**Προς:**

[Εισάγετε το όνομα του βουλευτή, εκπροσώπου, σενατόρου, δημάρχου κ.λπ.]

[Εισάγετε την επίσημη διεύθυνση]

**Από:**

[Το όνομά σας]

[Η διεύθυνσή σας]

Σχετικά με την ανάγκη άμεσης ρυθμιστικής δράσης και συνθηκών για την ανακοπή της γενικής τεχνητής νοημοσύνης (AGI, από τα αγγλικά)

Αγαπητέ/ή [Εισάγετε την επίσημη τιμή και το όνομα/διεύθυνση του βουλευτή, εκπροσώπου, σενατόρου, δημάρχου κ.λπ.],

Το όνομά μου είναι [το όνομά σας] και ζω στην πόλη/περιοχή [τοποθεσία].

Πρόσφατα, θορυβήθηκα για τους κινδύνους της τεχνητής νοημοσύνης (ΤΝ), τις βαθιές επιπτώσεις που θα έχει στην κοινωνία και την έλλειψη δημόσιας συζήτησης (πόσο μάλλον ενημερωμένης δημόσιας συζήτησης) για το θέμα αυτό.

Υπάρχουν πολλοί άμεσοι κίνδυνοι από τα ήδη υπάρχοντα συστήματα τεχνητής νοημοσύνης όπως το Chat-GPT ή το Midjourney, όπως η παραπληροφόρηση ή η ακατάλληλη εφαρμογή σε διάφορες επιχειρήσεις. Στο όχι και τόσο μακρινό μέλλον, θα πρέπει να ενεργοποιηθούν ορισμένα δίχτυα ασφαλείας (όπως μια πολιτική καθολικού βασικού εισοδήματος) σε περίπτωση μαζικής ανεργίας λόγω εκτόπισης θέσεων εργασίας από ρομπότ και συστήματα ΤΝ.

**Αλλά η μεγαλύτερη ανησυχία είναι η ταχύτητα με την οποία βαδίζουμε προς την AGI (τεχνητή γενική νοημοσύνη)** - συστήματα που θα έχουν γνωστικές ικανότητες στο επίπεδο του ανθρώπου ή και πάνω από αυτό. Η διαφορά μιας AGI από έναν άνθρωπο είναι ότι μια AGI χρειάζεται λίγα λεπτά για να αντιγράψει τον εαυτό της και να εξαπλωθεί σαν πυρκαγιά. Η AGI θα σκέφτεται με ταχύτητα υπολογιστή και δεν θα υπόκειται στην επιβολή οποιουδήποτε ανθρώπινου νόμου. Θα είναι σε θέση να κυριαρχήσει στο περιβάλλον μας και σε εμάς ως είδος.

Τους τελευταίους μήνες σημειώθηκε μεγάλη πρόοδος στην πορεία προς την AGI. Υπάρχουν γιγαντιαία πειράματα που έχουν αναληφθεί από εταιρείες όπως η Google και η Microsoft, τα οποία μέχρι στιγμής έχουν αποδώσει καρπούς στη δημιουργία πραγματικά έξυπνων συστημάτων. Επιπλέον, το OpenAI και άλλα ερευνητικά εργαστήρια που βρίσκονται πίσω από ένα μεγάλο μέρος της προόδου έχουν ρητό στόχο την ανάπτυξη AGI.

Οι πρόσφατες επιτυχίες τους προκάλεσαν ένα κύμα επενδύσεων σε δυνατότητες τεχνητής νοημοσύνης, ιδίως στις ΗΠΑ. Ενώ οι εταιρείες προβάλλουν ισχυρισμούς σχετικά με την ασφάλεια, την "ευθυγράμμιση" κ.λπ., πολλοί από αυτούς τους ισχυρισμούς δεν συμφωνούν με τις απόψεις των ερευνητών σχετικά με το θέμα της ευθυγράμμισης της τεχνητής νοημοσύνης. Οι εταιρείες γενικά λαμβάνουν αποφάσεις μονομερώς, χωρίς τη συγκατάθεση του κοινού μέσω οποιουδήποτε δημοκρατικού μέσου.

Η πλημμύρα χρημάτων που εξαπολύθηκε στη βιομηχανία ΤΝ τους τελευταίους μήνες θα τους αγοράσει λομπίστες και δημόσιους υποστηρικτές. **Επομένως, είναι ζωτικής σημασίας οι εκλεγμένοι αντιπρόσωποί μας να πάρουν τώρα τα πράγματα στα χέρια τους προς το συμφέρον του κοινού.**

Υπάρχουν πολλές συζητήσεις για το θέμα αυτό στο Twitter. Ελπίζω να διαβάσετε το άρθρο του Eliezer Yudkowsky με τίτλο **"Pausing AI Developments Isn't Enough. We Need to Shut it All Down".** Ακόμη και αν κάποιος διαφωνεί με τον κ. Yudkowsky, η λογική προσέγγιση θα ήταν να σταματήσει η ανάπτυξη και κλιμάκωση νέων δυνατοτήτων ΤΝ, ώστε να δοθεί χρόνος για δημόσιο διάλογο και γενική κατανόηση των κινδύνων.

Οι περισσότεροι άνθρωποι στη χώρα μας δεν κατανοούν τι θα σήμαινε η ΑΓΙ για τους ίδιους προσωπικά και για τον κόσμο γενικότερα. Δεν το έχουν σκεφτεί, δεν το διάβασαν στις ειδήσεις, δεν έχουν τεχνικές γνώσεις για να αξιολογήσουν είτε την πρόοδο προς αυτή την κατεύθυνση είτε τα πιθανά σενάρια.

Αυτό που επίσης είναι ανησυχητικό είναι οι παραδοχές των δημιουργών μιας τέτοιας τεχνολογίας. Το OpenAI παραδέχεται ότι δεν έχει έναν αλάνθαστο τρόπο να ευθυγραμμίσει ένα μοντέλο που είναι σημαντικά εξυπνότερο από εμάς. **Η Microsoft παραδέχεται ότι δεν καταλαβαίνει τι συμβαίνει στο εσωτερικό του GPT-4.**

Η λανθασμένη αποφασιστικότητα των εργαστηρίων τεχνητής νοημοσύνης, η πλημμύρα χρημάτων, η έλλειψη κανονισμών, όλα αυτά λειτουργούν αδυσώπητα για να επιταχύνουν το ρυθμό της τεχνικής προόδου προς την ΑΓΙ. Η κυβέρνησή μας πρέπει να δράσει εξαιρετικά γρήγορα για να εργαστεί πάνω σε συνθήκες που θα βάλουν φρένο στα γιγαντιαία πειράματα Τεχνητής Νοημοσύνης. Ο κ. Yudkowsky προτείνει μια τέτοια συνθήκη: **να σταματήσει η εκπαίδευση μεγάλων μοντέλων παγκοσμίως μέχρι να μπορέσουμε με βεβαιότητα να επιτύχουμε την ευθυγράμμιση των στόχων και των αξιών της ΤΝ με τις δικές μας.**

Υπάρχουν πολλές ομάδες σε όλο τον κόσμο που ξεπηδούν αυτή τη στιγμή και προωθούν μια λογική πολιτική προς αυτή την κατεύθυνση. Για παράδειγμα, η Εκστρατεία για την Ασφάλεια της Τεχνητής Νοημοσύνης κατέγραψε τις συστάσεις της. Το Ινστιτούτο Μέλλον της Ζωής έχει το δικό του έγγραφο.

Οι άνθρωποι δεν έζησαν ποτέ με "ψηφιακά μυαλά" χίλιες φορές ταχύτερα και οξύτερα από εμάς. Δεν μπορούμε να αφήσουμε μικρές ομάδες ανθρώπων να αποφασίσουν για τη μοίρα της ανθρωπότητας.

Με τους καλύτερους χαιρετισμούς,