

Μήνυμα του Αγίου Πατέρα Φραγκίσκου

για την 57η Παγκόσμια Ημέρα Ειρήνης

(1^η Ιανουαρίου 2024)

Τεχνητή νοημοσύνη και ειρήνη

Στην αρχή του νέου έτους, μιας περιόδου χάριτος που ο Κύριος δωρίζει στον καθένα μας, θα ήθελα να απευθυνθώ στον Λαό του Θεού, στα έθνη, στους αρχηγούς Κρατών και Κυβερνήσεων, στους Εκπροσώπους των διαφόρων θρησκειών και της κοινωνίας των πολιτών, καθώς και σε όλους τους άνδρες και τις γυναίκες της εποχής μας, για να εκφράσω τις καλύτερες ευχές μου για ειρήνη.

1. Η πρόοδος της επιστήμης και της τεχνολογίας ως δρόμος προς την ειρήνη

Η Αγία Γραφή μαρτυρεί ότι ο Θεός έδωσε το Πνεύμα Του στους ανθρώπους, για να έχουν “σοφία, κατανόηση και γνώση σε κάθε είδους εργασία” (Εξ 35,31). Η νοημοσύνη αποτελεί έκφραση της αξιοπρέπειας που μας έχει χαρίσει ο Δημιουργός, ο οποίος μας δημιούργησε κατ’ εικόνα και ομοίωσή του (πρβλ. Γεν. 1,26) και μας έδωσε τη δυνατότητα να ανταποκριθούμε στην αγάπη του μέσω της ελευθερίας και της γνώσεως. Η επιστήμη και η τεχνολογία εκδηλώνουν με ιδιαίτερο τρόπο αυτή τη θεμελιωδώς σχεσιακή ποιότητα της ανθρώπινης νοημοσύνης: είναι εξαιρετικά προϊόντα του δημιουργικού δυναμικού της.

Στην Ποιμαντική Διάταξη *Gaudium et Spes*, η Β’ Βατικανή Σύνοδος επιβεβαίωσε αυτή την αλήθεια, δηλώνοντας ότι “με το έργο και την εφευρετικότητά του ο άνθρωπος προσπάθησε πάντοτε να αναπτύξει την ίδια του τη ζωή”[1]. Όταν οι άνθρωποι, “με τη βοήθεια της τεχνολογίας”, προσπαθούν να καταστήσουν τη γη “κατοικία αντάξια ολόκληρης της ανθρώπινης οικογένειας”[2], ενεργούν σύμφωνα με το σχέδιο του Θεού και συνεργάζονται με τη θέλησή του για την ολοκλήρωση της δημιουργίας και τη διάδοση της ειρήνης μεταξύ των λαών. Ακόμη και η πρόοδος της επιστήμης και της τεχνολογίας, στο μέτρο που συμβάλλει στην καλύτερη τάξη της ανθρώπινης κοινωνίας, στην αύξηση της ελευθερίας και της αδελφικής κοινωνίας, οδηγεί έτσι στη βελτίωση του ανθρώπου και στη μεταμόρφωση του κόσμου. Δικαιολογημένα χαιρόμαστε και είμαστε ευγνώμονες για τα εξαιρετικά επιτεύγματα της επιστήμης και της τεχνολογίας, χάρη στα οποία έχουν θεραπευτεί αμέτρητα δεινά που ταλαιπωρούσαν την ανθρώπινη ζωή και προκαλούσαν μεγάλη ταλαιπωρία. Ταυτόχρονα, η τεχνολογική και επιστημονική πρόοδος, η οποία καθιστά δυνατή την άσκηση πρωτοφανούς μέχρι σήμερα ελέγχου της πραγματικότητας, θέτει στα χέρια του ανθρώπου ένα ευρύ φάσμα δυνατοτήτων, ορισμένες από τις οποίες μπορούν να αποτελέσουν κίνδυνο για την επιβίωση και να θέσουν σε κίνδυνο το κοινό σπίτι[3].

Η αξιοσημείωτη πρόοδος των νέων τεχνολογιών της πληροφορίας, ιδίως στον ψηφιακό τομέα, παρουσιάζει επομένως συναρπαστικές ευκαιρίες και σοβαρούς κινδύνους, με σοβαρές επιπτώσεις στην επιδίωξη της δικαιοσύνης και της αρμονίας

μεταξύ των ανθρώπων. Συνεπώς, πρέπει να τεθούν ορισμένα επείγοντα ερωτήματα. Ποιες θα είναι οι μεσοπρόθεσμες και μακροπρόθεσμες συνέπειες των νέων ψηφιακών τεχνολογιών; Και τι επιπτώσεις θα έχουν στη ζωή των ατόμων και της κοινωνίας, στη διεθνή σταθερότητα και ειρήνη;

2. Το μέλλον της τεχνητής νοημοσύνης ανάμεσα σε υποσχέσεις και κινδύνους

Η πρόοδος της πληροφορικής και η ανάπτυξη των ψηφιακών τεχνολογιών τις τελευταίες δεκαετίες έχουν ήδη αρχίσει να προκαλούν βαθιές μεταμορφώσεις στην παγκόσμια κοινωνία και τη δυναμική της. Τα νέα ψηφιακά εργαλεία αλλάζουν το πρόσωπο των επικοινωνιών, της δημόσιας διοίκησης, της εκπαίδευσης, της κατανάλωσης, των προσωπικών αλληλεπιδράσεων και αμέτρητων άλλων πτυχών της καθημερινής ζωής. Επιπλέον, οι τεχνολογίες που χρησιμοποιούν μια πληθώρα αλγορίθμων μπορούν να εξάγουν, από τα ψηφιακά ίχνη που αφήνουν στο Διαδίκτυο, δεδομένα που επιτρέπουν τον έλεγχο των νοητικών και σχεσιακών συνηθειών των ανθρώπων για εμπορικούς ή πολιτικούς σκοπούς, συχνά εν αγνοία τους, περιορίζοντας τη συνειδητή άσκηση της ελευθερίας των επιλογών τους. Πράγματι, σε έναν χώρο όπως το διαδίκτυο, που χαρακτηρίζεται από υπερφόρτωση πληροφοριών, μπορούν να δομούν τη ροή των δεδομένων σύμφωνα με κριτήρια επιλογής που δεν γίνονται πάντα αντιληπτά από τον χρήστη.

Πρέπει να θυμόμαστε ότι η επιστημονική έρευνα και οι τεχνολογικές καινοτομίες δεν είναι αποκομμένες από την πραγματικότητα και “ουδέτερες”[4], αλλά υπόκεινται σε πολιτισμικές επιρροές. Ως πλήρως ανθρώπινες δραστηριότητες, οι κατευθύνσεις που παίρνουν αντανακλούν επιλογές που εξαρτώνται από τις προσωπικές, κοινωνικές και πολιτιστικές αξίες κάθε εποχής. Το ίδιο ισχύει και για τα αποτελέσματα που επιτυγχάνουν: ακριβώς επειδή είναι αποτέλεσμα ειδικά ανθρώπινων προσεγγίσεων του κόσμου γύρω τους, έχουν πάντα μια ηθική διάσταση, στενά συνδεδεμένη με τις αποφάσεις εκείνων που σχεδιάζουν τον πειραματισμό και κατευθύνουν την παραγωγή προς συγκεκριμένους στόχους. Αυτό ισχύει και για τις μορφές τεχνητής νοημοσύνης. Μέχρι σήμερα, δεν υπάρχει ένας ενιαίος ορισμός της στον κόσμο της επιστήμης και της τεχνολογίας. Ο ίδιος ο όρος, ο οποίος έχει πλέον εισέλθει στην κοινή γλώσσα, περιλαμβάνει μια ποικιλία επιστημών, θεωριών και τεχνικών που αποσκοπούν στο να αναπαράγουν ή να μιμηθούν οι μηχανές τις γνωστικές ικανότητες του ανθρώπου κατά τη λειτουργία τους. Η χρήση του πληθυντικού τονίζει επίσης ότι αυτές οι πολύ διαφορετικές συσκευές πρέπει πάντα να θεωρούνται ως “κοινωνικοτεχνικά συστήματα”. Στην πραγματικότητα, ο αντίκτυπός τους, πέρα από τη βασική τεχνολογία, εξαρτάται όχι μόνο από τον σχεδιασμό τους, αλλά και από τους στόχους και τα συμφέροντα των ιδιοκτητών και των προγραμματιστών τους, καθώς και από τις καταστάσεις στις οποίες χρησιμοποιούνται. Επομένως, η τεχνητή νοημοσύνη πρέπει να γίνει αντιληπτή ως ένας γαλαξίας διαφορετικών πραγματικοτήτων και δεν μπορούμε να υποθέσουμε εκ των προτέρων ότι η ανάπτυξή της θα συμβάλει ευεργετικά στο μέλλον της ανθρωπότητας και στην ειρήνη μεταξύ των λαών. Ένα τέτοιο θετικό αποτέλεσμα θα είναι εφικτό μόνο εάν δείξουμε ότι είμαστε ικανοί να ενεργούμε υπεύθυνα και να σεβόμαστε τις θεμελιώδεις ανθρώπινες αξίες, όπως “η συμμετοχικότητα, η διαφάνεια, η ασφάλεια, η δικαιοσύνη, η

εμπιστευτικότητα και η λογοδοσία”[5]. Επίσης, δεν αρκεί να υποθέσουμε, εκ μέρους όσων σχεδιάζουν αλγόριθμους και ψηφιακές τεχνολογίες, ότι δεσμεύονται να ενεργούν ηθικά και υπεύθυνα. Είναι ανάγκη να ενισχυθούν ή, αν χρειαστεί, να δημιουργηθούν φορείς για την εξέταση των αναδυόμενων ηθικών ζητημάτων και την προστασία των δικαιωμάτων όσων χρησιμοποιούν την τεχνητή νοημοσύνη ή επηρεάζονται από αυτήν[6].

Η τεράστια επέκταση της τεχνολογίας πρέπει, επομένως, να συνοδεύεται από την κατάλληλη εκπαίδευση στην ευθύνη για την ανάπτυξή της. Η ελευθερία και η ειρηνική συνύπαρξη απειλούνται όταν τα ανθρώπινα όντα υποκύπτουν στον πειρασμό του εγωισμού, της ιδιοτέλειας, της δίψας για κέρδος και της δίψας για εξουσία. Έχουμε λοιπόν καθήκον να διευρύνουμε το βλέμμα μας και να κατευθύνουμε την τεχνική και επιστημονική έρευνα προς την επιδίωξη της ειρήνης και του κοινού καλού, στην υπηρεσία της ολοκληρωμένης ανάπτυξης του ανθρώπου και της κοινότητας[7]. Η εγγενής αξιοπρέπεια κάθε ανθρώπου και η αδελφοσύνη που μας συνδέει ως μέλη της μιας ανθρώπινης οικογένειας πρέπει να στηρίζουν την ανάπτυξη των νέων τεχνολογιών και να χρησιμεύουν ως αδιαμφισβήτητα κριτήρια για την αξιολόγησή τους πριν από τη χρήση τους, ώστε η ψηφιακή πρόοδος να πραγματοποιείται με σεβασμό προς τη δικαιοσύνη και να συμβάλλει στην υπόθεση της ειρήνης. Οι τεχνολογικές εξελίξεις που δεν οδηγούν σε βελτίωση της ποιότητας ζωής όλης της ανθρωπότητας, αλλά αντίθετα επιδεινώνουν τις ανισότητες και τις συγκρούσεις, δεν μπορούν ποτέ να θεωρηθούν πραγματική πρόοδος[8]. Η τεχνητή νοημοσύνη θα αποκτήσει ολοένα και μεγαλύτερη σημασία. Οι προκλήσεις που θέτει είναι τεχνικές, αλλά και ανθρωπολογικές, εκπαιδευτικές, κοινωνικές και πολιτικές. Υπόσχεται, για παράδειγμα, εξοικονόμηση εργασίας, αποδοτικότερη παραγωγή, ευκολότερες μεταφορές και δυναμικότερες αγορές, καθώς και επανάσταση στις διαδικασίες συλλογής, οργάνωσης και επαλήθευσης δεδομένων. Πρέπει να συνειδητοποιήσουμε τους ραγδαίους μετασχηματισμούς που συντελούνται και να τους διαχειριστούμε με τρόπο που να διασφαλίζει τα θεμελιώδη ανθρώπινα δικαιώματα, σεβόμενοι τους θεσμούς και τους νόμους που προάγουν την ολοκληρωμένη ανθρώπινη ανάπτυξη. Η τεχνητή νοημοσύνη θα πρέπει να βρίσκεται στην υπηρεσία του καλύτερου ανθρώπινου δυναμικού και των υψηλότερων προσδοκιών μας και όχι να τα ανταγωνίζεται.

3. Η τεχνολογία του μέλλοντος: μηχανές που μαθαίνουν μόνες τους

Στις πολλές μορφές της, η τεχνητή νοημοσύνη, που βασίζεται σε τεχνικές μηχανικής μάθησης, αν και βρίσκεται ακόμη σε πρωτοποριακό στάδιο, εισάγει ήδη σημαντικές αλλαγές στον ιστό των κοινωνιών, ασκώντας βαθιά επίδραση στους πολιτισμούς, την κοινωνική συμπεριφορά και την οικοδόμηση της ειρήνης. Εξελίξεις όπως η μηχανική μάθηση (*machine learning*) ή η βαθιά μάθηση (*deep learning*) εγείρουν ερωτήματα που υπερβαίνουν τα πεδία της τεχνολογίας και της μηχανικής και έχουν να κάνουν με μια κατανόηση που συνδέεται στενά με το νόημα της ανθρώπινης ζωής, τις βασικές διαδικασίες της γνώσης και την ικανότητα του νου να φθάνει στην αλήθεια. Η ικανότητα ορισμένων μηχανισμών να παράγουν συντακτικά και σημασιολογικά συνεκτικά κείμενα, για παράδειγμα, δεν αποτελεί εγγύηση αξιοπιστίας. Λέγεται ότι

είναι σε θέση να “δημιουργούν παραισθήσεις”, δηλαδή να παράγουν δηλώσεις που εκ πρώτης όψεως φαίνονται αληθοφανείς, αλλά στην πραγματικότητα είναι αβάσιμες ή προδίδουν προκαταλήψεις. Αυτό δημιουργεί σοβαρό πρόβλημα όταν η τεχνητή νοημοσύνη χρησιμοποιείται σε εκστρατείες παραπληροφόρησης που διαδίδουν ψευδείς ειδήσεις και οδηγούν σε αυξανόμενη δυσπιστία απέναντι στα μέσα ενημέρωσης. Η εμπιστευτικότητα, η κατοχή δεδομένων και η πνευματική ιδιοκτησία είναι άλλοι τομείς στους οποίους οι εν λόγω τεχνολογίες ενέχουν σοβαρούς κινδύνους, στους οποίους προστίθενται περαιτέρω αρνητικές συνέπειες της κακής χρήσης τους, όπως οι διακρίσεις, η παρέμβαση στις εκλογικές διαδικασίες, η επικράτηση μιας κοινωνίας που παρακολουθεί και ελέγχει τους ανθρώπους, ο ψηφιακός αποκλεισμός και η επιδείνωση ενός ατομικισμού που αποσυνδέεται όλο και περισσότερο από την κοινότητα. Όλοι αυτοί οι παράγοντες κινδυνεύουν να τροφοδοτούν τις συγκρούσεις και να εμποδίζουν την ειρήνη.

4. Η αίσθηση του ορίου στο τεχνοκρατικό παράδειγμα

Ο κόσμος μας είναι υπερβολικά μεγάλος, ποικιλόμορφος και πολύπλοκος για να είναι πλήρως γνωστός και ταξινομημένος. Το ανθρώπινο μυαλό δεν μπορεί ποτέ να εξαντλήσει τον πλούτο του, ακόμη και με τη βοήθεια των πιο προηγμένων αλγορίθμων. Αυτοί, μάλιστα, δεν προσφέρουν εγγυημένες προβλέψεις για το μέλλον, αλλά μόνο στατιστικές προσεγγίσεις. Δεν μπορούν να προβλεφθούν τα πάντα, δεν μπορούν να υπολογιστούν τα πάντα· τελικά, “η πραγματικότητα είναι ανώτερη από την ιδέα”[9] και, όσο θαυμαστή κι αν είναι η υπολογιστική μας δύναμη, θα υπάρχει πάντα ένα απρόσιτο υπόλειμμα που θα διαφεύγει από κάθε προσπάθεια μέτρησης. Επιπλέον, ο τεράστιος όγκος δεδομένων που αναλύεται από τις τεχνητές νοημοσύνες δεν αποτελεί από μόνος του εγγύηση αμεροληψίας. Όταν οι αλγόριθμοι εξάγουν πληροφορίες, διατρέχουν πάντα τον κίνδυνο να τις παραμορφώσουν, αναπαράγοντας τις αδικίες και τις προκαταλήψεις του εκάστοτε περιβάλλοντος από το οποίο προέρχεται. Όσο πιο γρήγοροι και πιο πολύπλοκοι γίνονται, τόσο πιο δύσκολο είναι να καταλάβουμε γιατί έχουν παράγει ένα συγκεκριμένο αποτέλεσμα.

Οι “ευφυείς” μηχανές μπορεί να εκτελούν τα καθήκοντα που τους ανατίθενται με αυξανόμενη αποτελεσματικότητα, αλλά ο σκοπός και το νόημα των λειτουργιών τους θα συνεχίσουν να καθορίζονται ή να ενεργοποιούνται από τα ανθρώπινα όντα που διαθέτουν το δικό τους σύμπαν αξιών. Ο κίνδυνος είναι ότι τα κριτήρια που διέπουν ορισμένες επιλογές γίνονται λιγότερο σαφή, ότι η ευθύνη λήψης αποφάσεων αποκρύπτεται και ότι οι παραγωγοί μπορεί να αποφύγουν την υποχρέωσή τους να ενεργούν για το καλό της κοινότητας. Κατά μία έννοια, αυτό ευνοείται από το τεχνοκρατικό σύστημα, το οποίο κάνει να συμμαχούν η οικονομία και η τεχνολογία και προκρίνει το κριτήριο της αποτελεσματικότητας, τείνοντας να αγνοεί οτιδήποτε δεν σχετίζεται με τα άμεσα συμφέροντά του[10]. Αυτό θα πρέπει να μας κάνει να σκεφτούμε μια πτυχή που τόσο συχνά παραβλέπεται στη σημερινή τεχνοκρατική και αποδοτική νοοτροπία, ως καθοριστική για την προσωπική και κοινωνική ανάπτυξη: την “αίσθηση των ορίων”. Ο άνθρωπος, στην πραγματικότητα, θνητός εξ ορισμού, που σκέπτεται να ξεπεράσει κάθε όριο χάρη στην τεχνολογία, κινδυνεύει, με την εμμονή του να θέλει να ελέγχει τα πάντα, να χάσει τον έλεγχο του εαυτού του·

αναζητώντας την απόλυτη ελευθερία, κινδυνεύει να πέσει στο σπιράλ μιας τεχνολογικής δικτατορίας. Η αναγνώριση και η αποδοχή του ορίου του ως πλάσμα είναι απαραίτητη προϋπόθεση για να φτάσει ο άνθρωπος στην πληρότητα, ή μάλλον να την υποδεχτεί ως δώρο. Αντίθετα, στο ιδεολογικό πλαίσιο ενός τεχνοκρατικού παραδείγματος, που εμφορείται από την προμηθεϊκή παραδοχή της αυτάρκειας, οι ανισότητες θα μπορούσαν να αυξηθούν υπέρμετρα και η γνώση και ο πλούτος να συσσωρευτούν στα χέρια λίγων, με σοβαρούς κινδύνους για τις δημοκρατικές κοινωνίες και την ειρηνική συνύπαρξη[11].

5. Καυτά θέματα για την ηθική

Στο μέλλον, η αξιοπιστία ενός που ζητάει δάνειο, η καταλληλότητα ενός ατόμου για μια θέση εργασίας, η πιθανότητα υποτροπής ενός καταδικασμένου ατόμου ή το δικαίωμα λήψης πολιτικού ασύλου ή κοινωνικής βοήθειας μπορεί να καθορίζονται από συστήματα τεχνητής νοημοσύνης. Η έλλειψη διαφορετικών επιπέδων διαμεσολάβησης που εισάγουν αυτά τα συστήματα είναι ιδιαίτερα επιρρεπής σε μορφές προκατάληψης και διακρίσεων: τα συστημικά σφάλματα μπορούν εύκολα να πολλαπλασιαστούν, παράγοντας όχι μόνο αδικίες σε μεμονωμένες περιπτώσεις αλλά και, μέσω του φαινομένου ντόμινο, πραγματικές μορφές κοινωνικής ανισότητας. Μερικές φορές, εξάλλου, οι μορφές τεχνητής νοημοσύνης φαίνεται να είναι σε θέση να επηρεάζουν τις αποφάσεις των ατόμων μέσω προκαθορισμένων επιλογών που συνδέονται με ερεθίσματα και αποτρεπτικούς παράγοντες ή μέσω συστημάτων ρύθμισης των προσωπικών επιλογών που βασίζονται στην οργάνωση των πληροφοριών. Αυτές οι μορφές χειραγώγησης ή κοινωνικού ελέγχου απαιτούν προσεκτική προσοχή και εποπτεία και συνεπάγονται σαφή νομική ευθύνη εκ μέρους των παραγωγών, αυτών που τις χρησιμοποιούν και των κυβερνητικών αρχών.

Η εξάρτηση από αυτόματες διαδικασίες που κατηγοριοποιούν τα άτομα, π.χ. μέσω της διάχυτης χρήσης της παρακολούθησης ή της υιοθέτησης συστημάτων κοινωνικής πιστοποίησης, θα μπορούσε επίσης να έχει βαθιές επιπτώσεις στον αστικό ιστό, σταθεροποιώντας τις ακατάλληλες κατατάξεις μεταξύ των πολιτών. Και αυτές οι τεχνητές διαδικασίες κατάταξης θα μπορούσαν επίσης να οδηγήσουν σε συγκρούσεις εξουσίας, καθώς δεν αφορούν μόνο εικονικά πρόσωπα, αλλά ανθρώπους με σάρκα και οστά. Ο θεμελιώδης σεβασμός της ανθρώπινης αξιοπρέπειας προϋποθέτει την άρνηση της ταύτισης της μοναδικότητας του ατόμου με ένα σύνολο δεδομένων. Οι αλγόριθμοι δεν πρέπει να επιτρέπεται να καθορίζουν τον τρόπο με τον οποίο αντιλαμβανόμαστε τα ανθρώπινα δικαιώματα, να παραμερίζουν τις βασικές αξίες της συμπόνιας, του ελέους και της συγχώρεσης ή να εξαλείφουν τη δυνατότητα ενός ατόμου να αλλάξει και να αφήσει πίσω του το παρελθόν. Σε αυτό το πλαίσιο, δεν μπορούμε να μην εξετάσουμε τον αντίκτυπο των νέων τεχνολογιών στον εργασιακό χώρο: εργασίες που κάποτε αποτελούσαν αποκλειστικό προνόμιο της ανθρώπινης εργασίας απορροφώνται ταχύτατα από βιομηχανικές εφαρμογές τεχνητής νοημοσύνης. Και εδώ υπάρχει σημαντικός κίνδυνος δυσανάλογου πλεονεκτήματος για λίγους εις βάρος της φτωχοποίησης πολλών. Ο σεβασμός της αξιοπρέπειας των εργαζομένων και η σημασία της απασχόλησης για την οικονομική ευημερία των ατόμων, των οικογενειών και των κοινωνιών, η ασφάλεια των θέσεων εργασίας και

οι δίκαιοι μισθοί θα πρέπει να αποτελούν υψηλή προτεραιότητα για τη διεθνή κοινότητα, ενώ αυτές οι μορφές τεχνολογίας διεισδύουν όλο και πιο βαθιά στους χώρους εργασίας.

6.Θα μετατρέψουμε τα σπαθιά σε αλέτρι;

Στις μέρες μας, κοιτάζοντας τον κόσμο γύρω μας, δεν μπορεί κανείς να ξεφύγει από τα σοβαρά ηθικά ζητήματα που σχετίζονται με την εξοπλιστική βιομηχανία. Η δυνατότητα διεξαγωγής στρατιωτικών επιχειρήσεων μέσω τηλεχειριζόμενων συστημάτων έχει οδηγήσει σε μια αυξημένη αντίληψη της καταστροφής που προκαλούν και της ευθύνης για τη χρήση τους, συμβάλλοντας σε μια ακόμη πιο ψυχρή και αποστασιοποιημένη προσέγγιση της τεράστιας τραγωδίας του πολέμου. Η έρευνα σχετικά με τις αναδυόμενες τεχνολογίες στον τομέα των λεγόμενων “θανατηφόρων αυτόνομων οπλικών συστημάτων”, συμπεριλαμβανομένης της χρήσης τεχνητής νοημοσύνης σε καιρό πολέμου, αποτελεί σοβαρό ηθικό ζήτημα. Τα αυτόνομα οπικά συστήματα δεν μπορούν ποτέ να είναι ηθικά υπεύθυνα υποκείμενα: η μοναδική ανθρώπινη ικανότητα για ηθική κρίση και ηθική λήψη αποφάσεων είναι κάτι περισσότερο από ένα πολύπλοκο σύνολο αλγορίθμων, και αυτή η ικανότητα δεν μπορεί να αναχθεί στον προγραμματισμό μιας μηχανής που, όσο “ευφυής” και αν είναι, παραμένει μια μηχανή. Για τον λόγο αυτόν, είναι επιτακτική ανάγκη να εξασφαλιστεί επαρκής, ουσιαστική και συνεπής ανθρώπινη εποπτεία των οπλικών συστημάτων. Επίσης, δεν μπορούμε να αγνοήσουμε την πιθανότητα να πέσουν εξελιγμένα όπλα σε λάθος χέρια, διευκολύνοντας, για παράδειγμα, τρομοκρατικές επιθέσεις ή επεμβάσεις με στόχο την αποσταθεροποίηση νόμιμων κυβερνητικών θεσμών. Εν ολίγοις, ο κόσμος δεν χρειάζεται πραγματικά τις νέες τεχνολογίες που συμβάλλουν στην άδικη ανάπτυξη της αγοράς και του εμπορίου όπλων, προωθώντας την τρέλα του πολέμου. Με αυτόν τον τρόπο, όχι μόνο η νοημοσύνη, αλλά και η ίδια η ανθρώπινη καρδιά, θα κινδυνεύσει να γίνει όλο και πιο “τεχνητή”. Οι πιο προηγμένες τεχνικές εφαρμογές δεν πρέπει να χρησιμοποιούνται για να διευκολύνουν τη βίαιη επίλυση συγκρούσεων, αλλά για να ανοίξουν τον δρόμο για την ειρήνη.

Σε μια οπτική πιο θετική, αν η τεχνητή νοημοσύνη χρησιμοποιούνταν για την προώθηση της ολοκληρωμένης ανθρώπινης ανάπτυξης, θα μπορούσε να εισαγάγει σημαντικές καινοτομίες στη γεωργία, την εκπαίδευση και τον πολιτισμό, τη βελτίωση του βιοτικού επιπέδου ολόκληρων εθνών και λαών και την ανάπτυξη της ανθρώπινης αδελφοσύνης και της κοινωνικής φιλίας. Τελικά, το πώς θα τη χρησιμοποιήσουμε για να συμπεριλάβουμε τους έσχατους, δηλαδή τους πιο αδύναμους και άπορους αδελφούς και αδελφές μας, είναι το μέτρο που αποκαλύπτει την ανθρωπιά μας. Το ανθρώπινο βλέμμα και η επιθυμία για ένα καλύτερο μέλλον για τον κόσμο οδηγούν στην ανάγκη ενός διεπιστημονικού διαλόγου με στόχο την ηθική ανάπτυξη των αλγορίθμων – την *αλγορητική* – στην οποία οι αξίες καθοδηγούν τα μονοπάτια των νέων τεχνολογιών[12]. Τα δεοντολογικά ζητήματα πρέπει να λαμβάνονται υπόψη από την αρχή της έρευνας, καθώς και στις φάσεις δοκιμής, σχεδιασμού, παραγωγής, διανομής και εμπορίας. Αυτή είναι η προσέγγιση της ηθικής

του σχεδιασμού, στην οποία τα εκπαιδευτικά ιδρύματα και οι υπεύθυνοι λήψης αποφάσεων έχουν ουσιαστικό ρόλο να διαδραματίσουν.

7. Προκλήσεις για την εκπαίδευση

Η ανάπτυξη της τεχνολογίας που σέβεται και υπηρετεί την ανθρώπινη αξιοπρέπεια έχει σαφείς συνέπειες για τα εκπαιδευτικά ιδρύματα και τον κόσμο του πολιτισμού. Πολλαπλασιάζοντας τις δυνατότητες επικοινωνίας, οι ψηφιακές τεχνολογίες έχουν επιτρέψει στους ανθρώπους να συναντηθούν μεταξύ τους με νέους τρόπους. Ωστόσο, εξακολουθεί να υπάρχει ανάγκη για συνεχή προβληματισμό σχετικά με το είδος των σχέσεων στις οποίες μας οδηγούν. Οι νέοι άνθρωποι μεγαλώνουν σε πολιτιστικά περιβάλλοντα που διαποτίζονται από την τεχνολογία και αυτό δεν μπορεί παρά να προκαλεί τις μεθόδους διδασκαλίας και κατάρτισης. Η εκπαίδευση στη χρήση μορφών τεχνητής νοημοσύνης θα πρέπει να στοχεύει πάνω απ' όλα στην προώθηση της κριτικής σκέψης. Είναι απαραίτητο οι χρήστες όλων των ηλικιών, αλλά κυρίως οι νέοι, να αναπτύξουν την ικανότητα διάκρισης κατά τη χρήση των δεδομένων και του περιεχομένου που συλλέγονται στο διαδίκτυο ή παράγονται από συστήματα τεχνητής νοημοσύνης.

Τα σχολεία, τα πανεπιστήμια και οι επιστημονικές εταιρείες καλούνται να βοηθήσουν τους φοιτητές και τους επαγγελματίες να κατανοήσουν τις κοινωνικές και ηθικές πτυχές της ανάπτυξης και της χρήσης της τεχνολογίας. Η εκπαίδευση στη χρήση των νέων εργαλείων επικοινωνίας θα πρέπει να λαμβάνει υπόψη όχι μόνο την παραπληροφόρηση και τις ψευδείς ειδήσεις, αλλά και την ανησυχητική αναζωπύρωση “προγονικών φόβων [...] που μπόρεσαν να κρυφτούν και να αναπτυχθούν πίσω από τις νέες τεχνολογίες”[13]. Δυστυχώς, για άλλη μια φορά βρισκόμαστε αντιμέτωποι με “τον πειρασμό να δημιουργήσουμε μια κουλτούρα τειχών, να υψώσουμε τείχη για να αποτρέψουμε τις συναντήσεις με άλλους πολιτισμούς, με άλλους ανθρώπους”[14] και την ανάπτυξη της ειρηνικής και αδελφικής συνύπαρξης.

8. Προκλήσεις για την ανάπτυξη του διεθνούς δικαίου

Η παγκόσμια εμβέλεια της τεχνητής νοημοσύνης καθιστά σαφές ότι, πέραν της ευθύνης των κυρίαρχων κρατών να ρυθμίζουν τη χρήση της στο εσωτερικό τους, οι διεθνείς οργανισμοί μπορούν να διαδραματίσουν καθοριστικό ρόλο στην επίτευξη πολυμερών συμφωνιών και στο συντονισμό της εφαρμογής και της υλοποίησής τους[15]. Στο πλαίσιο αυτό, καλώ την Κοινότητα των Εθνών να συνεργαστεί για την υιοθέτηση μιας δεσμευτικής διεθνούς συνθήκης που θα ρυθμίζει την ανάπτυξη και τη χρήση της τεχνητής νοημοσύνης στις πολλές μορφές της. Στόχος της ρύθμισης, βέβαια, δεν πρέπει να είναι μόνο η πρόληψη των κακών πρακτικών, αλλά και η ενθάρρυνση των καλών πρακτικών, η τόνωση νέων και δημιουργικών προσεγγίσεων και η διευκόλυνση των προσωπικών και συλλογικών πρωτοβουλιών[16]. Τελικά, στην αναζήτηση ρυθμιστικών μοντέλων που μπορούν να παρέχουν ηθική καθοδήγηση στους προγραμματιστές ψηφιακών τεχνο-λογιών, είναι επιτακτική ανάγκη να προσδιοριστούν οι ανθρώπινες αξίες που θα πρέπει να στηρίζουν τις προσπάθειες των κοινωνιών να διαμορφώσουν, να υιοθετήσουν και να επιβάλουν τα

απαραίτητα νομοθετικά πλαίσια. Το έργο της σύνταξης ηθικών κατευθυντήριων γραμμών για την παραγωγή μορφών τεχνητής νοημοσύνης δεν μπορεί να παραβλέπει την εξέταση βαθύτερων ζητημάτων που αφορούν το νόημα της ανθρώπινης ύπαρξης, την προστασία των θεμελιωδών ανθρωπίνων δικαιωμάτων και την επιδίωξη της δικαιοσύνης και της ειρήνης.

Αυτή η διαδικασία ηθικής και νομικής διάκρισης μπορεί να αποτελέσει ένα πολύτιμο προοίμιο για έναν κοινό προβληματισμό σχετικά με τον ρόλο που πρέπει να διαδραματίζει η τεχνολογία στην ατομική και κοινοτική μας ζωή και πώς η χρήση της μπορεί να συμβάλει στη δημιουργία ενός δικαιότερου και πιο ανθρώπινου κόσμου. Για τον λόγο αυτόν, οι φωνές όλων των ενδιαφερομένων, συμπεριλαμβανομένων των φτωχών, των περιθωριοποιημένων και άλλων που συχνά κανείς δεν τους ακούει στις παγκόσμιες διαδικασίες λήψης αποφάσεων, θα πρέπει να λαμβάνονται υπόψη στις συζητήσεις για τη ρύθμιση της τεχνητής νοημοσύνης.

* * * * *

Ελπίζω αυτός ο προβληματισμός να ενθαρρύνει ώστε η πρόοδος στην ανάπτυξη των μορφών τεχνητής νοημοσύνης να υπηρετήσει τελικά την υπόθεση της ανθρώπινης ειρήνης και ελευθερίας. Αυτό δεν είναι ευθύνη λίγων, αλλά ολόκληρης της ανθρώπινης οικογένειας. Η ειρήνη, στην πραγματικότητα, είναι ο καρπός των σχέσεων που αναγνωρίζουν και αποδέχονται τον άλλον στην αναφαιρέτη αξιοπρέπειά του, και της συνεργασίας και της δέσμευσης στην αναζήτηση της ολοκληρωμένης ανάπτυξης όλων των προσώπων και των λαών.

Η προσευχή μου στην αρχή του νέου έτους είναι να μην αυξήσει, η ραγδαία ανάπτυξη των μορφών τεχνητής νοημοσύνης, τις πολλές ανισότητες και αδικίες που ήδη υπάρχουν στον κόσμο, αλλά να συμβάλει στον τερματισμό των πολέμων και των συγκρούσεων και να ανακουφίσει πολλές μορφές πόνου που πλήττουν την ανθρώπινη οικογένεια. Είθε οι χριστιανοί πιστοί, οι πιστοί διαφόρων θρησκειών και οι άνδρες και οι γυναίκες καλής θελήσεως, να συνεργαστούν αρμονικά για να αδράξουν τις ευκαιρίες και να αντιμετωπίσουν τις προκλήσεις που θέτει η ψηφιακή επανάσταση και να παραδώσουν στις μελλοντικές γενεές έναν κόσμο μεγαλύτερης αλληλεγγύης, δικαιοσύνης και ειρήνης.

Από το Βατικανό, 8 Δεκεμβρίου 2023

ΦΡΑΓΚΙΣΚΟΣ