

Κατευθυνόμενες Απόψεις Και
Χειραγώγηση:
Ο Ρόλος Και Η Ηθική Πλευρά της
Τεχνητής Νοημοσύνης

Αλεξάκη Βασιλική AM:1097464
Μαντές Σωκράτης AM:1093421

Ιούνιος 2024



UNIVERSITY OF
PATRAS
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ

Περιεχόμενα

1	Υπόμνημα(434)	2
1.1	Επιλογή Άρθρων	2
1.2	Ομάδα	3
2	Περίληψη(195)	4
3	Βιβλιογραφική Παρουσίαση(1.061)	5
3.1	Βιβλιογραφική Ανασκόπηση Άρθρων	5
3.2	Συζήτηση	7
3.3	Κριτική συζήτηση με το ChatGPT	8
4	Βιβλιογραφία	12
5	Βιογραφικά(216)	14

Κεφάλαιο 1

Υπόμνημα(434)

1.1 Επιλογή Άρθρων

Η διμελής ομάδα μας για την εργασία στο μάθημα Συγγραφή και Παρουσίαση Τεχνικών Κειμένων, μετά από μία σύντομη μελέτη των δύο συλλογών άρθρων, αποφάσισε να επιλεγεί η συλλογή AI Ethics and more, λόγω του μεγαλύτερου ενδιαφέροντός μας για το γνωστικό της πεδίο. Στην συνέχεια για να καταλήξουμε στα τελικά τέσσερα (4) άρθρα που θα μελετηθούν χρησιμοποιήσαμε τον ιστότοπο: <https://www.fileformat.info/tool/hash.htm>.

Στο πρώτο πεδίο, String hash, στο κελί Text εισάγαμε αλφαβητικά τα ονοματεπώνυμα μας όπως παρακάτω:

Text: ΑΛΕΞΑΚΗ ΜΑΝΤΕΣ

Επιλέγουμε στην συνέχεια τα 6 πρώτα ΔΙΑΚΡΙΤΑ ψηφία του παραπάνω αριθμού, δοθέντος ότι δεν υπάρχουν στην συλλογή άρθρα με αριθμό f ή e. Τα ψηφία αυτά θα μας δώσουν τα τελικά 4 άρθρα που θα μελετήσουμε.

Διακριτά ψηφία: 6,5,9,8,7,4

SHA-1: 065590a68c974c8d85c3f1c38c706eb736a698b9

Επομένως καταλήγουμε στα:

6 Freedom and privacy in Ambient Intelligence

5 Cambridge Analytica's black box

9 Big Data ethics

8 Algorithmic amplification of politics on Twitter

7 The ethics of algorithms: key problems and solutions

4 What is Überveillance?(And What should be done about it) ?

Πίνακας 1.1:

Αριθμός Άρθρου	Τίτλος Άρθρου
6	Freedom and privacy in Ambient Intelligence
5	Cambridge Analytica's black box
9	Big Data ethics
8	Algorithmic amplification of politics on Twitter
7	The ethics of algorithms: key problems and solutions
4	What is Überveillance?(And What should be done about it?)

1.2 Ομάδα

Λειτουργώντας ως ομάδα, σκοπός μας ήταν από την αρχή η ισότιμη κατανομή αρμοδιοτήτων στα δύο μέλη και η υλοποίηση τους εντός του χρονικού περιθωρίου που είχε τεθεί. Για την ευκολότερη διαπεραίωση της εργασίας και την τήρηση συνέπειας δημιουργήσαμε ένα χρονοδιάγραμμα σύμφωνα με τις ανάγκες της εργασίας αλλά και τον συνολικό χρόνο που μπορούσε να αφιερώσει το κάθε μέλος.

Αρχικά, όσον αφορά την μελέτη των άρθρων, το κάθε μέλος ήταν υπεύθυνο να μελετήσει 2 από τα 4 άρθρα και να εντοπίσει τα πιο σημαντικά σημεία του. Έτσι είναι ευκολότερο και για το άλλο μέλος να τα κατανοήσει. Επιπλέον αρμοδιότητα μας ήταν ο καθένας να γράψει περιληπτικά ένα κείμενο σύντομης έκτασης που να περιγράφει το κεντρικό νόημα αλλά και πως συνδέονται νοηματικά τα δύο (2) άρθρα που είχε αναλάβει να μελετήσει. Στη συνέχεια κάναμε συνάντηση όπου και συνθέσαμε αρχικά την βιβλιογραφική παρουσίαση των τεσσάρων (4) άρθρων, μελετώντας από κοινού ξανά τα σημαντικά μέρη τους και παράλληλα παραθέτοντας τις δικές μας προσωπικές σημειώσεις. Στην συνέχεια, έγινε η συζήτηση με το ChatGPT όπου παραθέσαμε τις απαντήσεις του και σχολιάσαμε την εγκυρότητα του.

Για την σύνθεση της εργασίας έκανε ο καθένας μας μία προσωπική μελέτη του LaTeX, καθώς ήταν ένα εργαλείο με το οποίο ήρθαμε σε επαφή για πρώτη φορά. Έτσι, ήταν πιο εύκολο να συνθέσουμε την εργασία, αφού ο καθένας μπορούσε να συμβάλει στην σύνταξη της χωρίς να υπάρχουν περιορισμοί. Εξαιρετικά χρήσιμη αποδείχθηκε η δυνατότητα άμεσης διόρθωσης λαθών που προσφέρει το συγκεκριμένο περιβάλλον. , καθώς και την συζήτηση με το ChatGPT στο LaTeX. Τέλος, μελετήσαμε την παρουσίαση σε PowerPoint και δημιουργήσαμε το βίντεο για την παρουσίαση του κειμένου που είχαμε επιλέξει.

Κεφάλαιο 2

Περίληψη(195)

Κάνοντας μια ανασκόπηση της Βιβλιογραφικής Παρουσίασης που παρατίθεται πιο κάτω, προκύπτει ότι η κεντρική ιδέα που πραγματεύονται τα τέσσερα (4) άρθρα δεν είναι άλλη από την ανησυχία για τις μεθόδους που χρησιμοποιούνται προκειμένου να στοχεύεται και να κατευθύνεται μεγάλο κοινό ψηφοφόρων μέσω αλγορίθμων. Χαρακτηριστικό παράδειγμα αποτελεί το συμβάν με την Cambridge Analytica που βασίστηκε σε δεδομένα που λάμβανε από χρήστες του Facebook. Αντίστοιχα, για το Twitter υπήρξε έντονη συζήτηση σχετικά με την πιθανότητα ορισμένες πολιτικές ομάδες να επωφεληθούν περισσότερο από άλλες, εξαιτίας της χρήσης αλγορίθμων που δημιουργούν πολιτικό προφίλ για τον κάθε χρήστη με βάση τις πληροφορίες που δίνει. Κατατάσσοντας σταθερά ορισμένα περιεχόμενα υψηλότερα από κάποια άλλα, αυτοί οι αλγόριθμοι μπορούν να ενισχύσουν την ορατότητα ορισμένων μηνυμάτων αυτής της πολιτικής, μειώνοντας παράλληλα των άλλων. Παράλληλα με την εκθετική ανάπτυξη και εφαρμογή αλγορίθμων μηχανικής μάθησης, έχουν προταθεί νέα ηθικά προβλήματα και λύσεις σχετικά με την πανταχού παρούσα χρήση τους στην κοινωνία. Είναι γεγονός λοιπόν, πως η φύση των Big Data δημιουργεί νέες μορφές εξουσίας και ηθικής ευθύνης, διαμορφώνοντας εκ νέου τον τρόπο με τον οποίο αντιλαμβανόμαστε την ελεύθερη βούληση. Πλέον, η ψηφιακή καταγραφή της ζωής μας δημιουργεί νέες προκλήσεις στον τομέα του απορρήτου.

Κεφάλαιο 3

Βιβλιογραφική Παρουσίαση(1.061)

3.1 Βιβλιογραφική Ανασκόπηση Άρθρων

Υπάρχει άραγε αρκετή ασφάλεια των προσωπικών μας δεδομένων και πεποιθήσεων στον κυβερνοχώρο και ειδικά στα social media? Η άραγε η ραγδαία άνοδος της τεχνητής νοημοσύνης έχει επιτρέψει την πλήρη καταπάτηση του δικαιώματος της ιδιωτικότητας, της ελεύθερης βούλησης καθώς και την διαμόρφωση τάσεων για κατανάλωση πολιτικού περιεχομένου από το κοινό;

Σύμφωνα με το άρθρο 5, με αφορμή το σκάνδαλο Cambridge Analytica -Facebook του 2018, ανεβαίνει στο προσκήνιο η εκμετάλλευση δεδομένων χρηστών για πολιτική στόχευση, όπως συνέβη στις προεδρικές εκλογές των ΗΠΑ το 2016. Εκατομμύρια δεδομένα χρηστών συλλέχθηκαν χωρίς την συναίνεσή τους, μέσω της φαινομενικά αθώας εφαρμογής 'thisisyourdigitallife' στην πλατφόρμα. Οι χρήστες εν αγνοία τους, επέτρεπαν την πρόσβαση σε προσωπικά στοιχεία και επαφές, τα οποία χρησιμοποιήθηκαν για την ψυχογραφική στόχευση ψηφοφόρων (Electronic Privacy Information Center, 2019· Γκρανβίλ, 2018). Η Cambridge Analytica δημιούργησε εξαιρετικά στοχευμένα πολιτικά μηνύματα, προσαρμοσμένα φυσιικά, στις προσωπικότητες και τα ενδιαφέροντα αυτών. Η FTC επέβαλλε ποινή ύψους 5 δισεκατομμυρίων δολαρίων στο Facebook, μια κίνηση που κρίθηκε από τους εμπειρογνώμονες, οι οποίοι εξέφρασαν την ανησυχία τους για την αποτελεσματικότητα του οργανισμού στην ουσιαστική προστασία των πολιτών. Ο Ντέιβιντ Βλάντεκ σε απάντηση αυτών, υποστηρίζει πως οι περιορισμένες εξουσίες, πόρους και νομοθετικές ελλείψεις που διαθέτει η FTC την εμπόδισαν να ερευνήσει πλήρως τις πιο ανησυχητικές και σημαντικές πτυχές της υπόθεσης. Οι ειδικοί προτείνουν διάφορες μεταρρυθμίσεις για την ενίσχυση της αποτελεσματικότητας της FTC, συμπεριλαμβανομένης της αύξησης της χρηματοδότησης και της παροχής περισσότερων πόρων, όπως η πρόσληψη ειδικών (Barrett, 2019· Hoofnagle κ.ά., 2019), καθώς και την επέκταση των νομοθετικών εξουσιών της FTC για την καθορισμό και επιβολή παραβιάσεων της ιδιωτικότητας (Cicilline and McSweeney, 2018· Χαν,

2016· Pasquale, 2013). Το σκάνδαλο της Cambridge Analytica καταδεικνύει την επείγουσα ανάγκη για αυστηρότερη νομοθεσία και αυξημένη ευθύνη των εταιρειών στη διαχείριση των δεδομένων, αναδεικνύοντας τις αδυναμίες της τρέχουσας ρύθμισης και την ανάγκη για αποτελεσματικότερη προστασία της ιδιωτικότητας των χρηστών.

Μια παρόμοια οπτική αναλύεται και στο άρθρο 8 του David Laitin, (Department of Political Science, Stanford University, Stanford, CA) , το οποίο εξετάζει πώς οι αλγόριθμοι προσωποποίησης του Twitter ενδέχεται να ενισχύουν συγκεκριμένα πολιτικά μηνύματα και να μειώνουν την ορατότητα άλλων (political bias). Συγκεκριμένα, χρησιμοποιώντας δεδομένα από ένα εκτεταμένο χρονικά πείραμα με σχεδόν 2 εκατομμύρια ενεργούς χρήστες, έγινε αναλυτική μελέτη της διανομής περιεχομένου χωρίς αλγοριθμική προσωποποίηση. Τα ευρήματα δείχνουν ότι σε έξι από τις επτά χώρες που μελετήθηκαν, η αλγοριθμική ενίσχυση ευνοεί τη mainstream πολιτική δεξιά σε σχέση με την mainstream πολιτική αριστερά. Αυτό ισχύει και για τις ειδησεογραφικές πηγές στις ΗΠΑ, όπου οι δεξιές πηγές ενισχύονται περισσότερο από τους αλγόριθμους. Οι ερευνητές δεν βρήκαν στοιχεία που να υποστηρίζουν την υπόθεση ότι οι αλγόριθμοι ενισχύουν περισσότερο τις ακραίες πολιτικές ομάδες (αριστερές ή δεξιές) σε σύγκριση με τις 'μετριοπαθείς'. Το Twitter μάλιστα χρησιμοποιούσε τέτοιους αλγορίθμους από το 2016, όταν ξεκίνησε να παρουσιάζει περιεχόμενο με βάση την αλγοριθμική προσωποποίηση αντί για τη χρονολογική σειρά που είχε γίνει η ανάρτηση του συγκεκριμένου περιεχομένου. Οι αλγόριθμοι αυτοί φάνηκε πως καθορίζουν σε μεγάλο βαθμό την προτεραιότητα κάποιων tweets με βάση τα χαρακτηριστικά του περιεχομένου και πως μπορεί να συνδέεται αυτό με τις πολιτικές πεποιθήσεις των χρηστών. Η μελέτη αυτή προσφέρει νέα δεδομένα στη συζήτηση για την αλγοριθμική μεροληψία και την πολιτική ενίσχυση, επισημαίνοντας τη συστημική τάση ενίσχυσης της πολιτικής δεξιάς στα μέσα κοινωνικής δικτύωσης.

Το άρθρο 9 αναφέρει πως η ραγδαία ανάπτυξη των Μεγάλων Δεδομένων και των μέσων κοινωνικής δικτύωσης, έχει υπερβεί την ικανότητα του μέσου χρήστη να αντιληφθεί τις συνέπειες των ενεργειών του (Zwitter, 2014). Ακόμα μία ηθική πρόκληση ξεκινά από το Facebook μετά το περιστατικό στις 21 Σεπτεμβρίου 2012 και αυτή τη φορά αφορά και ανηλίκους (BBC, 2012). Δεδομένες έως τώρα ηθικές αρχές επαναπροσδιορίζονται καθώς οι αποφάσεις δεν λαμβάνονται πλέον μόνο συνειδητά, αλλά φαίνονται να έχουν χειραγωγηθεί από εξωτερικούς παράγοντες. Τα Big Data λοιπόν, εισάγουν το πρόβλημα της 'κατανεμημένης ευθύνης', όπου πολλοί εμπλέκονται σε ένα αποτέλεσμα, καθιστώντας την ευθύνη πιο περίπλοκη. Η ίδια η φύση τους δημιουργεί νέες μορφές εξουσίας και ηθικής ευθύνης, διαμορφώνοντας εκ νέου τον τρόπο με τον οποίο αντιλαμβανόμαστε την ελεύθερη βούληση. Η ψηφιακή καταγραφή της ζωής μας δημιουργεί νέες προκλήσεις στον τομέα του απορρήτου. Με την ανάλυση των δεδομένων, μπορούν να ανακαλυφθούν κρυφές συσχετίσεις και να εφαρμοστούν στοχευμένες δράσεις για την προαγωγή ή την αποθάρρυνση συγκεκριμένων συμπεριφορών. Αντίστροφα, ο Zwitter αναφέρει στρατηγικές υπερσυνδεσιμότητας με χαρακτηριστικό παράδειγμα τις ψεύτικες συζητήσεις των Twitter bots, μία πρακτική από ονομάστηκε Astroturfing. Υπάρχει μάλιστα πιθανότητα ένα τέτοιο bot δημιουργώντας μια εκστρατεία δυσφήμισης να κατάφερε όντως να αποφασίσει στις εκλογές του 2010 στη Μασαχουσέτη. Χαρα-

κτηριστικό παράδειγμα επίσης αποτελεί η ταινία *Minority Report* που προειδοποιεί για ένα μέλλον όπου οι αναλύσεις Big Data προβλέπουν εγκληματικές δραστηριότητες. Συμπερασματικά, οι κανόνες δεοντολογίας στην έρευνα δεν είναι αρκετά σαφείς για τη χρήση δεδομένων από κοινωνικά δίκτυα. Η προστασία της ιδιωτικής ζωής και η συναίνεση των συμμετεχόντων είναι ζητήματα που πρέπει να εξεταστούν προσεκτικά.

Τέλος, όλες αυτές οι ηθικές ανησυχίες οργανώνονται σε έναν εννοιολογικό χάρτη, ο οποίος παρουσιάζεται στο άρθρο 7, από τους Brent Mittelstadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter και Luciano Floridi. Συγκεκριμένα, εξετάζονται τα ηθικά ζητήματα που προκύπτουν από τη χρήση των αλγορίθμων σε διάφορους τομείς της κοινωνίας και παράλληλα οι συγγραφείς παρέχουν μια ολοκληρωμένη ανάλυση των προβλημάτων και των πιθανών λύσεων. Πρώτα, θίγεται το ζήτημα ότι τα δεδομένα που χρησιμοποιούνται από αλγόριθμους μπορεί να είναι ασαφή, με ελλείψεις ή λανθασμένα, οδηγώντας σε ανακριβή αποτελέσματα, με αποτέλεσμα η διαδικασία με την οποία οι αλγόριθμοι λαμβάνουν αποφάσεις να είναι συχνά αδιαφανής, καθιστώντας δύσκολη την κατανόηση και τον έλεγχο των αποτελεσμάτων. Εξαιτίας αυτού του μοντέλου, οι αλγόριθμοι μπορεί να παράγουν άδικα ή μεροληπτικά αποτελέσματα, επηρεάζοντας διαφορετικές ομάδες ανθρώπων με άνισο τρόπο. Στη συνέχεια, Οι συγγραφείς συζητούν τις ηθικές διαστάσεις των αλγορίθμων, εστιάζοντας στις έννοιες της αυτονομίας και της ιδιωτικότητας. Συγκεκριμένα, καταλήγουν στο ότι οι αλγόριθμοι μπορούν να περιορίσουν την αυτονομία των ατόμων εάν οι αποφάσεις που λαμβάνονται αυτόματα δεν επιτρέπουν τη συμμετοχή και τον έλεγχο από τους χρήστες, ενώ τονίζουν πως η συλλογή και επεξεργασία μεγάλων ποσοτήτων δεδομένων προσωπικού χαρακτήρα θέτει σοβαρές ανησυχίες για την ιδιωτικότητα. Τέλος, προτείνεται μια σειρά από λύσεις που εστιάζουν στην επαρκή παροχή πληροφοριών για το πώς λειτουργούν οι αλγόριθμοι και τα δεδομένα που χρησιμοποιούν, στην δημιουργία μηχανισμών λογοδοσίας για τις αλγοριθμικές αποφάσεις και στην ενσωμάτωση των ενδιαφερόμενων μερών στη διαδικασία σχεδίασης και ανάπτυξης αλγορίθμων. Οι συγγραφείς καταλήγουν ότι η ανάπτυξη δίκαιων και ηθικών αλγοριθμικών συστημάτων μπορεί να εξασφαλίσει την εξυπηρέτηση των συμφερόντων των χρηστών και της κοινωνίας.

3.2 Συζήτηση

Συνοψίζοντας, τα άρθρα που μας δόθηκαν θα πρέπει και θα ήταν καλό να συζητηθούν τόσο από φοιτητές όσο και από καθηγητές στις αίθουσες των πανεπιστημίων. Πιο συγκεκριμένα, τα άρθρα αυτά μπορούν να αποτελέσουν πηγή μάθησης και στοχασμού σχετικά με το πόσο εύκολα η τεχνητή νοημοσύνη μπορεί να κατευθύνει τις απόψεις μας. Ο κλάδος της τεχνητής νοημοσύνης στο τμήμα των Μηχανικών Η/Υ Και Πληροφορικής είναι βασικός και αποτελείται από μαθήματα που είναι απαραίτητα στην επιστήμη με την οποία ασχολούμαστε. Ως εκ τούτου, μπορεί να βοηθήσει τους καθηγητές ώστε να έχουν μια ευρεία και μεγάλη γκάμα σχετικά με τα εργαλεία που μπορούν να δώσουν στους φοιτητές. Αλλά και οι φοιτητές να έρθουν πιο κοντά στο αντικείμενο και την επιστήμη την οποία σπουδάζουν. Είναι απαραίτητο λοιπόν να συζητηθούν διάφορα άρθρα στις πανεπιστημιακές σχο-

λές και να γίνονται κατάλληλες ομιλίες και ενημερώσεις συστηματικά περί των συγκεκριμένων ζητημάτων.

3.3 Κριτική συζήτηση με το ChatGPT

Για την κριτική συζήτηση με το ChatGPT επιλέξαμε το 3ο από τα άρθρα, δηλαδή το άρθρο 9: Big Data ethics Παρακάτω παραθέτουμε τις ερωτήσεις που θέσαμε και τις αντίστοιχες απαντήσεις που παρήγαγε.

Ερώτηση 1

Εγώ: Δώσε μια σύντομη περίληψη σχετικά με το άρθρο *Big Data ethics* By *Andrej Zwitter*.

Chat:

Το άρθρο επισημαίνει τις ηθικές προκλήσεις που σχετίζονται με την ιδιωτικότητα, την ομαδική ιδιωτικότητα, τις προβλέψεις και την ηθική της έρευνας. Αναφέρει ότι η χρήση των Big Data μπορεί να οδηγήσει σε στοχευμένη διαφήμιση, πολιτική χειραγώγηση και προληπτική αστυνόμευση. Επίσης, συζητά τις αλλαγές στην κατανομή της εξουσίας που προκύπτουν από την υπερσυνδεδεμένη κοινωνία και πώς αυτό επηρεάζει την ηθική ευθύνη. Τέλος, τονίζει ότι οι ηθικολόγοι πρέπει να επανεξετάσουν τις παραδοσιακές έννοιες της ηθικής και της ατομικής ευθύνης σε μια εποχή όπου τα δεδομένα και οι τεχνολογίες δικτύωσης έχουν αλλάξει την φύση της εξουσίας και της γνώσης.

Σχολιασμός

Συγκρίνοντας την απάντηση που έδωσε το τσατ με το άρθρο που έχουμε ήδη διαβάσει, εύλογα συμπεράναμε ότι περιγράφει με ακρίβεια την κεντρική του ιδέα.

Ερώτηση 2

Εγώ: Συμφωνείς με την κεντρική ιδέα του άρθρου; Θεωρείς όντως ότι στην σημερινή κοινωνία έχει αλλάξει η ηθική των ανθρώπων εξαιτίας της τεχνητής νοημοσύνης;

Chat: Ναι, συμφωνώ. Η κεντρική ιδέα του άρθρου, ότι η εποχή των Big Data και της τεχνητής νοημοσύνης (TN) θέτει νέες ηθικές προκλήσεις και απαιτεί επανεξέταση των παραδοσιακών εννοιών της ηθικής και της ευθύνης, είναι εύλογη και σημαντική. Η τεχνητή νοημοσύνη έχει όντως αλλάξει πολλές πτυχές της κοινωνίας, και αυτό περιλαμβάνει και την ηθική συμπεριφορά και τις ηθικές αντιλήψεις των ανθρώπων. Υπάρχουν αρκετά σημεία που μπορούμε να εξετάσουμε:

Ιδιωτικότητα και προσωπικά δεδομένα: Η TN μπορεί να συλλέγει, να αναλύει και να προβλέπει τη συμπεριφορά των ανθρώπων σε πρωτοφανή κλίμακα. Αυτό θέτει σημαντικά ζητήματα ιδιωτικότητας και απαιτεί νέες προσεγγίσεις για την

προστασία των προσωπικών δεδομένων.

Ηθική των αλγορίθμων: Οι αλγόριθμοι που χρησιμοποιούνται στην ΤΝ μπορούν να περιέχουν προκαταλήψεις και να λαμβάνουν αποφάσεις που επηρεάζουν τις ζωές των ανθρώπων. Η διαφάνεια και η λογοδοσία των αλγορίθμων είναι κρίσιμα ζητήματα ηθικής.

Κοινωνική δικαιοσύνη: Η ΤΝ μπορεί να αναπαράγει ή να ενισχύσει υπάρχουσες ανισότητες. Για παράδειγμα, τα συστήματα προληπτικής αστυνόμευσης μπορεί να στοχεύουν δυσανάλογα συγκεκριμένες ομάδες πληθυσμού, οδηγώντας σε διακρίσεις.

Αλλαγές στην εργασία: Η αυτοματοποίηση και η χρήση της ΤΝ έχουν σημαντικές επιπτώσεις στην αγορά εργασίας, με την αντικατάσταση πολλών θέσεων εργασίας και την ανάγκη για νέες δεξιότητες. Αυτό θέτει ηθικά ζητήματα σχετικά με την οικονομική ανισότητα και την κοινωνική ευθύνη για την εκπαίδευση και την ανακατανομή των πόρων.

Ηθικές αποφάσεις: Σε τομείς όπως η ιατρική και οι μεταφορές, οι αλγόριθμοι ΤΝ καλούνται να λάβουν αποφάσεις που μπορεί να επηρεάσουν τη ζωή και την υγεία των ανθρώπων. Αυτό δημιουργεί ηθικά διλήμματα σχετικά με την ανθρωπινή επίβλεψη και την ευθύνη. Γενικά, η τεχνητή νοημοσύνη έχει επιφέρει σημαντικές αλλαγές στην κοινωνία και στις ηθικές αντιλήψεις. Η ανάγκη για επανεξέταση των παραδοσιακών εννοιών της ηθικής είναι προφανής, καθώς πρέπει να διασφαλιστεί ότι οι νέες τεχνολογίες χρησιμοποιούνται με τρόπους που είναι δίκαιοι, διαφανείς και ευθύνομενοι. Αυτό περιλαμβάνει τη δημιουργία νέων ηθικών πλαισίων και κανονιστικών ρυθμίσεων που να ανταποκρίνονται στις προκλήσεις της σύγχρονης εποχής.

Επομένως, είναι σημαντικό να αναγνωρίσουμε αυτές τις αλλαγές και να εργαστούμε προς την κατεύθυνση της ανάπτυξης πολιτικών και κανονισμών που θα διασφαλίσουν την ηθική χρήση της τεχνητής νοημοσύνης στην κοινωνία μας.

Σχολιασμός

Στην παραπάνω ερώτηση γίνεται κατανοητό ότι το ChatGPT υιοθετεί μια προσωπική θέση, καθώς συμφωνεί με την διαπίστωση που του θέτουμε. Παρατηρούμε ότι ως προϊόν του ανθρωπίνου νου, προσπαθεί μέσω αλγορίθμων να συντάξει μία απάντηση συλλέγοντας 'μηχανικά' δεδομένα σύμφωνα με την δική μας περιγραφή και να μιμνηθεί μία ανθρωπογενή συμπεριφορά. Ωστόσο, παρατηρούμε ότι δεν εκφράζει μια πλήρως απροσωποποιημένη και αντικειμενική άποψη, γεγονός που ίσως οφείλεται στην αναβαθμισμένη έκδοσή του την οποία χρησιμοποιούμε και που το καθιστά ακόμα πιο έξυπνο, φέρνοντας το ακόμα πιο κοντά στην ανθρωπινή λογική.

Ερώτηση 3

Εγώ: Ποιά πρακτική λύση μπορείς να προτείνεις ώστε να ελλατωθεί η χρήση της τεχνητής νοημοσύνης ως μέσο χειραγώγησης και προπαγάνδας;

Chat: Για να μειωθεί η χρήση της τεχνητής νοημοσύνης (TN) ως μέσο χειραγώγησης και προπαγάνδας, είναι απαραίτητο να αναπτυχθούν και να εφαρμοστούν πρακτικές λύσεις που περιλαμβάνουν ρυθμίσεις, εκπαίδευση και τεχνολογικές καινοτομίες. Μια συνολική προσέγγιση μπορεί να περιλαμβάνει τα εξής μέτρα:

1. Νομοθετικές Ρυθμίσεις και Κανονισμοί Δημιουργία αυστηρών κανονισμών: Εφαρμογή νόμων και κανονισμών που περιορίζουν τη χρήση της TN σε τομείς που ενέχουν κινδύνους χειραγώγησης, όπως τα μέσα κοινωνικής δικτύωσης και η πολιτική διαφήμιση. Διαφάνεια και λογοδοσία: Υποχρέωση των εταιρειών που αναπτύσσουν και χρησιμοποιούν TN να είναι διαφανείς σχετικά με τα δεδομένα και τους αλγόριθμους που χρησιμοποιούν. Επιβολή κανονισμών που απαιτούν την παροχή σαφών εξηγήσεων για τον τρόπο λήψης αποφάσεων από την TN.

2. Εκπαίδευση και Ενημέρωση του Κοινού Εκπαίδευση πολιτών: Εκστρατείες ενημέρωσης και εκπαιδευτικά προγράμματα για την ενίσχυση της ψηφιακής παιδείας, ώστε οι πολίτες να μπορούν να αναγνωρίζουν και να αντιστέκονται σε προσπάθειες χειραγώγησης. Εκπαίδευση επαγγελματιών TN: Ανάπτυξη εκπαιδευτικών προγραμμάτων για τους επαγγελματίες της TN που θα περιλαμβάνουν μαθήματα ηθικής και υπεύθυνης ανάπτυξης και χρήσης της τεχνολογίας.

3. Τεχνολογικές Καινοτομίες Αναγνώριση και πρόληψη χειραγώγησης: Ανάπτυξη αλγόριθμων και εργαλείων που μπορούν να ανιχνεύουν και να αποτρέπουν προσπάθειες χειραγώγησης, όπως ψευδείς ειδήσεις ή προκατειλημμένο περιεχόμενο. Ενίσχυση της ιδιωτικότητας: Εφαρμογή τεχνολογιών που προστατεύουν την ιδιωτικότητα των χρηστών και μειώνουν τη δυνατότητα συγκέντρωσης και κατάχρησης προσωπικών δεδομένων.

4. Δημιουργία Εποπτικών Αρχών Εποπτικές αρχές TN: Ίδρυση ανεξάρτητων εποπτικών αρχών που θα παρακολουθούν τη χρήση της TN, θα επιβλέπουν την εφαρμογή των κανονισμών και θα επιβάλλουν κυρώσεις σε περιπτώσεις κατάχρησης.

5. Συνεργασία και Διεθνής Συντονισμός Διεθνείς συμφωνίες και συνεργασίες: Συνεργασία μεταξύ κυβερνήσεων, διεθνών οργανισμών και εταιρειών τεχνολογίας για την ανάπτυξη και υιοθέτηση κοινών προτύπων και πρακτικών που θα περιορίζουν τη χρήση της TN για χειραγώγηση σε παγκόσμιο επίπεδο.

Η συνδυασμένη εφαρμογή αυτών των πρακτικών λύσεων μπορεί να συμβάλει στη μείωση της χρήσης της τεχνητής νοημοσύνης ως μέσο χειραγώγησης και στην προώθηση μιας πιο δίκαιης και διαφανούς κοινωνίας.

Σχολιασμός

Αναφορικά με το κείμενο και την απάντηση που δόθηκε παραπάνω καταληγουμε στο συμπέρασμα ότι το chatbot παρουσιάζει συγκεκριμένες λύσεις που έχουν κοινά χαρακτηριστικά με αυτές που παρουσιάζονται από τους συγγραφείς του άρθρου 7 και 9. Αυτό ενισχύει ακόμα περισσότερο την πειστικότητα και την εγκυρότητα της απάντησής του, καθώς αξιοποιεί πληροφορίες από επιβεβαιωμένα επιστημονικά άρθρα.

Τέλος, μπορούν να μελετηθούν και άλλες ενδιαφέρουσες πηγές για το θέμα, όπως τα άρθρα με αριθμό [4],[6].

Κεφάλαιο 4

Βιβλιογραφία

Στην παρακάτω σελίδα βρίσκονται όλες οι αναφορές της συλλογής (όλα τα 9 άρθρα της συλλογής), οι οποίες αποτελούν την βιβλιογραφία την οποία αξιοποιήσαμε στην καταγραφή των παραπάνω περιλήψεων.

Bibliography

- [1] V. Papakonstantinou, “States as information platforms: a political theory of information,” 2023.
- [2] K. Crawford, K. Miltner, and M. Gray, “Special section introduction,” *International Journal of Communication*, vol. 8, pp. 1663–1672, 2014.
- [3] R. McIlroy-Young, J. Kleinberg, S. Sen, S. Barocas, and A. Anderson, “Mimetic models: Ethical implications of ai that acts like you,” in *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, 2022, pp. 479–490.
- [4] R. Clarke, “What is überveillance?(and what should be done about it?),” *IEEE Technology and Society Magazine*, vol. 29, no. 2, pp. 17–25, 2010.
- [5] M. Hu, “Cambridge analytica’s black box,” *Big Data & Society*, vol. 7, no. 2, p. 2053951720938091, 2020.
- [6] P. Brey, “Freedom and privacy in ambient intelligence,” *Ethics and Information Technology*, vol. 7, pp. 157–166, 2005.
- [7] A. Tsamados, N. Aggarwal, J. Cows, J. Morley, H. Roberts, M. Taddeo, and L. Floridi, “The ethics of algorithms: key problems and solutions,” *Ethics, governance, and policies in artificial intelligence*, pp. 97–123, 2021.
- [8] F. Huszár, S. I. Ktena, C. O’Brien, L. Belli, A. Schlaikjer, and M. Hardt, “Algorithmic amplification of politics on twitter,” *Proceedings of the National Academy of Sciences*, vol. 119, no. 1, p. e2025334119, 2022.
- [9] A. Zwitter, “Big data ethics,” *Big data & society*, vol. 1, no. 2, p. 2053951714559253, 2014.

[1] [2] [3] [4] [5] [6] [7] [8] [9]

Κεφάλαιο 5

Βιογραφικά(216)

Ο Σωκράτης Μαντές είναι προπτυχιακός φοιτητής του Τμήματος ΜΗΥΠ του Πανεπιστημίου Πατρών. Κατά την φοίτησή του στο τμήμα έχει αποκτήσει σημαντική εμπειρία πάνω σε θεμελιώδη κομμάτια της επιστήμης των Η/Υ, έχοντας ταυτόχρονα αξιόλογη ακαδημαϊκή πορεία και συνέπεια στις υποχρεώσεις του προγράμματος σπουδών. Βασικές πηγές ενδιαφέροντός του είναι τομείς όπως οι Αρχές Ψηφιακών Ηλεκτρονικών, ο Λογικός Σχεδιασμός Κυκλωμάτων και τα Θέματα Υλικού. Κύριο μελλοντικό του στόχο αποτελεί η απόκτηση γνώσεων και η ενασχόληση με ζητήματα που αφορούν τον κλάδο των Δικτύων και της Κυβερνοασφάλειας. Δείχνει αφοσίωση στα προθέτες που συμμετέχει, προσαρμόζεται εύκολα στις αρμοδιότητες που αναλαμβάνει και επιλύει με ωριμότητα απροσδόκητα ζητήματα. Τέλος, διαθέτει κοινωνικές δεξιότητες όπως ομαδικότητα και άνεση στην επικοινωνία.



Η Βασιλική Αλεξάκη είναι προπτυχιακή φοιτήτρια του τμήματος Μηχανικών Η/Υ και Πληροφορικής του Πανεπιστημίου Πατρών. Από το 2021 έως σήμερα έχει αποκτήσει σημαντικές και σφαιρικές γνώσεις πάνω στους βασικούς τομείς της Επιστήμης των Υπολογιστών. Παρουσιάζει ιδιαίτερο ενδιαφέρον στα μαθήματα της Τεχνητής Νοημοσύνης και της Επεξεργασίας Σημάτων. Μελλοντικός στόχος είναι οι μεταπτυχιακές σπουδές στην Βιοϊατρική Μηχανική, κατόπιν η έρευνα και η εφαρμογή τεχνικών διαχείρισης δεδομένων και ανάλυσης Ιατρικών Σημάτων/Εικόνων. Η ακαδημαϊκή της πορεία χαρακτηρίζεται από αφοσίωση, οργάνωση και συνέπεια στην εκπόνηση των εργασιών. Η προσαρμοστικότητα και οι επικοινωνιακές δεξιότητες την καθιστούν ικανό και παραγωγικό μέλος μιας ομάδας αλλά και ικανή στον ρόλο ηγεσίας και λήψης αποφάσεων.

