονται με τρόπους που δίνουν προτεραιότητα στην ανθρώπινη ασφάλεια πάνω από τις ανατεθείσες εργασίες τους και τη δική τους ασφάλεια και που είναι επίσης σύμφωνες με τις αποδεκτές αρχές της ανθρώπινης ηθικής. Τα επιχειρήματα προέρχονται κυρίως από το ερώτημα εάν τα ρομπότ έχουν δικαιώματα όπως έχουν οι άνθρωποι και τα ζώα. Αυτές οι ερωτήσεις έχουν αποκτήσει μεγαλύτερη σημασία λόγω των αυξημένων δυνατοτήτων της AI και τώρα ινστιτούτα όπως το AI Now επικεντρώνονται στην εξερεύνηση αυτών των ερωτήσεων με ακαδημαϊκή αυστηρότητα.

Ένας πρωτοπόρος στον τομέα, ο συγγραφέας επιστημονικής φαντασίας Isaac Asimov συνέταξε τους τρεις νόμους της ρομποτικής για να καθοδηγήσει την ηθική συμπεριφορά των ρομπότ:

1. Τα ρομπότ δεν πρέπει ποτέ να βλάπτουν τα ανθρώπινα όντα ή, μέσω της αδράνειας, να επιτρέπουν στον άνθρωπο να βλάψει.
2. Τα ρομπότ πρέπει να ακολουθούν οδηγίες από ανθρώπους χωρίς να παραβιάζουν τον κανόνα 1.
3. Τα ρομπότ πρέπει να προστατεύονται χωρίς να παραβιάζουν τους άλλους κανόνες.

Το Βρετανικό Ινστιτούτο Προτύπων προσέφερε ένα άρθρο το "BS8611: Robots and Robotic Devices" που είναι πιο ολοκληρωμένο και προορίζεται για τους δημιουργούς ρομπότ να διασφαλίσουν ότι τα μηχανήματα τους συμπεριφέρονται δεοντολογικά. Το BS8611 προτείνει

ότι “τα ρομπότ δεν πρέπει να σχεδιάζονται αποκλειστικά ή κυρίως για να σκοτώνουν ή να βλάπτουν ανθρώπους. οι άνθρωποι, όχι τα ρομπότ, είναι οι υπεύθυνοι πράκτορες. θα πρέπει να είναι δυνατό να ανακαλυφθεί ποιος είναι υπεύθυνος για οποιοδήποτε ρομπότ και τη συμπεριφορά του”. Οι όροι περιλαμβάνουν την αρχή ότι ο σχεδιασμός τους δεν πρέπει να επιτρέπει πολιτιστικές, σεξουαλικές ή άλλες διακρίσεις. Στο άρθρο αυτό βλέπουμε τον προβληματισμό εάν τα ρομπότ πρέπει να είναι σχεδιασμένα για να ενθαρρύνουν συναισθηματικούς δεσμούς στους χρήστες.

Μεγάλο μέρος της ανησυχίας που οδήγησε στην ανάγκη για αυτούς τους κανόνες προέρχεται από ερωτήσεις εμπνευσμένες από καλλιτέχνες και συγγραφείς όπως ο Asimov και τα έργα τους. Καθώς τα ρομπότ γίνονται όλο και πιο αυτόνομα και η τεχνητή νοημοσύνη υπερβαίνει με πολλούς τρόπους τις ανθρώπινες ικανότητες, η ανάγκη για πρότυπα ρομποτικής καθίσταται πιο επιτακτική.

Οι φουτουριστές και οι τεχνολογικοί εμπειρογνώμονες, όπως ο Elon Musk, ο Steve Wozniak και ο Steven Hawking, εξέφρασαν ανησυχίες ότι αν δεν ελεγχθούν, τα ρομπότ θα μπορούσαν να οδηγήσουν στην πτώση των ανθρώπων. Οι πιο αισιόδοξες απόψεις περιλαμβάνουν την ελπίδα ότι τα προσεκτικά σχεδιασμένα ρομπότ θα μπορούσαν να βοηθήσουν τον κόσμο να ανακάμψει από τα ανθρώπινα προβλήματα.

Τα στοιχεία που υποστηρίζονται από την τεχνολογία AI μπορούν να παρέχουν αναλυτικές πληροφορίες σχετικά με την ανθρώπινη συμπεριφορά, ωστόσο, η κατάχρηση αναλυτικών στοιχείων για τον χειρισμό ανθρώπινων αποφάσεων είναι ηθικά λανθασμένη. Το πιο γνωστό παράδειγμα κακής χρήσης των αναλυτικών στοιχείων είναι το σκάνδαλο δεδομένων από το Facebook και την Cambridge Analytica. Η Cambridge Analytica πούλησε τα δεδομένα των Αμερικανών ψηφοφόρων που ανιχνεύθηκαν στο Facebook σε πολιτικές εκστρατείες και παρείχε βοήθεια και αναλυτικά στοιχεία στις προεδρικές εκστρατείες του 2016 των Ted Cruz και Donald Trump. Πληροφορίες σχετικά με την παραβίαση δεδομένων αποκαλύφθηκαν το 2018 και η Ομοσπονδιακή Επιτροπή Εμπορίου επέβαλε πρόστιμο στο Facebook ύψους 5 δισεκατομμυρίων δολαρίων λόγω παραβιάσεων της ιδιωτικής ζωής. [\[34\]](#)

Η χρήση της τεχνητής νοημοσύνης σε δικαστικά συστήματα σε όλο τον κόσμο αυξάνεται, δημιουργώντας περισσότερα ηθικά ερωτήματα για εξερεύνηση. Το AI θα μπορούσε πιθανώς να αξιολογήσει τις υποθέσεις και να εφαρμόσει τη δικαιοσύνη με καλύτερο, ταχύτερο και πιο αποτελεσματικό τρόπο από έναν δικαστή.

Οι μέθοδοι τεχνητής νοημοσύνης μπορούν ενδεχομένως να έχουν τεράστιο αντίκτυπο σε ένα ευρύ φάσμα τομέων, από τα νομικά επαγγέλματα και το δικαστικό σώμα έως τη βοήθεια στη λήψη αποφάσεων νομοθετικών και διοικητικών δημόσιων φορέων. Για παράδειγμα, μπορούν να αυξήσουν την αποτελεσματικότητα και την ακρίβεια των δικηγόρων τόσο στην παροχή συμβουλών όσο και στη δίκη, με οφέλη για τους δικηγόρους, τους πελάτες τους και την κοινωνία συνολικά. Τα υπάρχοντα συστήματα λογισμικού για δικαστές μπορούν να

συμπληρωθούν και να βελτιωθούν μέσω εργαλείων AI για να τους υποστηρίξουν στη σύνταξη νέων αποφάσεων. Αυτή η τάση προς την ολοένα αυξανόμενη χρήση αυτόνομων συστημάτων έχει περιγραφεί ως αυτοματοποίηση της δικαιοσύνης.

Ορισμένοι υποστηρίζουν ότι η τεχνητή νοημοσύνη θα μπορούσε να συμβάλει στη δημιουργία ενός δικαιότερου ποινικού δικαστικού συστήματος, στο οποίο οι μηχανές θα μπορούσαν να αξιολογήσουν και να σταθμίσουν τους σχετικούς παράγοντες καλύτερα από τον άνθρωπο, εκμεταλλευόμενοι την ταχύτητα και τη μεγάλη απορρόφηση δεδομένων. Επομένως, η AI θα λαμβάνει αποφάσεις με βάση τεκμηριωμένες αποφάσεις χωρίς καμία προκατάληψη και υποκειμενικότητα.

Υπάρχουν όμως πολλές ηθικές προκλήσεις:

- Έλλειψη διαφάνειας των εργαλείων AI: Οι αποφάσεις τεχνητής νοημοσύνης δεν είναι πάντοτε κατανοητές για τον άνθρωπο.
- Η τεχνητή νοημοσύνη δεν είναι ουδέτερη: οι αποφάσεις που βασίζονται στην τεχνητή νοημοσύνη είναι ευαίσθητες σε ανακρίβειες, αποτελέσματα που εισάγουν διακρίσεις, ενσωματωμένες ή παρεμβλημένες προκαταλήψεις.
- Πρακτικές παρακολούθησης για τη συλλογή δεδομένων και το απόρρητο των χρηστών του δικαστηρίου.
- Νέες ανησυχίες για τη δικαιοσύνη και τον κίνδυνο για τα ανθρώπινα δικαιώματα και άλλες θεμελιώδεις αξίες.

Σύμφωνα με όλα τα παραπάνω αν αναρωτηθούμε αν θα θέλαμε να κριθούμε από ένα ρομπότ σε δικαστήριο. Θα θέλαμε, ακόμη κι αν δεν είμαστε σίγουροι πώς καταλήγει στα συμπεράσματά του;

Συμπεράσματα

Ένα από τα κύρια συμπεράσματα αυτής της διπλωματικής είναι ότι ενώ οι αλγόριθμοι μπορούν πράγματι να παράγουν μετρήσεις και πληροφορίες για τα μοτίβα, δημιουργούν επίσης (εν μέρει) αυτά τα μοτίβα, διαμορφώνουν τη συμπεριφορά και βλέπουν τον κόσμο με συγκεκριμένο ποσοτικό, επεξεργασμένο τρόπο, δημιουργώντας δυσμενείς συνέπειες η διαδικασία. Δηλαδή, δεν μας λένε κάτι για τον κόσμο ως εξωτερικός αμερόληπτος παρατηρητής, αλλά συνδιαμορφώνουν ενεργά και αλληλοσυνδέονται με αυτό που περιγράφουν με πολλούς τρόπους. Αυτή η συνειδητοποίηση αποκαλύπτει την υποτιθέμενη ουδετερότητα και ορθολογισμό τους, καθώς δεν μπορούν να θεωρηθούν αντικειμενικά εργαλεία, ούτε να θεωρηθούν ότι λειτουργούν “έξω” από τα φαινόμενα στα οποία προορίζονται να παρέχουν πληροφορίες. Δεδομένου των αρνητικών συνεπειών που έχουμε παρατηρήσει. Αυτό θα πρέπει να προκαλέσει μια αλλαγή στην προοπτική ότι, τόσο στη φάση σχεδιασμού όσο και ως προς την κατάσταση, αναλύει θεμελιωδώς τους αλγόριθμους ως σημαντικά αντικείμενα που εμπλέκονται στην οικοδόμηση του κόσμου, την κατηγοριοποίηση κ.λπ. Υπό το πρίσμα αυτό, το δεύτερο υπο-ερώτημα με τη μορφή "ποιες είναι οι τρέχουσες εφαρμογές αλγορίθμων και ποιες είναι οι ηθικές συνέπειες αυτών;" έχει απαντηθεί επαρκώς, εξετάζοντας μια ποικιλία εφαρμογών και τα συνημμένα ηθικά ζητήματα. Η συνάφεια ενός τέτοιου ζητήματος υπογραμμίζεται από την πρόσφατη εργασία που επιχειρεί συγκεκριμένα να αναπτύξει και να υποστηρίξει μεθοδολογίες “για να παράγει πληροφορίες για τη φύση και το έργο των αλγορίθμων” [78]. Αυτή η εργασία συνέβαλε σε προσπάθειες να συνδεθούν αφηρημένες διαστάσεις με σαφείς, σχετικές περιπτώσεις προσφέροντας αξιολογήσεις στις οποίες μπορεί να οικοδομήσει μελλοντική εργασία. Δεν είναι πιθανό σενάριο να σταματήσει η ανάπτυξη αλγορίθμων και οι εφαρμογές τους, αλλά η απόκτηση γνώσης σχετικά με τη δυναμική του τρόπου με τον οποίο παράγουν ανεπιθύμητα αποτελέσματα θα μπορούσε να οδηγήσει σε πολιτική ή σε έναν ευρύτερο αλγοριθμικό “πολιτισμό” που παροτρύνει τους παραγωγούς να δώσουν μεγαλύτερη προσοχή πριν την εφαρμογή τους. Με αυτόν τον τρόπο, η ηθική των αλγορίθμων δεν θα πρέπει να είναι μια σκέψη, ή να διατηρείται αποκλειστικά στη σφαίρα της μελέτης των συνεπειών, αλλά μάλλον να λαμβάνεται υπόψη από την αρχή. Με αυτόν τον τρόπο, μια επακόλουθη ανάλυση μπορεί να συμβάλει στην πρακτική έννοια, βελτιώνοντας την ευαισθητοποίηση και βοηθώντας να αντλήσουμε μαθήματα προκειμένου να βρούμε τρόπους για την αντιμετώπιση των αρνητικών συνεπειών. Το ερώτημα που παρουσιάστηκε ως υποστήριξη της γενικής ερευνητικής ερώτησης, “ποιες είναι οι τρέχουσες εφαρμογές των αλγορίθμων και ποιες είναι οι ηθικές συνέπειες αυτών;” έχει έτσι απαντηθεί.

Βιβλιογραφία

- [1] Korzybski Alfred. "A Non-Aristotelian System and its Necessity for Rigour in Mathematics and Physics". In: Science and Sanity (1933).
- [2] A. B. Arndt. "Al-Khwarizmi". In: The Mathematics Teacher 76.9 (1983).
- [3] Nimrod Bar-Am. In Search of a Simple Introduction to Communication. Cham: Springer International Publishing, 2016. doi: 10.1007/978-3-319-25625-2.
- [4] Solon Barocas, Sophie Hood, and Malte Ziewitz. "Governing Algorithms: A Provocation Piece." In: Available at SSRN 2245322 (2013).
- [5] BBC. Google apologises for Photos app's racist blunder. 2015. url: <http://www.bbc.com/news/technology-33347866>.
- [6] Austin Blanton et al. "A Comparison of Human and Automated Face Verification Accuracy on Unconstrained Image Sets". In: (2016).
- [7] Andreas Blass and Yuri Gurevich. "Algorithms: A Quest for Absolute Definitions". In: Bulletin of European Association for Theoretical Computer Science 81 (2003).
- [8] Jean-François Bonnefon, Azim Shariff, and Iyad Rahwan. "Autonomous Vehicles Need Experimental Ethics: Are We Ready for Utilitarian Cars?" In: ([In press]).
- [9] Boyer, Dominic, and Ulf Hannerz. 2006. "Introduction: Worlds of journalism." Ethnography 7 (1): 5-17.
- [10] Franz Brentano, The Origin of the Knowledge of Right and Wrong, International Journal of Ethics
- [11] D. Bostock. Aristotle's Ethics, Oxford: Oxford University Press.
- [12] Seth Flaxman Bryce Goodman. "EU regulations on algorithmic decision-making and a "right to explanation". In: ICML Workshop on Human Interpretability in Machine (2016). url: <http://arxiv.org/abs/1606.08813v1>.
- [13] Bucher, Taina. 2012. "Want to be on the top? Algorithmic power and the threat of invisibility on Facebook." New Media and Society, April 8.
- [14] J. Burrell. "How the machine thinks: Understanding opacity in machine learning algorithms". In: Big Data & Society 3.1 (2016). issn: 2053-9517. doi: 10.1177/2053951715622512.
- [15] J. Cheney-Lippold. "Jus Algorithmi: How the National Security Agency Remade Citizenship". In: (2016).

- [16] Lars Rune Christensen. "Stigmergy in human practice: Coordination in construction work". In: *Cognitive Systems Research* 21 (2013), pp. 40–51. issn: 13890417. doi: 10.1016/j.cogsys.2012.06.004.
- [17] Brian Christian, "The A/B Test: Inside the Technology That's Changing the Rules of Business," *Wired.com* April 25, 2012. http://www.wired.com/business/2012/04/ff_abtesting/
- [18] Church A. A note on the Entscheidungsproblem. *Journal of Symbolic Logic*. 1936;1(1):40-41.
- [19] Jack B. Copeland. "Computation". In: *The Blackwell Guide to the Philosophy of Computing and Information*. Ed. by Luciano Floridi. Blackwell Publishing Ltd, 2004, pp. 3–17.
- [20] Jack B. Copeland. *The Church-Turing Thesis*. Ed. by Edward N. Zalta. 2015. url: <http://stanford.library.sydney.edu.au/archives/sum2014/entries/church-turing/>
- [21] Kate Crawford. "Can An Algorithm Be Agonistic? Ten Scenes of Contest in Calculated Publics". In: *Science, Technology & Human Values* 41 (2016).
- [22] Datta, Tschantz και Datta 2015; Lambrecht and Tucker 2019
- [23] Deloitte. *From Brawn to Brains - The Impact of Technology on Jobs in the UK*. 2015. url: <http://www2.deloitte.com/uk/en/pages/growth/articles/from-brawn-to-brains--the-impact-of-technology-on-jobs-in-the-u.html>.
- [24] Deloitte. *Technology and People - The Great Job-Creating Machine*. 2016. url: <http://www2.deloitte.com/uk/en/pages/finance/articles/technology-and-people.html>.
- [25] Nicholas Diakopoulos. "Accountability in Algorithmic Decision Making". In: *Queue* 13 (2015).
- [26] Nicholas Diakopoulos. "Algorithmic Accountability". In: *Digital Journalism* 3.3 (2014), pp. 398–415. issn: 2167-0811. doi: 10.1080/21670811.2014.976411.
- [27] Margery J. Doyle and Leslie Marsh. "Stigmergy 3.0: From ants to economies". In: *Cognitive Systems Research* 21 (2013), pp. 1–6. issn: 13890417. doi: 10.1016/j.cogsys.2012.06.001.
- [28] Edwards, *The Closed World: Computers and the Politics of Discourse in Cold War America*, 30.
- [29] *Encyclopædia Britannica* (2008). *The Britannica Guide to the 100 Most Influential Scientists*. Running Press. p. 12. ISBN 9780762434213.
- [30] Robert Epstein and Ronald E. Robertson. "The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections". In: *Proceedings of the National Academy of Sciences of the United States of America* 112.33 (2015), E4512–21. issn: 1091-6490. Doi: 10.1073/pnas.1419828112.

- [31] Nicholas Ernest and David Carroll. "Genetic Fuzzy based Artificial Intelligence for Unmanned Combat Aerial Vehicle Control in Simulated Air Combat Missions". In: *Journal of Defense Management* 06.01 (2016). issn: 21670374. doi: 10.4172/2167-0374.1000144.
- [32] Martín Escardó. *Foundations of Computer Science*. Birmingham, January 11th 2005.
- [33] Executive Office of the President. *Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights*. 5-2016
- [34] Facebook–Cambridge Analytica data scandal - Wikipedia [Internet]. En.wikipedia.org. 2022 [cited 27 February 2022]. Available from: https://en.wikipedia.org/wiki/Facebook%E2%80%93Cambridge_Analytica_data_scandal
- [35] Finkelstein, Seth. 2008. "Google, links, and popularity versus authority." In *The Hyperlinked Society: Questioning Connections in the Digital Age*, ed. Joseph Turow and Lokman Tsui, 104-120. Ann Arbor, MI; University of Michigan Press.
- [36] Floridi and MS A Lee 2020; Aggarwal 2020.
- [37] Floridi and Taddeo 2016
- [38] Green and Chen 2019) ; Yu and Du 2019
- [39] John Forge. "A note on the definition of "dual use"". In: *Science and engineering ethics* 16.1 (2010), pp. 111–118. issn: 1471-5546. doi: 10.1007/s11948-009-9159-9.
- [40] GCCS. *The Ethics of Algorithms: from radical content to self-driving cars: Final Draft Background Paper*. url: https://www.gccs2015.com/sites/default/files/documents/Ethics_Algorithmsfinal%20doc.pdf.
- [41] Gonzalo Genova and M. Rosario Gonzalez. "Teaching Ethics to Engineers: A Socratic Experience". In: *Science and engineering ethics* 22.2 (2016), pp. 567–580. issn: 1353-3452. doi: 10.1007/s11948-015-9661-1.
- [42] Giddens, Anthony. 1991. *The Consequences of Modernity*. Stanford, CA: Stanford University Press.
- [43] Tarleton Gillespie. "The Relevance of Algorithms". In: *Media Technologies* (Forthcoming).
- [44] A. Goldberg. "In defense of forensic social science". In: *Big Data & Society* 2.2 (2015). issn: 2053-9517. doi: 10.1177/2053951715601145.
- [45] Dina Goldin and Peter Wegner. "The Church-Turing Thesis: Breaking the Myth". In: *Conference on Computability in Europe* (2005).
- [46] P. Gottlieb (6 April 2018). *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University. Retrieved 6 April 2018 – via *Stanford Encyclopedia of Philosophy*.

- [47] J. P. Gownder et al. *The Future Of Jobs, 2025: Working Side By Side With Robots: Automation Won't Destroy All The Jobs, But It Will Transform The Workforce — Including Yours*. 2015.
- [48] H. Gunarto. *Ethical Issues in Cyberspace and IT Society*, Ritsumeikan Asia Pacific University
- [49] I. Kant. *Fundamental Principles of the Metaphysic of Morals*, Project Gutenberg.
- [50] Keynes, John Maynard, 1949, "My Early Beliefs,"
- [51] Yuri Gurevich. *What is an Algorithm?* 2014. url: <http://research.microsoft.com/en-us/um/people/gurevich/Opera/209a.pdf>.
- [52] Patrick Heavey. "Synthetic biology ethics: a deontological assessment". In: *Bioethics* 27.8 (2013), pp. 442–452. issn: 0269- 9702. doi: 10.1111/bioe.12052.
- [53] D. Hilbert and Ackermann. W. *The Principles of Mathematical Logic*. Chelsea Publishing Company, 1950.
- [54] Brad Hooker. "Ethics in Conflict". In: *Science and Technology Ethics*. Ed. by Raymond E. Spier. Routledge, 2002, pp. 89–106.
- [55] Hurka, Thomas *British Ethical Theorists from Sidgwick to Ewing*, Oxford: Oxford University Press.
- [56] R. Johnson (2008). "Kant's Moral Philosophy". *Stanford Encyclopedia of Philosophy*. Retrieved 11 September 2013.
- [57] D. Jungnickel. *Graphs, networks and algorithms*. Springer Science & Business Media, 2006.
- [58] Natascha Just and Michael Latzer. "Governance by Algorithms: Reality Construction by Algorithmic Selection on the Internet". In: *Media, Culture & Society* (2016).
- [59] Noel Kalicharan. *Learn to program with C*.
- [60] T. Karppi and K. Crawford. "Social Media, Financial Algorithms and the Hack Crash". In: *Theory, Culture & Society* 33.1 (2015), pp. 73–92. issn: 0263-2764. doi: 10.1177/0263276415583139.
- [61] Migga Joseph Kizza. *Ethical and Social Issue in the Information Age*. 4th ed. Springer, 2010.
- [62] Donald Knuth. *The Art of Programming: Volume 1: Fundamental Algorithms*. Addison-Wesley, 1973.0

[63] Felicitas Kraemer, Kees van Overveld, and Martin Peterson. "Is there an ethics of algorithms?" In: *Ethics and Information Technology* 13.3 (2011), pp. 251–260. issn: 1388-1957. doi: 10.1007/s10676-010-9233-7.

[64] Fuyuki Kurasawa. *The Algorithmic Shift: Theorizing the Social Implications of Big Data and Crowdsourcing*. 2014. url: <https://isaconf.confex.com/isaconf/wc2014/webprogram/Paper50155.html>.

[65] Jaron Lanier. *Who Owns the Future?* Simon and Schuster, 2014.

[66] Lazer, David, Alex Pentland, Lada Adamic, Sinan Aral, Albert-Laszlo Barabasi, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, Tony Jebara, Gary King, Michael Macy, Deb Roy, and Marshall Van Alstyne. 2009. "Computational Social Science." *Science* 323: 721-723.

[67] Light, Jennifer. 1999. "When computers were women." *Technology and Culture* 40 (3): 455-483.

[68] Ted G. Lewis. "Cognitive stigmergy: A study of emergence in small-group social networks". In: *Cognitive Systems Research* 21 (2013), pp. 7–21. issn: 13890417. doi: 10.1016/j.cogsys.2012.06.002.

[69] Seumas Miller, Michael J. Selgelid, and Koos van der Bruggen. *Report on Biosecurity and Dual Use Research*. 2011.

[70] G.E. Moore *Principia Ethica*, Cambridge: Cambridge University Press

[71] Nelson, Katherine, A., Trevino, Linda, K. (2004) *Managing Business Ethics* (pp: 7, 10, 11).

[72] Larson 2017; Prates, Avelar and Lamb 2019

[73] Latour, Bruno. 1986. "Visualization and cognition: Thinking with eyes and hands." *Knowledge and Society* 6: 1-40.

[74] M. Leunissen. 2017, *From Natural Character to Moral Virtue in Aristotle*, New York: Oxford University Press.

[75] Longford, Graham. 2005. "Pedagogies of digital citizenship and the politics of code." *Techné: Research in Philosophy and Technology* 9 (1): 68-96.

[76] Manovich, Lev. 1999. "Database as symbolic form." *Convergence: The International Journal of Research into New Media Technologies* 5 (2): 80-99.

[77] Mager, Astrid. 2012. "Algorithmic ideology: How capitalist society shapes search engines." *Information, Communication, and Society* 15 (5): 769-787.

[78] Marcus Morgan. "Humanising Sociological Knowledge". In: Social Epistemology (2016), pp. 1–17. issn: 0269-1728. doi: 10 . 1080/02691728.2015.1119911.

[79] Evgeny Morozov. Don't Be Evil. 2011. url: <https://newrepublic.com/article/91916/google-schmidt-obama-gates-technocrats>.

[80] Yiannis N. Moschovakis. "What is an Algorithm?" In: Mathematics unlimited–2001 and beyond (2001).

[81] Chantal Mouffe. "Deliberative Democracy or Agonistic Pluralism?" In: Social Research (1999).

[82] Vincent C. Müller and Thomas W. Simpson. "Autonomous Killer Robots Are Probably Good News". In: (2014).

[83] National Highway Traffic Safety Administration. U.S. Department of Transportation Releases Policy on Automated Vehicle Development. 2013. url: <http://www.nhtsa.gov/About+NHTSA/Press+Releases/U.S.+Department+of+Transportation+Releases+Policy+on+Automated+Vehicle+Development>.

[84] NETMARKETSHARE. 2016. url: <https://www.netmarketshare.com/>.

[85] Paraschakis 2017; Perra and Rocha 2019; Milano, Taddeo και Floridi 2020)

[86] Eli Pariser. The Filter Bubble: What the Internet is hiding from you. New York: Penguin, 2012. isbn: 978-0241954522.

[87] Frank A. Pasquale. "The Algorithmic Self". In: The Hedgehog Review 7.1 (2015).

[88] Patrick Lin, George Bekey, and Keith Abney. "Autonomous Military Robotics: Risk, Ethics, and Design". In: (2008).

[89] Michael Rundle. Facebook's political 'echo chamber' is your fault, not theirs. 2015. url: <http://www.wired.co.uk/article/facebook-echo-chamber-study>.

[90] Fiona Rutherford and Alan White. This Is Why Some People Think Google's Results Are "Racist". 2016. url: https://www.buzzfeed.com/fionarutherford/heres-why-some-peoplethink-googles-results-are-racist?utm_term=.kueMWP7yA#.jmdej0y8b.

[91] Christian Sandvig, Kevin Hamilton, and Karrie Karahalios. Can an Algorithm be Unethical? 2015.

[92] Silverstone, Roger. 1994. Television and Everyday Life. New York: Routledge.

[93] Walter Sinnott-Armstrong. Consequentialism. Ed. by Edward N. Zalta. 2015. url: <http://plato.stanford.edu/entries/consequentialism/>.

- [94] T. Striphas. "Algorithmic culture". In: *European Journal of Cultural Studies* 18.4-5 (2015), pp. 395–412. issn: 1367-5494. doi: 10.1177/1367549415577392.
- [95] Ioan Susnea. "Engineering Human Stigmergy". In: *International Journal of Computers Communications & Control* 10(3) (2015), pp. 420– 427.
- [96] Cass Sunstein. "Is Deontology a Heuristic? In *Psychology, Neuroscience, Ethics and Law*". In: (2013). url: https://dash.harvard.edu/bitstream/handle/1/13548959/sunstein_is_deontology_a_heuristic.pdf?sequence=1.
- [97] Tsjalling Swierstra and Arie Rip. "Nano-ethics as NEST-ethics: Patterns of Moral Argumentation About New and Emerging Science and Technology". In: *NanoEthics* 1.1 (2007), pp. 3–20. issn: 1871-4757. doi: 10.1007/s11569-007-0005-8.
- [98] *The Future of Jobs: Employment, Skills and Workforce Strategy for the Fourth Industrial Revolution*. January 2016.
- [99] Alan Mathison Turing and B. Jack Copeland. *The Essential Turing*. Oxford University Press, 2004.
- [100] D. McKerlie. "Aristotle and Egoism", *Southern Journal of Philosophy*
- [101] Obermeyer et al. 2019; Zhou et al. 2019; Morley, Machado, et al. 2019
- [102] Pariser, Eli. 2011. *The Filter Bubble: What the Internet Is Hiding from You*. London: Penguin Press.
- [103] Piccinini G. Computationalism, The Church–Turing Thesis, and the Church–Turing Fallacy. *Synthese*. 2007;154(1):97-120.
- [104] Python Program for QuickSort - GeeksforGeeks [Internet]. Available from: <https://www.geeksforgeeks.org/python-program-for-quicksort/>
- [105] QuickSort (With Code in Python/C++/Java/C) [Internet]. Programiz.com. Available from: <https://www.programiz.com/dsa/quick-sort>
- [106] Hastings Rashdall, *The Theory of Good and Evil*, *Hibbert Journal*
- [107] Rieder, Bernhard. 2012. "ORDER BY column_name: The relational database as pervasive cultural form." Presented at The Lived Logics of Database Machinery conference, London.
- [108] Solove, Daniel. 2004. *The Digital Person: Technology and Privacy in the Information Age*. New York: New York University Press.
- [109] Raymond Turner. *The Philosophy of Computer Science*. 2016. url: <http://stanford.library.sydney.edu.au/archives/sum2014/entries/computer-science/#Compu>.

- [110] Van Couvering, Elizabeth 2010 "Search engine bias: The structuration of traffic on the World Wide Web." PhD diss., London School of Economics and Political Science
- [111] Sunstein, Cass R. 2001. Republic.com 2.0. Princeton, NJ: Princeton University Press.
- [112] Peter-Paul Verbeek. "Expanding Mediation Theory". In: Foundations of Science 17.4 (2012), pp. 391–395. issn: 1233-1821. doi: 10.1007/s10699-011-9253-8.
- [113] Peter-Paul Verbeek. What Things Do: Philosophical reflections on technology, agency, and design. Penn State University Press, 2005.
- [114] Amy van Wijnsberghe and Michael Nagenborg. "Civilizing Drones by Design". In: Routledge, 2016. Chap. 8.
- [115] Michael J. Wilber, Vitaly Shmatikov, and Serge Belongie. "Can we still avoid automatic face detection?" In: (2016).
- [116] J. Weizenbaum, Eliza: a computer program for the study of natural language communication between man and machine.
- [117] J. Weizenbaum, "The Paradoxical Role of the Computer." Holst Memorial Lecture (1983).
- [118] J. Weizenbaum, Computer Power and Human Reason.
- [119] Weizenbaum and Wendt, Islands in the Cyberstream, "Television and Internet."
- [120] Wendy Hui Kyong Chun, Programmed Visions: Software and Memory
- [121] N. S. Yanofsky. "Towards a Definition of an Algorithm". In: Journal of Logic and Computation 21.2 (2011), pp. 253–286. issn: 0955-792X. doi: 10.1093/logcom/exq016.
- [122] M. Ziewitz. "Governing Algorithms: Myth, Mess, and Methods". In: Science, Technology & Human Values 41.1 (2016), pp. 3–16. issn: 0162-2439. doi: 10.1177/0162243915608948.
- [123] Zimmer, Michael 2008. "The externalities of search 2.0: The emerging privacy threats when the drive for the perfect search engine meets Web 2.0." First Monday 13 (3). Available at <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2136/1944>.
- [124] Υπολογιστική πολυπλοκότητα [Internet]. Slideshare.net. Available from: <https://www.slideshare.net/stefanidhes/ss-15014460>