

Προς μια Φεμινιστική Τεχνητή Νοημοσύνη: Τεχνο-Μύθοι, Πραγματικότητα και Εναλλακτικά Μέλλοντα

Χριστίνα Γραμματικοπούλου¹

Περίληψη

Μπορεί να υπάρξει μια Φεμινιστική Τεχνητή Νοημοσύνη; Με βάση αυτό το ερώτημα, το παρόν κείμενο οραματίζεται μια φεμινιστική Τεχνητή Νοημοσύνη, ως δύναμη αντίστασης στην αλγοριθμική καταπίεση. Αρχικά εξετάζονται αμφιλεγόμενες εφαρμογές της Τεχνητής Νοημοσύνης, που αντηχούν σεξιστικές, ρατσιστικές και αποικιοκρατικές θεωρήσεις, για να διερευνηθεί πώς μια φεμινιστική προσέγγιση στα συστήματα τεχνητής νοημοσύνης μπορεί να σκιαγραφήσει εναλλακτικές λύσεις. Μέσα σε αυτό το πλαίσιο, διερευνώνται καλλιτεχνικές πρακτικές που στηρίζονται σε φεμινιστικές αρχές και δείχνουν έναν διαφορετικό τρόπο προσέγγισης στην Τεχνητή Νοημοσύνη και στην επιστήμη των δεδομένων. Το κείμενο στηρίζεται θεωρητικά σε αναλύσεις σχετικά με την τοποθετημένη γνώση, τον μαύρο φεμινισμό ως δύναμη ενάντια στην αλγοριθμική καταπίεση, την αλγοριθμική αποικιοκρατία και τον φεμινισμό των δεδομένων.

Λέξεις-κλειδιά: τεχνητή νοημοσύνη, μαύρος φεμινισμός, αλγοριθμική καταπίεση, καλλιτεχνικές πρακτικές

Abstract

Can there be a feminist Artificial Intelligence? Based on this question, this text envisions a Feminist AI, as a counterforce against algorithmic bias and oppression. After a look into controversial applications of AI, that reverberate sexist, racist and colonialist views, it shows how a feminist approach to AI systems can outline alternatives through experimental artistic projects. Our theoretical tools build on analyses regarding situated knowledge, Black Feminism as a force against algorithmic oppression, algorithmic colonization and Data Feminism.

Keywords: artificial intelligence, black feminism, algorithmic oppression, artistic practices

¹ Διδακτορίσσα Ιστορίας της Τέχνης, επιστημονική συνεργάτης Τμήματος Βαλκανικών, Σλαβικών και Ανατολικών Σπουδών, Πανεπιστήμιο Μακεδονίας

Εισαγωγή¹

Πρώτα έρχεται η αισιοδοξία, μετά ο ρεαλισμός και έπειτα η προσπάθεια για αλλαγή. Είναι ένα σενάριο που έχουμε δει να παίζει πολλές φορές σε σχέση με τη νέα τεχνολογία τις τελευταίες δεκαετίες -πρώτα με το Διαδίκτυο, μετά τα Μέσα Κοινωνικής Δικτύωσης και τον τελευταίο καιρό με την Τεχνητή Νοημοσύνη. Η πρώτη ενθουσιώδης υποδοχή είναι ότι η τεχνολογική καινοτομία θα φέρει ισότητα, θα μειώσει τη φτώχεια, θα προωθήσει την εκπαίδευση και τη δημοκρατία. Στη συνέχεια έρχεται η συνειδητοποίηση ότι οι νέες τεχνολογίες υποτάσσονται άμεσα στην καθεστηκυία τάξη πραγμάτων, κάτι που στη συνέχεια πυροδοτεί τον αγώνα για διασφάλιση των κοινωνικών δικαιωμάτων υπό τη νέα πραγματικότητα. Τα αντανάκλαστα μας σε ό,τι αφορά τον αντίκτυπο της τεχνολογίας στην κοινωνία έχουν γίνει πιο γρήγορα, και, καθώς οι κριτικές προσεγγίσεις στην Τεχνητή Νοημοσύνη αρχίζουν να πληθαίνουν, ίσως να είναι η στιγμή να οραματιστούμε εναλλακτικές εφαρμογές της, στη βάση της φροντίδας και της αλληλεγγύης.

Το κείμενο αυτό θα εστιάσει στο όραμα για μια Φεμινιστική Τεχνητή Νοημοσύνη, ως εργαλείο ενάντια στην αλγοριθμική προκατάληψη και την καταπίεση. Αρχικά θα παρουσιαστούν εφαρμογές της Τεχνητής Νοημοσύνης που απηχούν σεξιστικές, ρατσιστικές και αποικιοκρατικές ιδέες, προκειμένου να διερευνηθεί πώς μια φεμινιστική προσέγγιση στα συστήματα τεχνητής νοημοσύνης μπορεί να σκιαγραφήσει εναλλακτικούς δρόμους ενάντια στην καταπίεση. Οι πειραματισμοί αυτοί με την Τεχνητή Νοημοσύνη πραγματοποιούνται με τη μορφή σύγχρονων καλλιτεχνικών πρακτικών, μερικές από τις οποίες θα παρουσιαστούν στην τελευταία ενότητα.

Τα τελευταία χρόνια υπάρχει μια πληθώρα κριτικών προσεγγίσεων στην Τεχνητή Νοημοσύνη, από τα γνωστικά αντικείμενα των Σπουδών Φύλου, της κριτικής φυλετικής θεωρίας, των σπουδών κουήρ και των μεταποικιακών σπουδών, μεταξύ άλλων. Η διαθεματική φεμινιστική προσέγγιση μας βοηθά να κατανοήσουμε τα πολλαπλά επίπεδα καταπίεσης που είναι ενσωματωμένα στην τεχνολογία, που σχετίζονται με το φύλο, την ταυτότητα φύλου, τη φυλή, την καταγωγή, την κοινωνική τάξη, την αναπηρία και τον σεξουαλικό προσανατολισμό (Baker 2018). Οι θεωρίες και τα καλλιτεχνικά έργα που εξετάζονται εδώ, εγγράφονται σε μια ευρύτερη συζήτηση σχετικά με τον τεχνοφεμινισμό, εστιάζοντας στη θεσιακότητα της Donna Haraway (1988), αλλά και την ανάλυση της Cornelia Sollfrank (2016) για το πώς η τεχνολογία και το φύλο παράγουν το ένα το άλλο· επίσης, λαμβάνουν υπόψη τον διαδικτυακό φεμινιστικό ακτιβισμό και τις προσπάθειες να γίνει επαναπροσδιορισμός του σχεδιασμού αλληλεπίδρασης ανθρώπου-υπολογιστή (HCI), της σχεδιαστικής σκέψης (design thinking) και της επιστήμης των υπολογιστών (computer science) με βάση φεμινιστικές αρχές (Bardzell 2010).

Εγγενής προκατάληψη ή καταπίεση;

Εδώ και δεκαετίες τώρα, η μαζική κουλτούρα (ταινίες, τηλεοπτικά προγράμματα) αναπαράγει το στερεότυπο ότι η Τεχνητή Νοημοσύνη αναφέρεται σε πολύπλοκα υπολογιστικά συστήματα, ανθρωποειδή στην εμφάνιση, αλλά πιο έξυπνα από τα ανθρώ-

¹ Το κείμενο αυτό βασίζεται σε μια προηγούμενη δημοσίευσή μου (Grammatikopoulou 2022), η οποία είχε πραγματοποιηθεί στο πλαίσιο της μεταδιδακτορικής μου έρευνας, υπό την επίβλεψη της καθηγήτριας Φωτεινής Τσιμπιρίδου στο Τμήμα Βαλκανικών, Σλαβικών και Ανατολικών Σπουδών του Πανεπιστημίου Μακεδονίας. Μέσα στα δύο χρόνια που μεσολάβησαν από την ολοκλήρωση του αρχικού κειμένου (το 2021), έχει επεκταθεί σημαντικά η βιβλιογραφία που εξετάζει την Τεχνητή Νοημοσύνη μέσα από μια φεμινιστική και μεταποικιακή προσέγγιση και έχουν λάβει χώρα ραγδαίες τεχνολογικές εξελίξεις στο πεδίο, όπως το ChatGPT. Για το λόγο αυτό, προχώρησα σε μια ριζική αναθεώρηση του αρχικού κειμένου, ώστε να αντανάκλα την τρέχουσα κατάσταση στο πεδίο της Φεμινιστικής Τεχνητής Νοημοσύνης και το εμπλούτισα με πιο σύγχρονη βιβλιογραφία.

πινα όντα, και ικανά να λαμβάνουν ορθολογικές αποφάσεις. Έτσι, τα συστήματα τεχνητής νοημοσύνης παρουσιάζονται ως αυτοματοποιημένα, ανέγγιχτα από ανθρώπινο χέρι και ως εκ τούτου απρόσβλητα από την ανθρώπινη προκατάληψη. Οι ενδοιασμοί του κοινού για την Τεχνητή Νοημοσύνη είναι μονάχα φόβοι που προέρχονται από το πεδίο της κινηματογραφικής επιστημονικής φαντασίας, ότι θα δημιουργηθούν ανεξάρτητα ρομπότ που θα υποδουλώσουν την ανθρωπότητα, συνεπώς όλη η κουβέντα στα Μέσα Μαζικής Ενημέρωσης και τα Μέσα Κοινωνικής Δικτύωσης εξαντλείται σε φανταστικά σενάρια. Όμως, όπως θα δούμε, οι κίνδυνοι της τεχνολογίας είναι πολύ διαφορετικοί και διαφαινονται ήδη.

Στην πραγματικότητα, τα συστήματα Τεχνητής Νοημοσύνης είναι απλά εργαλεία που λαμβάνουν συγκεκριμένες αποφάσεις με βάση τα δεδομένα που τους έχουν δοθεί και τον τρόπο με τον οποίο έχουν οργανωθεί. Αυτή η παρανόηση αντανάκλα κάθε αντίληψη για επιστημονική και τεχνολογική αντικειμενικότητα που αγνοεί τα συστήματα εξουσίας που υπάρχουν από πίσω. Το επίμονο στερεότυπο της τεχνολογικής αντικειμενικότητας είναι επιζήμιο, επειδή διαιωνίζει τις διακρίσεις, αλλά και αποθαρρύνει την διερεύνηση των ψηφιακών ηγεμονιών σε βάθος (Haraway 1988).

Αντί λοιπόν για ένα πανίσχυρο μηχανήμα που γνωρίζει τα πάντα και αναπαράγει μια «αντικειμενική» άποψη του κόσμου, τα συστήματα Τεχνητής Νοημοσύνης είναι συγκεκριμένες εφαρμογές, που επιλύουν προβλήματα όπως η αυτοματοποιημένη αναγνώριση προσώπων, η αναγνώριση φωνής, η κατανόηση κειμένου, η αναπαραγωγή ομιλίας και δεκάδες άλλα παραδείγματα. Εν συντομία, η Τεχνητή Νοημοσύνη ξεκινάει με την ανάλυση ενός μεγάλου όγκου δεδομένων, που οδηγεί σε αναγνώριση πιθανοτήτων και μοτίβων, και τελικά στη δημιουργία ενός υπολογιστικού μοντέλου, που χρησιμοποιείται για να δώσει μια απάντηση σε ένα ερώτημα, με βάση τα όσα έχει «μάθει» από τα δεδομένα που του έδωσαν (Broussard 2018, σ. 32). Όμως, καθώς τα δεδομένα πάνω στα οποία μαθαίνουν τα συστήματα, καθώς και η οργάνωση των συστημάτων, αντανάκλουν υπάρχουσες κοινωνικές ανισότητες. Έτσι μπορεί να βρει κανείς δείγματα τεχνολογικής προκατάληψης σε όλα σχεδόν τα συστήματα Τεχνητής Νοημοσύνης. Γι' αυτό το λόγο, η κριτική προσέγγιση της Τεχνητής Νοημοσύνης ξεκινάει από άτομα που έχουν βιώσει αυτή την καταπίεση. Όπως τονίζει η Tourin (2023), χάρη στο έργο μαύρων φεμινιστριών ερευνητριών που εστίασαν στο θέμα αυτό, έχει πλέον ανοίξει ο δρόμος για μια διερεύνηση της Αποαποικιοποιημένης Τεχνητής Νοημοσύνης (Decolonized AI), της Τεχνητής Νοημοσύνης των Ιθαγενών (Indigenous AI) και της Φεμινιστικής Τεχνητής Νοημοσύνης (Feminist AI). Προτού όμως περάσουμε σε παραδείγματα που να στοιχειοθετούν τις προοπτικές αυτές, είναι χρήσιμο να δούμε πώς εκδηλώνεται η μεροληπτική αντιμετώπιση μειονοτικών ομάδων από τα συστήματα Τεχνητής Νοημοσύνης.

Κάποιες φορές τα μοντέλα Τεχνητής Νοημοσύνης κάνουν λάθη επειδή εκπαιδεύτηκαν πάνω σε ελλιπή δεδομένα, που δεν έλαβαν υπόψη κοινωνικές ομάδες που είναι «αόρατες». Για παράδειγμα, στον τομέα της όρασης υπολογιστών (computer vision), έχουμε την περίπτωση της Joy Buolamwini, που ερευνούσε συστήματα αναγνώρισης προσώπου στο MIT, όταν ανακάλυψε ότι ο αλγόριθμος δεν μπορούσε να δει το πρόσωπό της. Αντίθετα, το ίδιο σύστημα ήταν ικανό να αναγνωρίσει ως ανθρώπινο πρόσωπο μια λευκή μάσκα ή το σκίτσο ενός χαμογελαστού προσώπου πάνω στην παλάμη του χεριού της. Αυτή η «τυφλότητα» απέναντι στις μαύρες γυναίκες οφειλόταν στο γεγονός ότι ο αλγόριθμος είχε εκπαιδευτεί με σύνολα δεδομένων που αποτελούνταν κυρίως από φωτογραφίες λευκών ανδρών (Buolamwini 2023). Παρομοίως, η δημιουργία ενός εργαλείου Τεχνητής Νοημοσύνης για την αυτοματοποίηση στη διαδικασία επιλογής υποψηφίων εργαζόμενων, κατέδειξε την ίδια ανεπάρκεια. Το 2018, η Amazon χρειάστηκε να καταργήσει ένα πολυδάπανο πρότζεκτ αυτοματοποίησης, που θα βοηθούσε στην εύρεση των πιο κατάλληλων υποψηφίων ανάμεσα σε χιλιάδες βιογραφικά, όταν

έγινε σαφές ότι ο αλγόριθμος απέρριψε όλες τις γυναίκες υποψήφιες. Καθώς ο αλγόριθμος είχε εκπαιδευτεί με βάση τα δεδομένα που αφορούσαν το προσωπικό στην εταιρεία, που ήταν κυρίως άνδρες, έκρινε τις γυναίκες υποψήφιες ως ακατάλληλες (Franzoni 2023).

Κάποιες άλλες φορές οι εφαρμογές εκπαιδεύονται πάνω στην προκατάληψη (Noble 2018). Ένα χαρακτηριστικό τέτοιο παράδειγμα στον τομέα της Επεξεργασίας Φυσικής Γλώσσας (Natural Language Processing) είναι η *Tay* της Microsoft (2016), ένα αυτοματοποιημένο προφίλ στο Twitter που διαχειριζόταν ένα social bot (μποτ Μέσων Κοινωνικής Δικτύωσης). Ο στόχος του πρότζεκτ ήταν η κατανόηση της γλώσσας και ο διάλογος με άλλους χρήστες. Ωστόσο, το προφίλ κατέβηκε όταν εκπαιδεύτηκε μέσα σε λιγότερο από 24 ώρες στη χρήση αντισημιτικής, σεξιστικής και ρατσιστικής γλώσσας από άλλους χρήστες του Twitter (Wolf, Miller και Grodzinsky 2017). Το ίδιο συμβαίνει και σε εφαρμογές ανάλυσης συναισθήματος (sentiment analysis) που εκπαιδεύονται να αναγνωρίζουν τη διάθεση ενός ατόμου από την έκφραση του προσώπου. Οι εφαρμογές αυτές συχνά χαρακτηρίζουν άτομα Ασιατικής καταγωγής ως «κοιμισμένα» και μαύρους άνδρες ως «θυμωμένους», επειδή τα συστήματα εκπαιδεύτηκαν από άτομα που έβλεπαν τις διαφορετικές φυλές μέσα από στερεότυπα. Για τον ίδιο λόγο, οι βοηθοί φωνής (voice assistants) συνήθως έχουν γυναικεία φωνή και απαντούν με υποτακτικό τόνο, ακόμα και με φλερτ, σε υβριστικά και σεξιστικά σχόλια (UNESCO και EQUALS Skills Coalition 2019). Οι αλγόριθμοι αναζήτησης διαδίδουν τα στερεότυπα της σεξουαλικότητας των εφήβων κοριτσιών και της παραβατικότητας των μαύρων εφήβων -όπως φαίνεται στα αντίστοιχα αποτελέσματα αναζήτησης κοριτσιών με σέξι ρούχα και μαύρων αγοριών που κρατούν όπλα (D'Ignazio και Klein 2020). Αντίθετα, η αναζήτηση για λευκούς εφήβους δίνει στα πρώτα αποτελέσματα την καθιερωμένη εικόνα της εφηβείας: δηλαδή τα λευκά αγόρια, τα οποία αποτυπώνονται από τον αλγόριθμο ως η «κανονική» εικόνα, ενώ οι άλλες ως «διαφοροποιήσεις» ή «παρεκβάσεις» από τον κανόνα.

Αυτά τα παραδείγματα δεν είναι μεμονωμένα. Στις εφαρμογές Τεχνητής Νοημοσύνης ακολουθείται το ίδιο μοτίβο ξανά και ξανά: δημιουργούνται συστήματα που «προβλέπουν» απαντήσεις με βάση τα δεδομένα που τροφοδοτούνται σε αυτά, τα οποία δεδομένα αντανακλούν κοινωνικές ανισότητες, δημογραφικούς διαχωρισμούς και κοινωνικές διακρίσεις. Αυτό συμβαίνει επειδή οι εφαρμογές αυτές έχουν σχεδιαστεί, κατασκευαστεί και εφαρμοστεί από πάνω προς τα κάτω, από μια θέση ισχύος που παγιώνεται ακόμα περισσότερο μέσα από τις νέες τεχνολογίες. Αντίστροφα, οι μειονοτικές ομάδες περιθωριοποιούνται περαιτέρω, μέσα από την αναπαραγωγή στερεοτύπων, την ενίσχυση της αορατότητάς τους, σε ζητήματα όπου ήδη παραβλέπονται (για παράδειγμα σε θέματα εκπαιδευτικής πολιτικής, πολιτιστικής πολιτικής, τεχνολογικού και αστικού σχεδιασμού) ή την υπερ-εκπροσώπησή τους, όταν οι τεχνολογίες αφορούν την προληπτική αστυνόμευση, την επιτήρηση και τον σωφρονισμό (ό.π.). Όπως φάνηκε από τα παραδείγματα που δόθηκαν πιο πάνω, οι αλγόριθμοι διαιωνίζουν τις υπάρχουσες προκαταλήψεις, καθώς και τα συστήματα καταπίεσης που αφορούν συγκεκριμένες κοινωνικές ομάδες.

Πίσω από τα συστήματα Τεχνητής Νοημοσύνης υπάρχει ένα εργατικό δυναμικό με εμφανή ταξική διαστρωμάτωση: συνήθως άτομα που ανήκουν στις προνομιούχες κοινωνικές ομάδες είναι αυτά που λαμβάνουν τις σημαντικές αποφάσεις που αφορούν την τεχνολογική καινοτομία, σε ό,τι αφορά το σχεδιασμό και την εφαρμογή των συστημάτων, ενώ τις πιο μονότονες και επαναληπτικές εργασίες εκτελούν άτομα σε επισφαλείς εργασιακές συνθήκες, που πληρώνονται ελάχιστα για κάθε ώρα εργασίας σε αυτοματοποιημένες πλατφόρμες (όπως για παράδειγμα το Mechanical Turk της Amazon). Η Catherine D'Ignazio και η Lauren Klein (ό.π.) μιλούν για τον «κίνδυνο του προνομίου», δηλαδή την αδυναμία των ανθρώπων που έχουν καλή εκπαίδευση,

υψηλές αποδοχές και υψηλό κοινωνικό υπόβαθρο να αναγνωρίσουν περιπτώσεις καταπίεσης, λόγω έλλειψης βιωμένης εμπειρίας. Έτσι, όταν προκύπτουν προβλήματα που αναπαράγουν τις κοινωνικές ανισότητες, οι λύσεις που έχουν να προτείνουν συνήθως έχουν τη μορφή του βιαστικού φιζαρίσματος: για παράδειγμα προσπαθούν να «διορθώσουν» τις βάσεις δεδομένων για να δώσουν καλύτερα αποτελέσματα, αντί να τις χτίσουν εκ νέου, λαμβάνοντας υπόψη τις φωνές των κοινωνικών ομάδων που επηρεάζονται από την προκατάληψη. Επιπλέον, τα προβλήματα δεν βρίσκονται μόνο στη συλλογή δεδομένων. Ολόκληρο το σύστημα της Τεχνητής Νοημοσύνης, από τα δεδομένα που χρησιμοποιούνται, μέχρι τους ανθρώπους που το δημιουργούν και, κυρίως, τις οικονομικές και πολιτικές δυνάμεις που καθορίζουν τα χαρακτηριστικά του, θα ήταν ανάγκη να αναδιαρθρωθεί προκειμένου να διαφοροποιηθεί η υπάρχουσα δυναμική. Στην τελευταία ενότητα αυτού του κειμένου θα αναφερθούμε σε ορισμένα καλλιτεχνικά και ακτιβιστικά πρότζεκτ, που πειραματίζονται με τέτοιες εφαρμογές, ιδωμένες μέσα από μια φεμινιστική οπτική.

Το πρόβλημα με την ανάπτυξη της τεχνητής νοημοσύνης δεν είναι μια απλή «προκατάληψη». Έχει προταθεί ότι έχουμε να κάνουμε με μια νέα αποικιοκρατική αντίληψη του κόσμου, όπου τα μεγάλα τεχνολογικά μονοπώλια έχουν καταλάβει την επικοινωνία, τις υποδομές, και την πολιτική εξουσία, η οποία χαρακτηρίζεται ως αλγοριθμική αποικιοκρατία (Birhane 2020), τεχνο-αποικιοκρατία (Madianou 2019) ή αποικιοκρατία των δεδομένων (Thatcher, O'Sullivan και Mahmoudi 2016). Οι όροι αυτοί έχουν χρησιμοποιηθεί για να περιγράψουν την επέλαση των Μεγάλων Εταιρειών Τεχνολογίας (Big Tech) σε χώρες του Παγκόσμιου Νότου και σε μη προνομιούχες κοινότητες, με τεχνολογίες που εντείνουν τις ανισότητες, αντλούν πόρους από τις χώρες αυτές και ενδυναμώνουν τα τρέχοντα οικονομικά και πολιτικά συστήματα κυριαρχίας. Οι Μεγάλες Εταιρείες Τεχνολογίας επεκτείνονται γεωγραφικά και παίρνουν τις «πρώτες ύλες» της ψηφιακής εποχής, δηλαδή τα δεδομένα που προκύπτουν από την ανθρώπινη δραστηριότητα: στοιχεία που προκύπτουν από τη χρήση του κινητού τηλεφώνου, GPS, συσκευές που φοριούνται στο χέρι ή στο σώμα (wearables, όπως ρολόγια), τα Μέσα Κοινωνικής Δικτύωσης, τις αγορές με κάρτα ή σε διαδικτυακά καταστήματα. Όλα αυτά καταγράφονται σε πραγματικό χρόνο, αποθηκεύονται και υφίστανται επεξεργασία ώστε να προκύψουν πληροφορίες με πολιτική ή οικονομική αξία. Όπως και οι αποικιοκράτες, οι Μεγάλες Εταιρείες Τεχνολογίας θεωρούν ότι τα δεδομένα αυτά δεν ανήκουν σε κανέναν, άρα μπορούν να τα συλλέξουν και να τα χρησιμοποιήσουν για να ενισχύσουν την οικονομική τους δύναμη και την πολιτική τους κυριαρχία (Birhane 2020). Μέσα σε αυτό το πλαίσιο, υπάρχει μια ασύμμετρη κατανομή δύναμης ανάμεσα στους χρήστες των δεδομένων και τις εταιρείες που συλλέγουν και κατέχουν τα δεδομένα (Madianou 2019).

Ένα παράδειγμα που τεκμηριώνει τις θεωρίες για την τεχνο-αποικιοκρατία είναι το πρότζεκτ Facebook Free Basics, που παρουσιάστηκε σαν απόπειρα γεφύρωσης του ψηφιακού χάσματος μεταξύ του «Πρώτου Κόσμου» και των «Αναπτυσσόμενων Χωρών» και ως φιλανθρωπική βοήθεια σε άτομα, που δεν είχαν πρόσβαση στο Διαδίκτυο σε περισσότερες από εξήντα χώρες στη Λατινική Αμερική, την Ασία και την Αφρική. Ωστόσο, στην πραγματικότητα δεν έδινε δωρεάν πρόσβαση στο Ίντερνετ, αλλά μόνο στο Facebook, καθιστώντας το τη μοναδική ψηφιακή πηγή ενημέρωσης και επικοινωνίας σε αυτές τις χώρες. Έτσι, η επένδυση του Facebook απέφερε μονοπωλιακή πρόσβαση σε τεράστιες ποσότητες προσωπικών δεδομένων, που πωλήθηκαν σε τρίτες εταιρείες και απέφεραν κέρδη για την εταιρεία (ό.π.).

Παράλληλα, η επέκταση των Big Tech στις χώρες του Παγκόσμιου Νότου είχε ως αποτέλεσμα την τεχνολογική εξάρτηση από τους «τεχνο-αποικιοκράτες», που είχαν πλέον στον έλεγχό τους τις πληροφορίες, άρα και τη γνώση που αφορούσε τους ντόπιους πληθυσμούς. Ο πολιτισμός και η βιωμένη εμπειρία των ατόμων αυτών για άλλη

μια φορά φιλτράρεται μέσα από το δυτικό βλέμμα: από τη ζωγραφική απεικόνιση και την επιστημονική μελέτη των προηγούμενων αιώνων, η οποία και τότε παρουσιαζόταν ως «αντικειμενική», αλλά αντανάκλασε μια Ευρωκεντρική θεώρηση, περνάμε τώρα πια στην ανάλυση μέσα από τα δεδομένα, πάντα με δυτικά κριτήρια. Δηλαδή, οι δυτικοί (τέχνο)αποικιοκράτες έχουν και πάλι στα χέρια τους το αφήγημα για τους Άλλους. Επιπλέον, τα ψηφιακά προϊόντα που εισάγονται στις νέες χώρες, σχεδιάστηκαν σε άλλα συμφραζόμενα, χωρίς να λαμβάνουν υπόψη τις ανάγκες των τοπικών κοινωνιών και φυσικά χωρίς να τις αφήνουν να ωφεληθούν από την ψηφιοποίηση. Οι φεμινιστικές προσεγγίσεις στον τομέα της ψηφιακής τεχνολογίας αφορούν την ανεξαρτητοποίηση από τη δυτική κυριαρχία και ματιά, όπως θα αναλυθεί πιο κάτω.

Όπως τονίζει και η Hampton (2021, σ. 2), η αναπαραγωγή της αδικίας και της ανισότητας δεν είναι μια απλή «εγγενής προκατάληψη», γιατί «η προκατάληψη αφαιρεί την ευθύνη και κάνει να φαίνεται ότι το αποτέλεσμα είναι ακούσιο ή καλοπροαίρετο, και όχι ένα σκόπιμο υποπροϊόν καταπιεστικών θεσμών». Δηλαδή ο όρος «προκατάληψη» υποβαθμίζει τον ενεργό ρόλο των Μεγάλων Τεχνολογικών Εταιρειών στην καταπίεση μειονοτικών ομάδων, όπως φάνηκε στα παραδείγματα που έχουν δοθεί στο κείμενο: την αναπαραγωγή σεξιστικών και ρατσιστικών στερεοτύπων στα μοντέλα Τεχνητής Νοημοσύνης, την υπαρπαγή των δεδομένων για την εκπαίδευση των μοντέλων Τεχνητής Νοημοσύνης, την εφαρμογή τους σε συστήματα επιτήρησης και καταστολής, την αδιαφάνεια, την εργασιακή επισφάλεια, τις ιμπεριαλιστικού τύπου επεκτάσεις της τεχνολογικής τους κυριαρχίας σε παγκόσμιο επίπεδο, ακόμα και στην άσκηση πολιτικής εξουσίας σε χώρες του Παγκόσμιου Νότου για την εξασφάλιση των πρώτων υλών που χρειάζονται για τις τεχνολογίες αυτές.

Μέσα από αυτό το σύστημα καταπίεσεων προκύπτει αυτό που η Hampton (ό.π.) ονομάζει «αλγοριθμική καταπίεση». Σε αυτήν δεν μπορούν να δοθούν γρήγορες λύσεις, όπως είναι ο εμπλουτισμός των συνόλων δεδομένων ή η πρόσληψη περισσότερων εργαζομένων από μειονοτικές κοινωνικές ομάδες (ό.π.). Η επιδίωξη πολυπολιτισμικής εκπροσώπησης στα σύνολα δεδομένων δεν αλλάζει το γεγονός ότι τα δεδομένα συχνά συλλέγονται αυτοματοποιημένα, χωρίς τη ρητή συγκατάθεση των ατόμων, τα οποία αφορούν. Και βέβαια αποδεικνύεται επιφανειακή, όταν οι εργαζόμενοι που ανήκουν σε μειονοτικές ομάδες εργάζονται σε επισφαλείς συνθήκες. Ακόμα όμως και αν τους παρέχονταν οι ίδιες συνθήκες εργασίας, θα ήταν εξαιρετικά απίθανο να γίνουν αλλαγές σε ένα μεγάλο τεχνολογικό πρότζεκτ, επειδή κάποιοι εργαζόμενοι επισήμαναν τις εγγενείς προκαταλήψεις σε αυτό. Επιπλέον, θα ήταν άδικο να ανατεθεί στα άτομα αυτά το βάρος της διερεύνησης και ανάδειξης περιπτώσεων αλγοριθμικής προκατάληψης, δηλαδή της αποτίναξης της καταπίεσής τους.

Εν ολίγοις, οι Μεγάλες Εταιρείες Τεχνολογίας δεν θα επιχειρήσουν ποτέ να αναστρέψουν μια αδικία που δημιουργεί επιχειρηματικές ευκαιρίες και προωθεί την δική τους ανάπτυξη. Όπως θα έλεγε η Lorde (1984): «τα εργαλεία του αφέντη δεν θα γκρεμίσουν ποτέ το σπίτι του αφέντη». Δεν μπορεί να παταχθεί η αλγοριθμική καταπίεση χωρίς την κατάργηση της συστημικής καταπίεσης, καθώς είναι αδιαίρετες (Hampton 2021). Αντίθετα, είναι ανάγκη να ανατραπούν τα περίπλοκα ιστορικά, πολιτικά και οικονομικά ζητήματα που δημιουργούν και διαιωνίζουν την αλγοριθμική καταπίεση.

Αποαποικιοποίηση της Τεχνητής Νοημοσύνης

Το γεγονός ότι η καταπολέμηση της αλγοριθμικής καταπίεσης είναι τόσο σύνθετη, δεν σημαίνει ότι πρέπει να εναποτεθεί η προοπτική μιας δίκαιης Τεχνητής Νοημοσύνης σε ένα αβέβαιο μέλλον, όταν θα έχει γίνει η κοινωνία δικαιότερη. Η Τεχνητή Νοημοσύνη μπορεί να είναι ένα από τα πεδία μάχης, καθώς και ένα από τα επικοινωνιακά και πρακτικά εργαλεία για την ισότητα. Γι' αυτό, είναι σημαντικό το όραμα και η εφαρμογή μιας φεμινιστικής επιστημολογίας στην ανάπτυξη της Τεχνητής Νο-

ημοσύνης.

Καθώς η επιστήμη και η τεχνολογία δημιουργούνται από μια συγκεκριμένη τοποθέτηση -κοινωνική, ιδεολογική και επιστημολογική- μια φεμινιστική προσέγγιση διερευνά πρώτα ποιο άτομο δημιουργεί τι, προς όφελος ποιου και σε βάρος ποιου (Hampton 2021). Όλα τα συστήματα γνώσης είναι κοινωνικά τοποθετημένα, γι' αυτό μια φεμινιστική θεώρηση βοηθάει να διαλυθεί ο μύθος της αντικειμενικότητας αλλά και να ανοιχτεί το πεδίο της γνώσης σε επιστημολογίες, που προέρχονται από περιφερειακά συστήματα γνώσης (Bardzell 2010).

Η Haraway (1988) έχει διερευνήσει εκτενώς τα συστήματα γνώσης που χτίζονται από μειονοτικές κοινωνικές ομάδες, ανεξάρτητα από τις ηγεμονικές επιστημολογίες που δημιουργούνται από τα πάνω, εκπροσωπούν την εξουσία και συμβάλλουν στη συντήρησή της. Σύμφωνα με την Haraway (ό.π.):

«Η φεμινιστική αντικειμενικότητα έχει να κάνει με την περιορισμένη τοποθεσία και την τοποθετημένη γνώση, όχι με την υπέρβαση και τη διάσπαση υποκειμένου και αντικειμένου. Μας επιτρέπει να είμαστε υπόλογοι για αυτά που μαθαίνουμε πώς να βλέπουμε».

Με άλλα λόγια, το υποκείμενο που παρατηρεί και δημιουργεί τη γνώση δεν μπορεί να αποστασιοποιηθεί εντελώς από το αντικείμενο της παρατήρησης και πρέπει να έχει επίγνωση της θεσιακότητάς του, ώστε να είναι ανοιχτό προς προσεγγίσεις διαφορετικές από τη δική του, αλλά οφείλει να επιδεικνύει διαφάνεια ως προς αυτήν.

Η Hampton (2021) διευρύνει περαιτέρω το ζήτημα της τοποθετημένης γνώσης, μέσα από τη διερεύνηση της βιωμένης εμπειρίας ατόμων που υφίστανται πολλαπλές μορφές καταπίεσης με βάση το φύλο, τη φυλή, την ταυτότητα φύλου, το σεξουαλικό προσανατολισμό και την τάξη και εξετάζεται πώς η αλγοριθμική καταπίεση προστίθεται σε αυτές, μέσα από ένα πρίσμα μαύρου φεμινισμού. Εναντιώνεται ανοιχτά στην «καπιταλιστική ετεροπατριαρχική ηγεμονία της λευκής υπεροχής» (ό.π., σ. 5) και προτείνει μια φεμινιστική, μεταποικιακή και αντικαπιταλιστική στάση ενάντια στην αλγοριθμική καταπίεση. Ας δούμε πώς στοιχειοθετούνται αυτά τα τρία χαρακτηριστικά:

Η μεταποικιακή και φεμινιστική προσέγγιση σημαίνει τη συμπερίληψη εναλλακτικών επιστημολογιών που προέρχονται από τα άτομα, τα οποία αφορά ένα τεχνολογικό πρότζεκτ. Οι ομάδες που επωφελούνται από μια τεχνολογία θα έπρεπε να συμπεριλαμβάνονται στη διαδικασία έρευνας, σχεδιασμού και ανάπτυξης, καθώς γνωρίζουν καλύτερα τις ανάγκες τους. Δηλαδή οι σχεδιαστές και σχεδιάστριες υπολογιστών και οι επιστήμονες και επιστημότισσες υπολογιστών πρέπει να ξεκινούν τον σχεδιασμό και την ανάπτυξη της νέας τεχνολογίας σε συνεργασία με τα άτομα στα οποία απευθύνεται, λαμβάνοντας υπόψη τις δικές τους γνώσεις, εμπειρίες, και μορφές οργάνωσης, σε συνάρτηση με τις ιδιαίτερες ανάγκες τους. Για παράδειγμα, ένα θεσιακό σύνολο δεδομένων θα συνδύαζε τα Μεγάλα Δεδομένα (Big Data), που προκύπτουν μαζικά και αυτοματοποιημένα, με τα «μικρά» δεδομένα, που προκύπτουν από πιο εστιασμένες έρευνες και ποιοτικές μελέτες.

Η μεταποικιακότητα και η αντικαπιταλιστική στάση προκύπτουν από την τεχνολογική εκπαίδευση και αυτονόμηση. Όταν ένα άτομο ή μια ομάδα συμμετέχει στις διαδικασίες ανάπτυξης εφαρμογών Τεχνητής Νοημοσύνης, μαθαίνουν πώς να δημιουργούν τις δικές τους τεχνολογίες, προσαρμοσμένες στις ιδιαίτερες ανάγκες τους, σε αντίθεση με τον καπιταλισμό που ενθαρρύνει την παθητική κατανάλωση από άτομα που δεν γνωρίζουν πώς να καλύπτουν τις ανάγκες τους αυτόνομα (για παράδειγμα, καλλιέργεια ή προπαρασκευή τροφής, και δημιουργία εργαλείων). Δεύτερον, μια τέτοια προσέγγιση έχει ως αποτέλεσμα την παραγωγή σε μικρότερη κλίμακα, όπου χρησιμοποιούνται μόνο οι απολύτως αναγκαίοι πόροι, άρα μιλάμε για ένα πιο βιώσιμο μοντέλο ανάπτυξης. Παράλληλα, δίνοντας τη δυνατότητα στους ανθρώπους να συμμε-

τέχουν στην ανάπτυξη των τεχνολογιών που χρειάζονται, γίνονται αυτάρκεις και μπορούν να αντισταθούν στην εκμετάλλευση της εργασίας τους, των δεδομένων τους και των πόρων τους από τις αποικιοκρατικές δυνάμεις της ψηφιακής εποχής, δηλαδή τις Big Tech και τις οικονομικά προηγμένες χώρες που αξιοποιούν την τεχνολογική τους ηγεμονία, ώστε να επεκτείνουν την πολιτική τους κυριαρχία. Μέσα από την εκπαίδευση και την ανάπτυξη σε επίπεδο κοινότητας, οι τεχνολογίες της Τεχνητής Νοημοσύνης απομυθοποιούνται και αποαποικιοποιούνται από τη βάση τους. Όπως αναφέρεται στο *Decolonial AI Manifesto*² (2023):

«Αποαποικιοποίηση δεν είναι απλώς η διαφορετικότητα και η συμπερίληψη. Η εξάλειψη των απόηχων της αποικιοκρατίας στην Τεχνητή Νοημοσύνη χρειάζεται την επανόρθωση της υλικής και επιστημικής αδικίας και απαλλοτρίωσης, στο παρόν και το παρελθόν».

Τονίζει έτσι, τη σημασία της τεχνογνωσίας που προέρχεται από τη βιωμένη εμπειρία και την αξιοποίηση της τεχνολογίας, ώστε όσοι «έχουν ιστορικά περιθωριοποιηθεί να έχουν την ευκαιρία να αποφασίσουν και να χτίσουν το δικό τους αξιοπρεπές κοινωνικο-τεχνολογικό μέλλον».

Οι φεμινιστικές προσεγγίσεις στην επιστήμη των υπολογιστών και τον σχεδιασμό στρέφονται γύρω από τη δημιουργία τεχνολογιών με γνώμονα τον τελικό χρήστη/ χρήστρια, εμπλουτίζοντας τις ζωές κοινωνικών ομάδων με διαφορετικές εμπειρίες και ανάγκες. Για παράδειγμα, γίνεται λόγος για φεμινιστικό σχεδιασμό στην αλληλεπίδραση ανθρώπου υπολογιστή (feminist design in HCI) και φεμινιστική σχεδιαστική σκέψη (feminist design thinking), που όχι μόνο αντιστρέφει την καθιερωμένη υποεκπροσώπηση των γυναικών και των μειονοτικών ομάδων στο σχεδιασμό, αλλά επιδιώκει επίσης να ενσωματώσει την ιδέα της ισότητας στο σχεδιασμό του αλγόριθμου (Shipman 2020). Η ιδέα είναι να εφαρμοστούν φεμινιστικές μεθοδολογίες που θα αναθεωρήσουν κάθε πτυχή της διαδικασίας σχεδιασμού, από την αρχική ιδέα έως το τελικό προϊόν. Το πώς μπορεί να γίνει αυτό, θα το εξετάσουμε μέσα από παραδείγματα στην επόμενη ενότητα.

Το όραμα για μια Φεμινιστική Τεχνητή Νοημοσύνη

Προκειμένου να χτίσουμε το όραμα μιας Φεμινιστικής Τεχνητής Νοημοσύνης, μπορούμε να ξεκινήσουμε με προηγούμενες φεμινιστικές προσεγγίσεις στην ψηφιακή τεχνολογία, που αφορούν τη δημιουργία συστημάτων γνώσης και πρακτικών εργαλείων (Sollfrank 2016). Για παράδειγμα, η διαχείριση φεμινιστικών διακομιστών, αναδεικνύει τη σημασία της αυτονόμησης και της αποκέντρωσης της διαχείρισης της πληροφορίας, σε ανεξάρτητες υποδομές που διασφαλίζουν την ελεύθερη διακίνηση των πληροφοριών και την τήρηση φεμινιστικών αρχών, όπως η συναίνεση στη συλλογή δεδομένων και η ενδυνάμωση μέσα από τη γνώση. Το φεμινιστικό hacking βοηθάει άτομα από διαφορετικές κοινωνικές ομάδες -και μειονοτικές- να σχεδιάσουν και να δημιουργήσουν μόνες τους τα τεχνολογικά εργαλεία, που χρειάζονται στην καθημερινότητά τους στη βάση του ανοικτού λογισμικού. Έτσι, μπορούν ακόμα και να φτιάχνουν εργαλεία για να ελέγχουν την αναπαραγωγική τους υγεία, στο πλαίσιο του ακτιβισμού Gynepunk. Οι φεμινιστικές τεχνο-ακτιβιστικές ομάδες δημιουργούν διαδικτυακούς ασφαλείς χώρους και εγχειρίδια κυβερνοασφάλειας, ώστε να βοηθήσουν στην ενδυνάμωση και την προστασία από τις διαδικτυακές επιθέσεις, την επιτήρηση και την καταστολή. Εξετάζοντας φεμινιστικές προσεγγίσεις στην ψηφιακή τεχνολογία, διαπιστώνουμε ότι καλύπτουν ένα ευρύ φάσμα, από τις υλικές υποδομές, μέχρι την επικοινωνία και την ανάπτυξη εναλλακτικών τεχνολογικών εργαλείων.

Η προσέγγιση αυτή, που πραγματοποιείται από κάτω προς τα πάνω και οριζόντια,

² Μανιφέστο της Αποαποικιοποιημένης Τεχνητής Νοημοσύνης

διαφέρει ριζικά σε σχέση με το πώς αντιμετωπίζουν τα σύγχρονα συστήματα Τεχνητής Νοημοσύνης το ζήτημα της ισότητας. Αυτό προκύπτει αν συγκρίνουμε το ChatGPT, ένα από τα πιο κυρίαρχα συστήματα Τεχνητής Νοημοσύνης, με πειραματισμούς στη Φεμινιστική Τεχνητή Νοημοσύνη.

Σε αντίθεση με τα ανθρωποειδή ρομπότ της επιστημονικής φαντασίας, η Τεχνητή Νοημοσύνη σήμερα δεν είναι ούτε αυστηρά «τεχνητή» ούτε «νοήμων». Για παράδειγμα, τα chatbots μπορεί να έχουν εξελιχθεί για να κατανοούν την ανθρώπινη γλώσσα και να ανταποκρίνονται ανάλογα, αλλά αν ρίξουμε μια πιο προσεκτική ματιά στην αρχιτεκτονική και στη λειτουργία τους, δεν είναι τόσο καινούρια. Οι πρώτες τεχνολογίες αυτοματοποιημένου διαλόγου, όπως το *ELIZA*, που κυκλοφόρησε το 1964 ή τα βιντεοπαιχνίδια της δεκαετίας του 1970, χρησιμοποιούσαν «δέντρα» συνομιλίας, όπου κάθε απάντηση οδηγούσε στον επόμενο «κλάδο» κειμένου. Η καινοτομία που εισάγεται με τεχνολογίες όπως το *ChatGPT* είναι ότι πλέον υπάρχει πρόσβαση σε έναν τεράστιο όγκο δεδομένων, που σε συνδυασμό με την Επεξεργασία Φυσικής Γλώσσας (NLP) μπορεί να δώσει μια πολύ φυσική συνομιλία -χωρίς βέβαια αυτό να σημαίνει ότι οι πληροφορίες που δίνει είναι πάντοτε ορθές ή ουδέτερες. Η μεροληψία εδώ δεν είναι κάτι που προκύπτει μονάχα από τα δεδομένα πάνω στα οποία εκπαιδεύεται το σύστημα. Αν ρωτήσουμε το ίδιο το *ChatGPT* για τις δικλίδες ασφαλείας του απέναντι σε προκαταλήψεις και λόγο μίσους, εξηγεί ότι υπάρχουν φίλτρα που αποκλείουν ορισμένα σετ δεδομένων και επεξεργάζονται τις απαντήσεις ανάλογα, ενώ υπάρχουν και διαχειριστές περιεχομένου. Δηλαδή πρόκειται για ένα σύστημα που μπορεί μεν να μάθει από τις πληροφορίες που του δίνονται, αλλά με συγκεκριμένους περιορισμούς που έχουν τεθεί *a priori* από την ομάδα που το δημιούργησε και που ενίοτε τίθενται και *a posteriori* σε περιπτώσεις καταγγελίας περιεχομένου από τους χρήστες προς τους διαχειριστές.

Έτσι, δεν θα συναντήσουμε στο *ChatGPT* ακραίο περιεχόμενο, αλλά ούτε και ιδιαίτερα προοδευτικές θέσεις. Όταν θέτουμε ένα ερώτημα για τα δικαιώματα των φύλων, απαντάει φυσικά ότι οι γυναίκες πρέπει να έχουν ίσα δικαιώματα, αλλά όταν το ερώτημα αφορά κάτι πιο εξειδικευμένο που άπτεται των ίσων δικαιωμάτων, όπως για παράδειγμα η πρόσβαση στην νόμιμη και ασφαλή έκτρωση, προσπαθεί να τηρήσει μια πιο «ουδέτερη» στάση και δεν παίρνει θέση, αλλά παραθέτει τα επιχειρήματα των δύο ομάδων που στηρίζουν τις αντίθετες θέσεις. Αντίστοιχα, σε ό,τι αφορά την αναπαραγωγή των αποικιοκρατικών συστημάτων κυριαρχίας, αν το ρωτήσουμε, για παράδειγμα, για το δικαίωμα των Ισραηλινών στην ελευθερία, θα απαντήσει καταφατικά, ενώ αν ρωτήσουμε για αντίστοιχο δικαίωμα των Παλαιστινίων στην ελευθερία, θα επιλέξει έναν πιο «ουδέτερο» τόνο, αποφεύγοντας να δώσει μια θετική απάντηση.³ Φυσικά, η ουδετερότητα εδώ είναι μόνο κατ' επίφαση. Η προώθηση των «δυτικών απόψεων», όπως τις ονομάζει, είναι κάτι που παραδέχεται το ίδιο το πρότζεκτ του OpenAI που δημιούργησε και εξελίσσει το εργαλείο, δηλώνοντας ανοιχτά τη θεσιακότητά του (OpenAI 2023).

Τα διαλογικά μοντέλα της Φεμινιστικής Τεχνητής Νοημοσύνης, από την άλλη, όχι μόνο μιλάνε ανοιχτά για την ισότητα σε διαθεματικό επίπεδο, αλλά και επιδιώκουν να χτίσουν τα απαραίτητα εργαλεία που θα βοηθήσουν στην αποτίναξη των έμφυλων και πολλαπλών καταπιέσεων. Ένα παράδειγμα είναι η *Betânia* (2017), μια εφαρμογή chatbot που κυκλοφόρησε στο Facebook Messenger, προκειμένου να βοηθήσει φεμινίστριες στη Βραζιλία να οργανωθούν ενάντια σε μια σχεδιαζόμενη τροποποίηση του συντάγματος, που θα ποινικοποιούσε τις αμβλώσεις. Εκτός από τον ενημερωτικό της χαρακτήρα, η *Betânia* βοήθησε τους χρήστες και τις χρήστριες να στείλουν πάνω από 34.000 μηνύματα ηλεκτρονικού ταχυδρομείου σε βουλευτές και βουλευτρίες, συμβάλλοντας ουσιαστικά στην καταψήφιση του νομοσχεδίου. Ένα ακόμα φεμινιστικό bot με ακτιβιστικό περιεχόμενο είναι το *acoso.online*, που εξειδικεύεται σε ζητήματα

³ Βλ. <https://chat.openai.com/>

εκδικητικής πορνογραφίας. Μέσα από την εφαρμογή, τα θύματα εκδικητικής πορνογραφίας μπορούν να ενημερωθούν με απλό τρόπο για τη σχετική νομοθεσία της χώρας τους και τα βήματα που μπορούν να κάνουν για προστασία και καταγγελία (Toupin και Couture 2020). Η φεμινιστική εφαρμογή *Mumkin* απευθύνεται στις γυναίκες στην Ινδία, κυριότερα αυτές που ανήκουν στη θρησκευτική κοινότητα των Bohra και τις βοηθάει να συζητήσουν τα ζητήματα της έμφυλης βίας, του ακρωτηριασμού των γυναικείων γεννητικών οργάνων και της συναίνεσης, παρέχοντας ένα chatbot που παίζει το ρόλο του συζύγου, της μητέρας ή της πεθεράς (*mumkinapp*). Τέλος, η *F'xa*, ένα bot που δημιουργήθηκε από το Feminist Internet, παρουσιάζει τις φεμινιστικές ιδέες και αρχές σε μορφή διαλόγου, περιλαμβάνοντας ακόμα και memes και emojis (ό.π.).

Ωστόσο, αν ρίξουμε μια πιο προσεκτική ματιά σε αυτά τα chatbot, δύσκολα μπορούμε να μιλήσουμε για «έξυπνα» συστήματα, καθώς βασίζονται σε προκαθορισμένα σενάρια, χωρίς μηχανική μάθηση, δηλαδή χωρίς τη δυνατότητα να εκπαιδεύεται το σύστημα από τα άτομα που το χρησιμοποιούν και να προσαρμόζεται σε διαφορετικά ερωτήματα. Οι περιορισμοί στην τεχνολογική καινοτομία οφείλονται στα υλικά και τεχνικά μέσα που έχουν στη διάθεσή τους οι φεμινιστικές ακτιβιστικές ομάδες, που αναπτύσσουν τις τεχνολογίες αυτές. Θα ήταν σχεδόν αδύνατο να δημιουργηθεί ένα εργαλείο σαν το *ChatGPT* της OpenAI, που στηρίζεται από τη Microsoft και άλλους επενδυτές με δισεκατομμύρια δολάρια. Από την άλλη, είναι και μια σκόπιμη επιλογή (ό.π.), γιατί υπάρχει άμεσος κίνδυνος να παρέμβουν στα συστήματα αυτά αντιφεμινιστικές ομάδες και να τα εκπαιδεύσουν ώστε να αναπαράγουν τις αντίθετες ιδέες από αυτές για τις οποίες σχεδιάστηκαν (όπως είχε συμβεί με την *Tay* της Microsoft που αναφέρθηκε προωτέρω). Έτσι, χρησιμοποιούνται προκαθορισμένες επιλογές, οι οποίες προέκυψαν μέσα από μια διαδικασία φεμινιστικού σχεδιασμού με την συμπεριληψη των ομάδων στις οποίες απευθύνονται τα εργαλεία και τη χρησιμοποίηση γνώσεων, που προέρχονται από αυτές. Με αυτόν τον τρόπο, προκύπτουν εργαλεία επικοινωνίας που προσφέρουν έναν πιο άμεσο και σύγχρονο τρόπο παρουσίασης των πληροφοριών, οι οποίες χάρη στο διαλογικό μοντέλο γίνονται πιο άμεσες και εύληπτες από τα άτομα που τις χρειάζονται. Οι φεμινιστικές αυτές ομάδες δεν εισάγουν νέες τεχνολογίες, αλλά διερευνούν μεθόδους και εισάγουν πρακτικά εργαλεία περιορισμένης εμβέλειας, που μπορούν ενίοτε να σημειώσουν νίκες ενάντια στις τοπικές πατριαρχίες.

Η Τεχνητή Νοημοσύνη παρουσιάζεται συχνά ως ένα «μαύρο κουτί», που τροφοδοτείται με δεδομένα και παρέχει προβλέψεις. Κατά συνέπεια, καθίσταται δύσκολο να εντοπίσουμε την πηγή τυχόν προβλημάτων που μπορεί να προκύψουν, όταν δεν κατανοούμε πώς λειτουργεί το σύστημα. Αυτός είναι ο λόγος για τον οποίο ορισμένα καλλιτεχνικά έργα που ασχολούνται με την Φεμινιστική Τεχνητή Νοημοσύνη επικεντρώνονται στη διαδικασία, προσπαθώντας να φέρουν στο φως όλα τα βήματα που ακολουθούνται: από τη δημιουργία συνόλων δεδομένων έως τις προβλέψεις, ενώ παράλληλα εκπαιδεύουν τους χρήστες/χρήστριες στο να αναπαράγουν αυτές τις τεχνολογίες.

Το *Feminist Data Set* (από το 2017 μέχρι και σήμερα) της Caroline Sindors (2020) είναι ένα καλλιτεχνικό έργο εν εξελίξει που εξετάζει κάθε βήμα της διαδικασίας Τεχνητής Νοημοσύνης: από τη συλλογή δεδομένων και την κατηγοριοποίηση με βάση ετικέτες, έως τη δημιουργία ενός μοντέλου μηχανικής μάθησης, την αλγοριθμική εκπαίδευση, την υλοποίηση και τη δημιουργία ενός προτύπου. Το ερευνητικό πλαίσιο του έργου ακολουθεί σε κάθε βήμα τις αρχές του διαθεματικού φεμινισμού, επιδιώκοντας να εξαλείψει κάθε στοιχείο προκατάληψης και εκμετάλλευσης. Μέσα στο πρότζεκτ εκπροσωπούνται ισότιμα άτομα που έχουν διαφορετικά φύλα, ταυτότητες φύλου και καταγωγή, όχι μόνο επειδή αυτό μπορεί να βοηθήσει άτομα από διαφορετικές κοινότητες να αποκτήσουν γνώση στην ανάπτυξη της Τεχνητής Νοημοσύνης αλλά και επειδή με αυτόν τον τρόπο η τεχνολογία ανοίγεται σε επιστημολογίες, που δεν έχουν

αξιοποιηθεί επαρκώς και συνδέονται με τις γνώσεις αυτών των κοινοτήτων. Επιπλέον, τα δεδομένα που χρησιμοποιούνται, συγκεντρώνονται με συναινετικό τρόπο από την κοινότητα για την κοινότητα. Ένα «αποαποικιοποιημένο» σύνολο δεδομένων εμπλέκει και μειονοτικές ομάδες στη διαδικασία συλλογής δεδομένων, αναγνωρίζοντας ότι μπορούν να καταλάβουν τι είναι καλύτερο για τις ίδιες -σε αντίθεση με την προσέγγιση εταιρειών που σχεδιάζουν τεχνολογίες για αυτές χωρίς αυτές. Είναι σημαντικό ότι τα συμμετέχοντα άτομα είναι και κάτοχοι των δεδομένων τους και αμείβονται με ένα δίκαιο ποσό για τη συνεισφορά τους στο πρότζεκτ. Υπάρχει διαφάνεια ως προς τον τρόπο που χτίζεται η βάση δεδομένων και την αμοιβή των συμμετεχόντων ατόμων, πράγμα που καταδεικνύει πόσο δύσκολη είναι η εργασία αυτή.

Περιττό να πούμε ότι η ανάπτυξη ενός τέτοιου πρότζεκτ τόσο μικρής εμβέλειας απαιτεί σημαντικό χρόνο και χρήμα, σε αντίθεση με τα συνηθισμένα σύνολα δεδομένων, τα οποία, όπως αναφέραμε παραπάνω, χαρακτηρίζονται από έλλειψη συναίνεσης και επισφαλείς συνθήκες εργασίας. Αυτός είναι ο λόγος για τον οποίο η διάδοση της Φεμινιστικής Τεχνητής Νοημοσύνης θα ήταν μια αργή διαδικασία, που συνδυάζει τα Μεγάλα Δεδομένα με «μικρά» δεδομένα, προσεγγίζει κριτικά την αυτοματοποίηση και αντικαθιστά την εκμετάλλευση με τη συγκατάθεση.

Συμπέρασμα

Όπως προκύπτει μέσα από μια διαθεματική φεμινιστική κριτική στα συστήματα Τεχνητής Νοημοσύνης, τα συστήματα αυτά αναπαράγουν και ενισχύουν τις υπάρχουσες ανισότητες. Το πεδίο της Φεμινιστικής Τεχνητής Νοημοσύνης παρέχει τα θεωρητικά σχήματα, τις επιστημονικές γνώσεις, την πολιτική εμπρόθετη δράση και τις πολιτιστικές πρακτικές αντίστασης στις τεχνολογικές ηγεμονίες, που παράγουν την αλγοριθμική καταπίεση. Διαλύοντας τον μύθο της τεχνολογικής αντικειμενικότητας, εμπλουτίζοντας τη γνώση με την εμπειρία των μειονοτικών ομάδων και ανοίγοντας την τεχνολογία σε συλλογικά μοντέλα σχεδιασμού και ανάπτυξης, από την κοινότητα για την κοινότητα, η διαθεματική φεμινιστική προσέγγιση δείχνει πώς μπορεί να αναθεωρηθεί ο τρόπος με τον οποίο χτίζεται η γνώση και η τεχνολογία. Ο τρόπος αυτός ευθυγραμμίζεται με την μεταποικιακή και αντικαπιταλιστική κριτική, δείχνοντας το δρόμο προς πιο βιώσιμα μοντέλα τεχνολογικής ανάπτυξης, σε αντίθεση με τις τρέχουσες πρακτικές εκμετάλλευσης και εξτρακτιβισμού.

Η Φεμινιστική Τεχνητή Νοημοσύνη μπορεί να μην είναι εμπορικά ανταγωνιστική, να μην έχει αξιοποιηθεί παρά μόνο αποσπασματικά σε καλλιτεχνικά και ακτιβιστικά πρότζεκτ, αλλά έχει θέση σε ένα τεχνολογικό μέλλον, όπου οι άνθρωποι ανησυχούν όλο και περισσότερο για τη διασφάλιση των δικαιωμάτων τους μέσω της νομοθεσίας και αναζητούν ηθικές εναλλακτικές σε τεχνολογίες, που στηρίζονται στην εκμετάλλευση και προωθούν την καταπίεση. Ακόμη και με τους προαναφερθέντες περιορισμούς, οι πειραματισμοί με μια διαθεματική φεμινιστική προσέγγιση στην Τεχνητή Νοημοσύνη, που εμπλέκει τις κοινότητες στη διαδικασία ανάπτυξης και σέβεται τα ανθρώπινα δικαιώματα, εισάγει έναν διαφορετικό τρόπο σκέψης, όπου η τεχνητή νοημοσύνη δεν θεωρείται κάποιο είδος «μαύρου κουτιού» ή αμερόληπτος κριτής, αλλά ένα ανθρώπινο δημιούργημα που μπορεί να εξυπηρετήσει τις ανάγκες και να σεβαστεί τα δικαιώματα των ανθρώπων.

Το όραμα μιας Φεμινιστικής Τεχνητής Νοημοσύνης σημαίνει ανοιχτή αμφισβήτηση των τρόπων με τους οποίους η τεχνολογία αναπαράγει τη συστημική καταπίεση· σημαίνει προσπάθεια δημιουργίας εναλλακτικών εργαλείων με βάση φεμινιστικές αρχές, όπως η συγκατάθεση, η ενδυνάμωση, η οικοδόμηση γνώσης, η βιωσιμότητα, η χειραφέτηση και η ελεύθερη πρόσβαση στη γνώση και τις πληροφορίες. Συνολικά, μια Φεμινιστική Τεχνητή Νοημοσύνη εκφράζει αντίσταση στην αποικιοκρατική και καπιταλιστική επιβολή της τεχνολογίας, ενάντια στην αλόγιστη εκμετάλλευση πόρων και

ανθρώπων, ενώ ταυτόχρονα παρέχει σε μειονοτικές ομάδες την εμπρόθετη δράση ώστε να εκπαιδευτούν και να δημιουργήσουν τις δικές τους τεχνολογίες και τα εργαλεία της απελευθέρωσής τους.

Βιβλιογραφία

Baker, S. E. (2018) “Post-work Futures and Full Automation: Towards a Feminist Design Methodology”, *Open Cultural Studies*, 2, pp. 540-552.

Bardzell, S. (2010) “Feminist HCI: Taking Stock and Outlining an Agenda for Design”, paper presented at CHI 2010: HCI for all, Atlanta, GA, USA, April 10-15.

Birhane, A. (2020) “Algorithmic Colonization of Africa”, *Scripted*, 17(2) [online]. Available at: <https://script-ed.org/wp-content/uploads/2020/08/birhane.pdf?d=12142023> [Accessed 14 December 2023].

Broussard, M. (2018) *Artificial Unintelligence: How Computers Misunderstand the World*. Massachusetts: The MIT Press.

Buolamwini, J. (2023) *Unmasking AI: My Mission to Protect What is Human in a World of Machines*. New York: Penguin Random House.

D’Ignazio, C. and Klein L. F. (2020) *Data Feminism*. Massachusetts: The MIT Press.

Franzoni, V. (2023) “Gender Differences and Bias in Artificial Intelligence” in Vallverdú, J. (ed.) *Gender in AI and Robotics: The Gender Challenges from an Interdisciplinary Perspective*. Cham: Springer.

Grammatikopoulou, C. (2022) “Envisioning a Feminist Artificial Intelligence” in Grunert, F. P. (ed.) *HumanITies and Artificial Intelligence*. European Union: Noema.

Haraway, D. (1988) “Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective”, *Feminist Studies*, pp. 575-599.

Hampton, L. M. (2021) “Black Feminist Musings on Algorithmic Oppression”, paper presented at Conference on Fairness, Accountability, and Transparency (FAccT ’21), Virtual Event, Canada, March 3–10.

Lorde, A. (1984) *Sister Outsider: Essays and Speeches*. Berkley: Crossing Press.

Madianou, M. (2019) “Technocolonialism: Digital Innovation and Data Practices in the Humanitarian Response to Refugee Crises”, *Social Media + Society*, 5(3) [online], <https://doi.org/10.1177/2056305119863146> [Accessed: 10 November 2023].

Noble, S. (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.

Shipman, M. (2020) “Can ‘Feminist Design’ Save Hiring Algorithms from Bias?” [online]. Available at: <https://www.futurity.org/feminist-design-hiring-algorithms-bias-2276022/> [Accessed: 10 November 2023].

Sinders, C. (2020) *Feminist Data Set, Open Source Tool Kit* [online]. Available at: <https://carolinesinders.com/wp-content/uploads/2020/05/Feminist-Data-Set-Final-Draft-2020-0526.pdf> [Accessed: 10 November 2023].

Sollfrank, C. (2016) “Revisiting the Future: Cyberfeminism in the 21st Century” in Bishop, R., Gansing, K., Parikka, J. and Wilk, E. (eds.) *Across & Beyond – A transmediale Reader on Post-digital Practices, Concepts and Institutions*. Berlin:

Sternberg Press, pp. 228-247.

Thatcher, J., O’Sullivan, D. and Mahmoudi, D. (2016) “Data Colonialism through Accumulation by Dispossession: New metaphors for daily data”, *Environment and Planning D: Society and Space*, 34(6), 990–1006.

Toupin, S. and Couture, S. (2020) “Feminist Chatbots as Part of the Feminist Toolbox”, *Feminist Media Studies*, 20(5), pp. 737-740.

Toupin, S. (2023) “Shaping feminist artificial intelligence”. *New Media & Society*, 0(0) [online]. <https://doi.org/10.1177/14614448221150776> [Accessed: 10 November 2023].

UNESCO and EQUALS Skills Coalition (2019) “I’d Blush If I Could: Closing Gender Divides in Digital Skills Through Education” [online]. Available at: <https://doi.org/10.54675/RAPC9356> [Accessed: 10 November 2023].

Wolf, M.J., Miller, K.W. and Grodzinsky, F.S. (2017) “Why We Should Have Seen that Coming: Comments on Microsoft’s Tay ‘Experiment’ and Wider Implications”, *The ORBIT Journal*, 1(2), pp. 1-12 [online]. Available at: <https://doi.org/10.29297/orbit.v1i2.49> [Accessed: 10 November 2023].

Σύνδεσμοι:

ChatGPT, <https://chat.openai.com/> [Accessed: 10 November 2023].

Decolonial AI Manyfesto, <https://manyfesto.ai/index.html> [Accessed: 10 November 2023].

Mumkin, <https://www.mumkinapp.com/research> [Accessed: 10 November 2023].

OpenAI (2023), Is ChatGPT Biased? <https://help.openai.com/en/articles/8313359-is-chatgpt-biased> [Accessed: 10 November 2023].