ePublishing

Παιδαγωγικός Λόγος

Τόμ. 32, Αρ. 1 (2026)

Λόγος περί της Τεχνητής Νοημοσύνης

Τεχνητή Νοημοσύνη και Αριστοτελική Επιείκεια

Αλέξανδρος Νούνεσης

doi: 10.12681/plogos.33695

Copyright © 2026, Αλέξανδρος Νούνεσης

Άδεια χρήσης Creative Commons Attribution-NonCommercial-NoDerivatives 4.0.

Βιβλιογραφική αναφορά:

Νούνεσης Α. (2026). Τεχνητή Νοημοσύνη και Αριστοτελική Επιείκεια. *Παιδαγωγικός Λόγος*, *32*(1), 65–80. https://doi.org/10.12681/plogos.33695

Αλέξανδρος ΝΟΥΝΕΣΗΣ

Τεχνητή Νοημοσύνη και Αριστοτελική Επιείκεια

doi:https://doi.org/10.12681/plogos.33695

Πρόλογος

Η ΚΙΝΑ ΠΕΡΗΦΑΝΕΎΕΤΑΙ ΟΤΙ ΚΑΤΑΦΕΡΕ ΝΑ ΔΗΜΙΟΥΡΓΉΣΕΙ Τεχνητή Νοημοσύνη (AI) που μπορεί να πάρει τη θέση ενός δικαστή. Ανακάλυψα αυτό το γεγονός στις αρχές του 2022 καθώς ερευνούσα βιβλιογραφία σε σχέση με την πτυχιακή μου εργασία (καθώς και μια ακόμη εργασία που σχετιζόταν με την Τεχνητή Νοημοσύνη). Συνέχεια έπεφτα πάνω σε άρθρα ειδήσεων που μιλούσαν για αυτήν την AI της Κίνας, και όλα τα άρθρα επισήμαναν πως αυτή η AI υποτίθεται ότι φτάνει σε σωστές δικαστικές αποφάσεις με ποσοστό επιτυχίας 97% (Chen 2021; Newman 2021; Samson 2021; The Korea Times 2021).

Αυτό το 97% μου έκανε πραγματικά εντύπωση. Τι σημαίνει αυτή η επιτυχία; Η AI-δικαστής εκδικάζει δίκαια 97% όλων των υποθέσεων; Άρα τρεις στις εκατό φορές κάνει λάθος; Και τι είναι όμως το λάθος και το σωστό στο πλαίσιο της Δικαιοσύνης; Πως ορίζει η AI το τι είναι το Δίκαιο και το Άδικο;

Κάπως έτσι σπίθισε μέσα μου η ανάγκη να αναλύσω στον ελεύθερο μου χρόνο την έννοια της Δικαιοσύνης, και να τη συσχετίσω με την υπόσταση της AI. Ήθελα να καταλάβω αν γίνεται μια μηχανή να φέρεται Δίκαια, κι αν μπορεί να μας δείξει και σε μας επιτέλους τι είναι το Δίκαιο (μιας και βασανιζόμαστε χιλιετίες να το καταλάβουμε από μόνοι μας).

Η Δικαιοσύνη είναι μια έννοια αναπόσπαστη από μια ουσιαστική κουβέντα περί Ηθικής. Δεκάδες φιλόσοφοι έχουν παρουσιάσει τις δικές τους ερμηνείες πάνω σε αυτό το θέμα μέσα στις χιλιετίες της ανθρώπινης ιστορίας. Ένας από τους διάσημους και σημαντικότερους φιλοσόφους που

ανέλυσε εκτεταμένα το τι είναι Δικαιοσύνη ήταν ο Αριστοτέλης. Έχοντας λοιπόν προσωπικά μεγάλη εκτίμηση για τη φιλοσοφική σκέψη του Σταγειρίτη, θα ήθελα να εξετάσω σε αυτό το δοκίμιο τέτοιες απόπειρες δημιουργίας AI-δικαστών (όπως αυτές τις Κίνας) μέσα από το πρίσμα της Αριστοτελικά ορισμένης Δικαιοσύνης, όπως αυτή παρουσιάστηκε στο μνημειώδες έργο του, τα Ηθικά Νικομάχεια.[1]

Η Δικαιοσύνη μέσα από το Πρίσμα της Αριστοτελικής Ηθικής

Ο Αριστοτέλης, διαφοροποιούμενος σε ένα βαθμό από το δάσκαλο του Πλάτωνα, καταλάβαινε τη Δικαιοσύνη ως κάτι που δεν έχει μεταφυσική υπόσταση: κάτι που δεν υπάρχει ακριβώς «εκεί έξω». Ο Αριστοτέλης θεωρεί τη Δικαιοσύνη ως κάτι άρρηκτα και αποκλειστικά συνδεδεμένο με τον άνθρωπο. Είναι μια από τις μορφές/εκφάνσεις της Αρετής (και, με έναν τρόπο, η κυρίαρχη) που σημαίνει ότι είναι δυνάμει ένα κομμάτι του χαρακτήρα του ανθρώπου που προϋπάρχει ως δυνατότητά του, και μπορεί τελεολογικά να αναπτυχθεί και να γίνει πραγματικά κομμάτι του χαρακτήρα του, αρκεί να την καλλιεργήσει στη καθημερινότητά του (και από όσο πιο νεαρή ηλικία γίνεται) με Ενάρετες πράξεις.[2] Μερικές άλλες ηθικές Αρετές είναι η Ανδρεία, η Φιλικότητα και η Μεγαλοψυχία. Εάν επιτύχει κάποιος να γίνει Ενάρετος (και συνεπώς Δίκαιος, Ανδρείος κλπ.), αυτό θα αποδεικνύεται συνέχεια πλέον στις πράξεις του και το περιεχόμενο των επιθυμιών του. Η Αρετή του δηλαδή δεν θα είναι κρυφή, αλλά θα φαίνεται από το χαρακτήρα του όταν πρέπει να εκδηλώνεται προς τον κοινωνικό του περίγυρο.[3]

Πέραν του ότι ο Αριστοτέλης θεωρεί τη Δικαιοσύνη ως μια από τις εκφάνσεις της Αρετής, πρέπει να καταλάβουμε τι ακριβώς είναι η Δικαιοσύνη αυτοτελώς. Δίκαιος λοιπόν κατά τον Αριστοτέλη είναι κάποιος που

[1] Μια εκτενής φιλοσοφική σύνοψη και χαρτογράφηση των ηθικών προβληματισμών που εγείρονται από τη δυνατότητα δικαστηριακής χρήσης της AI έχουν παρουσιάσει οι Άλκης Γούναρης και Γιώργος Κωστελέτος στο άρθρο τους «Γράφοντας τον αλγόριθμο του Καλού: Η Τεχνητή Νοημοσύνη ως μηχανή απόδοσης Δικαιοσύνης» (2024). Οι συγγραφείς του άρθρου επιχειρούν να εξετάσουν αυτούς τους προβληματισμούς και μέσα από το πρίσμα κανονιστικών ηθικών θεωριών διαφορετικών από την Αριστοτελική. Το άρθρο τους προσφέρεται ως ένα θεωρητικό υπόβαθρο που συμπληρώνει τη δική μου μελέτη που εστιάζει συγκεκριμένα στη σχέση Αριστοτελικής επιείκειας και AI.

[2] Αριστοτέλης, *Ηθικά Νικομάχεια*, Β.1, 1103a14–1103b1· Β.4, 1105a30–1105b11.

[3] Αριστοτέλης, *Ηθικά Νικομάχεια*, Z.13, 1144b4–17.

ξέρει πότε ένα άτομο αδικείται όταν αυτό συμβαίνει. Ξέρει δηλαδή πότε ένα άτομο *κλέβει* ουσιαστικά (ακόμα και άθελα του) με κάποιο τρόπο ένα άλλο άτομο (ή ακόμα και τον ίδιο του τον εαυτό).

Προφανώς, μια τέτοια κλοπή μπορεί να συμβεί είτε σε θέματα υλικών αγαθών και σωματικής υγείας. Αλλά και, ενδεχομένως, κάποιος θα μπορούσε να υποστηρίξει πως τέτοιες κλοπές μπορούν να συμβούν ακόμα και σε θέματα ψυχολογικής υγείας και επιλογών στην ζωή. Ένας γονιός που καταπιέζει το παιδί του σε συγκεκριμένες δραστηριότητες και κατευθύνσεις στη ζωή, παρόλο που το παιδί δυσανασχετεί… ή ένας νεαρός Αφροαμερικάνος που καταπιάνεται με εγκληματικές δραστηριότητες διότι γεννήθηκε και μεγάλωσε σε γκέτο του Ντιτρόιτ, είναι δυο παραδείγματα αυτών των πιο ασαφών μορφών Αδικίας.

Εν κατακλείδι, ο Ενάρετος (κάποιος που έχει δηλαδή αναπτύξει και την αρετή της Δικαιοσύνης και φέρεται Δίκαια), είναι κάποιος που γνωρίζει πρώτον πως να *μην* κλέψει, και δεύτερον, πως να επανορθωθεί η κλοπή (δηλαδή η Αδικία) που μπορεί να έχει προξενηθεί σε ένα άτομο (για οποιονδήποτε λόγο). Επίσης *πράττει* και *θέλει να πράττει* με βάση αυτές του τις γνώσεις, γιατί αυτές οι πράξεις του προσφέρουν την ύψιστη ευχαρίστηση.[4] Θέλει λοιπόν να φέρεται με ένα τρόπο που αποσκοπεί στο να διατηρείται μια αρμονία και ισότητα μέσα του, και στο κόσμο γύρω του ανεξαρτήτως πλαισίου, άρα ουσιαστικά να λαμβάνει ο καθένας ότι του αξίζει στο μεγαλύτερο δυνατό βαθμό.[5]

Τώρα, σε δεύτερο χρόνο, ο Αριστοτέλης μιλάει για τη Δικαιοσύνη στο πλαίσιο της πολιτικής. Η πολιτική δικαιοσύνη είναι ο τρόπος ρύθμισης και οργάνωσης με στόχο τη διατήρηση μιας γενικής ισότητας μεταξύ πολιτών… δηλαδή να μην κλέβονται, έμμεσα ή άμεσα, οι πολίτες μεταξύ τους. Η πραγμάτωση της πολιτικής δικαιοσύνης φαίνεται να οδηγεί στη θέσπιση και εφαρμογή νόμων, οι οποίοι θα προνοούν για να διατηρείται η Δίκαιη ισότητα μεταξύ των πολιτών, καθώς και να τιμωρούνται οι Αδικίες (με μια επανορθωτική λογική φυσικά) που συμβαίνουν ανάμεσά τους.[6]

Εδώ όμως πιστεύω μπορούμε να επαγάγουμε τρία καίρια σημεία που ξεκαθαρίζουν την διαφορά της Δικαιοσύνης ως Αρετή από το πολιτικό και, επακολούθως, νομικό δίκαιο.

Το πρώτο σημείο είναι ότι η ίδια η θέσπιση νόμων είναι μια πράξη (ή

[4] Αριστοτέλης, *Ηθικά Νικομάχεια*, Α.8, 1099a7–30· Β.3, 1104b5–12.

[5] Αριστοτέλης, *Ηθικά Νικομάχεια*, Β.1, 1103b3–6· Ε.1, 1129a33–1130a13.

[6] Ο Αριστοτέλης αποκαλεί τις δυο αυτές μορφές της Δικαιοσύνης Διανεμητικό και Επανορθωτικό Δίκαιο αντίστοιχα: *Ηθικά Νικομάχεια*, Ε.2, 1130b30–1131a9· Ε.3-4, 1131a29–1132a29.

μια σειρά πράξεων) που συμπωματικά μπορεί να είναι Δίκαιη ή Άδικη. Αυτό σημαίνει πως ένας νομοθέτης μπορεί κάλλιστα να έχει στο μυαλό του να Αδικήσει τους συμπολίτες του για το δικό του κέρδος, και έτσι να θεσπίσει νόμους που ουσιαστικά Αδικούν. Σε άλλη περίπτωση, ένας άλλος νομοθέτης, λόγω άγνοιας, μπορεί απλά να μη καταφέρει να προβλέψει πως η θέσπιση κάποιων νόμων θα προκαλέσουν μεγάλες Αδικίες στο μέλλον.

Αυτός που όντως θα θεσπίσει δίκαιους νόμους είναι κάποιος που είναι ήδη Δίκαιος, και, ως Δίκαιος, θέλει και ξέρει πως να πράξει Δίκαια. Άρα θεσπίζει νόμους που αποσκοπούν και όντως καταφέρνουν στο να διατηρούνται οι Δίκαιες σχέσεις μεταξύ των πολιτών. Επομένως καταλαβαίνουμε πως το νομικό δίκαιο δεν είναι απαραίτητα Δίκαιο όπως το έχει ορίσει ο οποιοσδήποτε, αλλά μόνο συμπωματικά αν, και μόνο αν, έχει θεσπιστεί από Ενάρετους ανθρώπους.[7]

Το δεύτερο καίριο σημείο που πρέπει να έχουμε υπόψιν μας είναι πως όταν κάποιος ακολουθεί απολύτως ρητά τους νόμους, *ακόμα κι αν αυτοί είναι Δίκαιοι*, δεν τον καθιστά εκείνον πραγματικά Αριστοτελικά Δίκαιο, αλλά μόνο νομικά δίκαιο. Όπως είπαμε, οι νόμοι, ακόμα και οι καλύτεροι και Δικαιότεροι, είναι αποτελέσματα των πράξεων των Δίκαιων ανθρώπων. Αν κάποιος είναι Ενάρετος και Δίκαιος, γνωρίζει ήδη τι είναι Δίκαιο και θέλει να το υπηρετεί. Δεν χρειάζεται λοιπόν να έχει το βλέμμα του στραμμένο προς τους νόμους περιμένοντας να του υποδείξουν εκείνοι πως να φερθεί Δίκαια. Τέτοιες υποδείξεις τις χρειάζονται, με διαφορετικούς τρόπους και για διαφορετικούς λόγους, οι μη-Ενάρετοι.[8]

Τέλος, το τρίτο καίριο σημείο που πρέπει να έχουμε στο μυαλό μας είναι ότι οι νόμοι έχουν πάντα έναν αόριστα καθολικό χαρακτήρα που στοχεύει να δημιουργεί και να διατηρεί ισότητα *γενικά*, αλλά και να προσφέρει επανορθωτικές λύσεις στις περισσότερες εκφάνσεις Αδικίας. Ο Αριστοτέλης όμως πιστεύει πως στη πραγματική ζωή πάντα θα εμφανίζονται επιμέρους περιπτώσεις που, όχι μόνο ο νόμος δε θα προσφέρει τη πιο Δίκαιη λύση, αλλά μπορεί κιόλας να γίνει Άδικος ο ίδιος. Αυτό σημαίνει ότι όταν εφαρμόζονται σε τέτοιες συγκεκριμένες περιπτώσεις, ακόμα και οι φαινομενικά δίκαιοι νόμοι που έχουν πραγματικά στόχο τη γενική διατήρηση ισότητας, μπορούν να αποδειχτούν ουσιαστικά Άδικοι.[9]

---

[7] Αριστοτέλης, *Ηθικά Νικομάχεια,* Ε.1, 1129a4-12, 1129a32-b2, 1129b27-1130a13· Ε.5, 11134a1-7· Ε.6, 1134a24-b8· Ε.8, 11135a18-b25.

[8] Αριστοτέλης, *Ηθικά Νικομάχεια,* Ε.1, 1129b15-27· Ε.8, 1136a1-4· Κ.9, 1179b1-1181b13.

[9] Αριστοτέλης, *Ηθικά Νικομάχεια,* Β.2, 1104a1-9· Ε.7, 1134b18-1135a16.

Η Κορύφωση της Αριστοτελικής Δικαιοσύνης

Εδώ είναι λοιπόν που ο Αριστοτέλης παρουσιάζει επιτέλους ποια είναι η κορύφωση της Αρετής της Δικαιοσύνης, μια κορύφωση που ξεδιαλύνει τις επιπλοκές που εμφανίζονται λόγω του τρίτου προαναφερθέντος καίριου σημείου στη κατανόηση μας της Αριστοτελικής Δικαιοσύνης. Αυτή η κορύφωση είναι η Επιείκεια.[10]

Η Επιείκεια, ως προς το νομικό δίκαιο, μοιάζει σαν μια μικρή ελίτσα που «χαλάει» ένα τέλειο αρμονικό πρόσωπο, ενώ ταυτόχρονα, ανεξήγητα, το κάνει ακόμα πιο ωραίο… ή μια λάθος αρμονικά νότα που όμως χρωματίζει και εξυψώνει με μοναδικό τρόπο μια μελωδία, και την κάνει απαράμιλλη. Έτσι και η Επιείκεια είναι μια εξυψωτική δυνατότητα που όμως διαφεύγει από κανόνες και αρμονίες και λογικές.

Ο Επιεικής είναι ο πραγματικά Ενάρετος και Δίκαιος άνθρωπος, κι είναι αυτός που αναγνωρίζει μεν τον γενικά και *σχεδόν τέλειο δυνάμει* Δίκαιο χαρακτήρα των νόμων, αλλά καταλαβαίνει δε ότι υπάρχουν και θα υπάρχουν γεγονότα και συγκυρίες οι οποίες θα τον ξεπερνούν. Άρα δε θα ναι αρκετός ο νόμος για να ανταποκριθεί σε τέτοια γεγονότα και συγκυρίες. Σε αυτές τις περιπτώσεις έρχεται ο πραγματικά Δίκαιος (και άρα Επιεικής), αναγνωρίζοντας την αδυναμία και ανελαστικότητα του νόμου, και δείχνει σε αυτές τις μοναδικές επιμέρους περιπτώσεις Επιείκεια κατά την κρίση του και κατά την αντίστοιχη με τη κρίση του πράξη.[11]

Προφανώς και είναι πρακτικά αδύνατο να κανονικοποιήσουμε ποιες είναι οι Επιεικείς, και άρα Δίκαιες, πράξεις. Στο τέλος της ημέρας, οι Επιεικείς κρίσεις είναι εντελώς στην ευθύνη του Δίκαιου. Και ο Δίκαιος δεν μπορεί απαραίτητα να εξηγήσει γιατί τις κάνει, πέραν της προαίρεσης του για Δίκαιες πράξεις που έχει αναπτύξει. Για αυτόν τον λόγο είναι τόσο δύσκολο να αποδεχτούμε φιλοσοφικά την έννοια της Επιείκειας όταν προσπαθούμε να θεσπίσουμε νόμους και ένα σταθερό σύστημα Δικαιοσύνης. Γιατί ο καθένας θα μπορούσε να επικαλεστεί την δήθεν Επιεική κρίση του για να δικαιολογήσει τις πράξεις του ή τις πράξεις των άλλων.

---

[10] Θα ήθελα να σημειώσω εδώ πως η έννοια της Επιείκειας στο πλαίσιο της Αριστοτελικής Ηθικής δεν αντιστοιχεί στη συνηθισμένη (και ξεκάθαρα πιο στενή) έννοια του «να δείχνεις έλεος». Επιείκεια μπορεί να σημαίνει πράγματι μια συγχωρητική στάση απέναντι σε αυτούς που έπραξαν άδικα ως προς το νόμο, αλλά μπορεί κάλλιστα να σημαίνει και το αντίθετο, όπου ο Επιεικής χρειάζεται να επιβάλλει μεγαλύτερη τιμωρία σε αυτόν που έκανε άδικη πράξη από ότι προβλέπει ο νόμος. Με άλλα λόγια, η Αριστοτελική Επιείκεια μεταφράζεται σωστότερα στα αγγλικά ως “Equity” παρά ως “Mercy”.

[11] Αριστοτέλης, *Ηθικά Νικομάχεια,* Ε.10 1137a31-1138a2.

Εν πάση περιπτώσει, έχοντας σε γενικές γραμμές καταλάβει την Αριστοτελικά ορισμένη Δικαιοσύνη, μας μένει να την εξετάσουμε στο πλαίσιο της Τεχνητής Νοημοσύνης και των AI-δικαστών. Πρώτα όμως πρέπει να ξεκαθαρίσουμε τι ακριβώς είναι η Τεχνητή Νοημοσύνη.

Οι Δυο Τύποι Τεχνητής Νοημοσύνης

Στις μέρες μας, ο όρος «Τεχνητή Νοημοσύνη» χρησιμοποιείται αδιακρίτως για να χαρακτηρίσει οποιαδήποτε τεχνολογική εξέλιξη που συσχετίζεται με την λύση προβλημάτων ως προς την ανάλυση και επεξεργασία δεδομένων. Γι' αυτό το λόγο είναι σημαντικό να εντοπίσουμε σημαντικές διαφοροποιήσεις που προκύπτουν στην φιλοσοφική ανάλυση του ορού. Μελετώντας τα κείμενα διάφορων ακαδημαϊκών και Φιλοσόφων πάνω στο θέμα, όπως το άρθρο των Selmer Bringsjord και Naveen Sundar Govindarajulu με τον ευθύ τίτλο *"Artificial Intelligence",* αντιλήφθηκα πως είναι απαραίτητο να επισημανθεί η ύπαρξη δυο πολύ διαφορετικών ειδών AI, ειδών που έχουν εδραιωθεί και θεμελιωθεί πλέον στην φιλοσοφική κοινότητα. Τα δυο είδη είναι η «Ασθενής» (Weak) και η «Ισχυρή» (Strong) AI. Ο πρώτος που ξεχώρισε τα δυο αυτά είδη ήταν ο John Searle, ένας διαβόητος φιλόσοφος του δευτέρου μισού του 20^{ου} αιώνα, που ασχολήθηκε περισσότερο με φιλοσοφικά ζητήματα γύρω από το Νου (Searle 1980, 417, 421, 423–424).

Η Ασθενής AI είναι τεχνολογία που έχει σχεδιαστεί με σκοπό την επίλυση πολύ συγκεκριμένων τύπων προβλημάτων. Η αποδοτικότητα της Ασθενούς AI στο να επιλύει τέτοια προβλήματα είναι ασύγκριτα μεγαλύτερη από την αντίστοιχη φυσικώς ανθρώπινη αποδοτικότητα. Όμως η πιο σημαντική δυνατότητα της Ασθενούς AI (που την ξεχωρίζει άλλωστε από τους συνηθισμένους υπολογιστές) είναι ότι, μέσω της ανάπτυξης της τεχνολογίας Μηχανικής Μάθησης ("machine learning"), δύναται να εξελίσσει αυτονόμως τις ικανότητές της στο να επιλύει προβλήματα, και άρα να επαυξάνει την αποδοτικότητά της ανωτέρω (Bringsjord and Govindarajulu 2018, sec. 4.1). Μορφές Ασθενούς AI ήδη χρησιμοποιούνται σε δεκάδες, επαγγελματικούς τομείς, όπως στην Ιατρική (όπου η Ασθενής AI μπορεί να πραγματοποιήσει πολύ πιο αποτελεσματικά χειρουργικές επεμβάσεις από ανθρώπους γιατρούς) και τον Πρωταθλητισμό (όπου η Ασθενής AI μπορεί πλέον να ανταγωνιστεί και εύκολα να κατατροπώσει τους μεγαλύτερους ανθρώπους πρωταθλητές δεκάδων διαφορετικών παιχνιδιών, όπως το σκάκι) (Bringsjord and Govindarajulu 2018, sec. 1.5).

Παρά όλα αυτά, πρέπει να υπογραμμιστεί πως κάθε εκδοχή Ασθενούς AI είναι απόλυτα περιορισμένη στο να λύνει ένα και μόνο μοναδικό τύπο προβλημάτων. Αυτό σημαίνει ότι μια εξειδικευμένη AI στο να κερδίζει παρτίδες σκάκι είναι παντελώς άχρηστη στο να κερδίζει παρτίδες τάβλι (πόσο μάλλον να πραγματοποιεί χειρουργικές επεμβάσεις).

Σε αντίθεση με την Ασθενή AI (που έχει πραγματικά εφευρεθεί), η Ισχυρή AI υπάρχει μόνο στην θεωρία. Θεωρητικά λοιπόν, κάθε Ισχυρή AI δεν περιορίζεται σε ένα και μόνο συγκεκριμένο τύπο προβλημάτων, αλλά μπορεί να αναγνωρίσει και να ανταποκριθεί σε οποιοδήποτε πρόβλημα μπορεί να προκύψει, ακριβώς όπως ο άνθρωπος. Αναγνωρίζει έτσι τις ριζικές διαφορές μεταξύ επιτραπέζιων παιχνιδιών, χειρουργικών επεμβάσεων, μαγειρικής, πυρηνικής φυσικής και οποιουδήποτε άλλου τύπου προβλήματος μπορούμε να φανταστούμε, καθώς και ποιοι παράγοντες και μεταβλητές έχουν σχέση με κάθε συγκεκριμένο τύπο προβλήματος.[12]

Επιπλέον, πέραν των φαινομενικά τεχνικών της ικανοτήτων, η Ισχυρή AI θα έχει τη δυνατότητα υπαρξιακής αυτογνωσίας και συνείδησης. Επομένως, ως νοημοσύνη, η Ισχυρή AI δε θα διαφέρει ουσιαστικά από την ανθρώπινη (πέραν του ότι θα είναι αδιανόητα πιο έξυπνη) και θα πρέπει πρακτικά να συνυπολογίζεται από τον άνθρωπο ως ίσο αισθανόμενο ον.

Όμως η Ισχυρή AI, όπως τονίσαμε, υπάρχει μόνο στη θεωρία (και σε ιστορίες επιστημονικής φαντασίας) και το αν είναι κάποτε εφικτή ρεαλιστικά η δημιουργία της είναι ένα φιλοσοφικό (κι όχι τεχνικό όπως πολλοί νομίζουν) ερώτημα που δεν μας απασχολεί στην παρούσα εργασία μας.

[12] Το τεράστιο επιστημολογικό πρόβλημα που προέκυψε από την έρευνα για το πώς θα μπορούσε να εφευρεθεί AI που δεν θα περιορίζεται σε έναν και μόνο τύπο προβλημάτων έχει αναγνωριστεί εδώ και δεκαετίες: Ο άνθρωπος μπορεί σε ριπή οφθαλμού να ξεχωρίσει τις ριζικές διαφορές ανάμεσα σε θεωρητικά άπειρους τύπους προβλημάτων, αλλά και ανάμεσα στους εξίσου άπειρους παράγοντες και μεταβλητές που σχετίζονται με αυτά τα προβλήματα. Για παράδειγμα, όταν ο άνθρωπος θέλει να παίξει σκάκι, δε χρειάζεται να σκεφτεί συνειδητά ότι το αλάτι, τα μαχαίρια, ο φούρνος κι οτιδήποτε άλλο που βρίσκεται στην κουζίνα του και σχετίζεται με τη μαγειρική είναι άσχετα με το σκάκι. Ξέρει (και μαθαίνει) σχεδόν αυτόματα τι είναι σχετικό με το σκάκι, τι με τη μαγειρική, και τι είναι άσχετο και με τα δύο. Αυτή την επιστημολογική ικανότητα των ανθρώπων (που φαίνεται τόσο αυτονόητη στην καθημερινή ανθρώπινη εμπειρία μας) παραμένει αδιευκρίνιστη ως προς τον τρόπο που προκύπτει, και συνεπώς παραμένει άγνωστο πώς θα μπορούσε να προγραμματιστεί σε μια AI. Την πιο καθαρή διατύπωση του προβλήματος παρουσίασε ο Daniel Dennett στο δοκίμιό του *"Cognitive Wheels: The Frame Problem of AI"* (1987). Για μια αναλυτική επισκόπηση του προβλήματος, δείτε επίσης: *"The Frame Problem"* (2016) του Murray Shanahan στο *The Stanford Encyclopedia of Philosophy*. Και τα δύο κείμενα περιλαμβάνονται στη βιβλιογραφία του παρόντος δοκιμίου.

Στη περίπτωση της εξέτασης λοιπόν της AI-δικαστή, όπως αυτή δύναται να προγραμματιστεί βάσει των τεχνολογικών δυνατοτήτων της εποχής μας, πρέπει να εστιάσουμε μόνο στην περίπτωση της Ασθενούς AI.

Μια ρεαλιστικά δυνάμει AI-δικαστής λοιπόν θα είναι μια Ασθενής AI, δηλαδή μια AI που θα είναι συγκλονιστικά αποτελεσματική στο να εκδικάζει δικαστικές υποθέσεις (με την εκδίκαση δικαστικών υποθέσεων να είναι ο απόλυτα μοναδικός σκοπός της). Δε πρέπει σε καμία περίπτωση να μπερδευτούμε και να σκεφτούμε ότι μια AI-δικαστής, όπως αυτή ρεαλιστικά δύναται, κατέχει κάποια μορφή συνείδησης και βούλησης (Bringsjord and Govindarajulu 2018, sec. 8.1.1).

Πώς θα Λειτουργήσει μια AI-δικαστής

Μπορούμε να σκεφτούμε πολλούς λόγους για τους οποίους θα θέλουμε να σχεδιάσουμε μια AI-δικαστή. Για παράδειγμα, επειδή δεν εμπιστευόμαστε τις ανθρώπινες προκαταλήψεις που μπορεί να έχουν άνθρωποι δικαστές εις βάρος (ή υπέρ) μειονοτήτων ή άλλων ανθρώπινων χαρακτηριστικών… ή ίσως ελπίζουμε ότι θα καταπολεμήσουμε τη διαφθορά που μπορεί να υπάρχει στο δικαστικό μας σύστημα λόγω της ανθρώπινης φύσης των δικαστών. Ο πιο σημαντικός όμως λόγος κατά την άποψη μου, που θα θέλαμε να έχουμε AI-δικαστή είναι διότι η AI θα είναι πάρα πολύ πιο αποδοτική στο να εκδικάζει υποθέσεις: Μια AI θα «βλέπει» μια υπόθεση, θα αναλύει τα δεδομένα της υπόθεσης (όπως καταθέσεις και αποδεικτικά στοιχεία) και μέσα σε μερικά λεπτά (αν όχι γρηγορότερα) θα την εκδικάζει, χωρίς ποτέ να αισθάνεται κούραση.[13]

Για να φτάσει σε αυτή την ικανότητα όμως, η AI-δικαστής θα πρέπει πρώτα να «εκπαιδευθεί» στο πως να είναι δικαστής. Αυτό επιτυγχάνεται πράγματι χάρη στη μέθοδο Μηχανικής Μάθησης. Μια AI-δικαστής, για να «εκπαιδευτεί» λοιπόν θα πρέπει να έχει πρόσβαση σε όσα περισσότερα δεδομένα data είναι προσβάσιμα από όλες τις υποθέσεις και εκδικάσεις που έχουν γίνει στο παρελθόν σε ολόκληρη την υφήλιο, και για τις οποίες υπάρχουν αποθηκευμένα και καταγεγραμμένα data. Η AI-δικαστής θα

[13] Για μια πιο ενδελεχή τεκμηρίωση της θέσης ότι η Τεχνητή Νοημοσύνη μπορεί να βελτιώσει ουσιαστικά το δικαστικό σύστημα μιας κοινωνίας (ειδικά ως προς την επιτάχυνση της εκδίκασης νομικών υποθέσεων) δείτε το άρθρο των Cinara Rocha και João Alvaro Carvalho με τίτλο *"Artificial Intelligence in the Judiciary: Uses and Threats"* (2022).

«μάθει» έτσι να κάνει σωστή συσχέτιση νόμων με τις εκδικάσεις των υποθέσεων. Έπειτα, έχοντας τους αλγόριθμους της παραμετροποιημένους από όλους τους νόμους μιας χώρας στην οποία θα λειτουργήσει ως δικαστής, η AI θα μπορούσε να βαλθεί να εκδικάσει (σε προσομοίωση) όλες τις υποθέσεις της χώρας που αντιστοιχούν στους νόμους της και που έχουν εκδικαστεί στο παρελθόν. Τελικώς, συγκρίνοντας τις εκδικάσεις της AI με τις παλιές και τωρινές εκδικάσεις ανθρώπων δικαστών, μπορούμε άμεσα να εξετάσουμε αν φτάνει στις ίδιες αποφάσεις που είχαν φτάσει και φτάνουν άνθρωποι δικαστές.[14]

Αν παρατηρήσουμε εν τέλει σε αυτές τις προσομοιώσεις τρανά ποσοστά αντιστοιχίας μεταξύ των εκδικάσεων της AI με τις εκδικάσεις των ανθρώπων δικαστών (όπως για παράδειγμα 97%), τότε θα μπορούσαμε να συμπεράνουμε ότι η AI ακολουθεί και εφαρμόζει το νόμο πρακτικά ακριβώς όπως οι άνθρωποι δικαστές. Έχοντας στο μυαλό μας και τα προαναφερθέντα πλεονεκτήματα μιας AI-δικαστή (ως προς την ταχύτητα εκδικάσεων και την αδιάφθορη και απροκατάληπτη φύση της) θα φάνταζε πράγματι πολύ λογικό να θέλουμε και στην πράξη να αντικαταστήσουμε τους ανθρώπους δικαστές με μια AI-δικαστή.

Αν όμως ασπαζόμαστε τις φιλοσοφικές θέσεις του Αριστοτέλη ως προς τη Δικαιοσύνη και την Αρετή, τότε μια τέτοια αντικατάσταση δεν θα ήταν καθόλου απλή να την αποδεχτούμε. Το πρόβλημα είναι το εξής: Η AI-δικαστής θα φέρεται πάντα με γενική και καθολική λογική, όσο πολύπλοκη και διεκπεραιωτική μπορεί να φαίνεται πως είναι. Διότι η AI έχει σχεδιαστεί και προγραμματιστεί εξαρχής με έναν πολύ συγκεκριμένο στόχο, καθώς και παραμέτρους που θα υποδεικνύουν πότε και πόσο επιτυγχάνεται αυτός ο στόχος. Αυτός ο στόχος είναι η εφαρμογή του νομικού δικαίου, στην αυθαίρετη μορφή που της έχει παρουσιαστεί. Συνεπώς έχει σχεδιαστεί, όχι να είναι Δίκαια η ίδια, αλλά να μιμηθεί τις εκδικάσεις παλιών υποθέσεων βάσει του εκάστοτε νομικού δικαίου, και να προβλέψει βάσει

[14] Το Κοινοβούλιο της Ευρωπαϊκής Ένωσης έχει αναγνωρίσει ότι η χρήση εφαρμογών AI (ιδίως όταν πρόκειται για AI που προορίζεται για χρήση στο πλαίσιο της απονομής νομικής δικαιοσύνης) ενέχει υψηλό ρίσκο για την παραβίαση ανθρωπίνων δικαιωμάτων και την υπονόμευση δημοκρατικών θεσμών. Συνεπώς, έχουν ήδη προβλεφθεί και κατοχυρωθεί νομικά στην Ευρωπαϊκή Ένωση μέσω του Artificial Intelligence Act (Κανονισμός ΕΕ 2024/1689) παρόμοιες διαδικασίες εκπαίδευσης και ελέγχου AI με αυτή που περιέγραψα συνοπτικά παραπάνω (αν και οι διαδικασίες της ΕΕ είναι, προφανώς, πολύ πιο σύνθετες, θεσμικά αυστηρές και τεχνικά απαιτητικές από τη δικιά μου απλουστευμένη περιγραφή). Ο προαναφερθείς κανονισμός περιλαμβάνεται στη βιβλιογραφία του παρόντος δοκιμίου.

αυτών των εκδικάσεων πως θα εκδικάζονταν μελλοντικές υποθέσεις.

Η ΑΙ ουσιαστικά λοιπόν ακολουθεί, έστω εμμέσως, *οδηγίες*. Της θέτουμε εμείς τι θεωρούμε γενικά δίκαιο. Άρα τις θέτουμε τους νόμους που έχουμε θεσπίσει. Στη συνέχεια της λέμε τι παράγοντες της πραγματικότητας αντιστοιχούν στους νόμους, για να μπορεί να τους προσμετρήσει. Και τέλος, η ΑΙ απλά εκτελεί εκδικάσεις βάσει αυτών των οδηγιών με απίστευτη αποτελεσματικότητα και αποδοτικότητα. Εκεί είναι το ταβάνι της.

Η συσχέτιση της ΑΙ-δικαστή με την Αριστοτελική Δικαιοσύνη

Έχοντας καταλάβει ποιο είναι το έργο της ΑΙ-δικαστή, ας δούμε διεξοδικά γιατί δεν μπορεί να ταυτιστεί καθόλου με τη Αριστοτελική Δικαιοσύνη. Για να το πετύχουμε αυτό, θα συσχετίσουμε την ΑΙ-δικαστή με τα τρία καίρια σημεία της Αριστοτελικής Δικαιοσύνης που προανέφερα σε προηγούμενες παραγράφους μέσω μερικών ρεαλιστικών παραδειγμάτων.

Πρώτα ξεκαθαρίσαμε ότι το νομικό δίκαιο μπορεί μόνο συμπωματικά να αντιπροσωπεύει το πραγματικό Δίκαιο. Είναι εύκολο να αναλογιστούμε δεκάδες ιστορικά παραδείγματα τυραννιών, και να καταλάβουμε πως, πράγματι, σε τέτοιες τις περιπτώσεις, ένα ΑΙ-δικαστής θα ήταν απλά ένα πανίσχυρο εργαλείο καταπίεσης στα χέρια των τυράννων. Φανταστείτε για παράδειγμα τον Κιμ Γιονγκ Ουν της Βόρειας Κορέας να είχε στη διάθεσή του μια τέτοια ΑΙ-δικαστή που, εντελώς αυτοματοποιημένα, και μέσω της αστραπιαίας και απίστευτα ακριβούς ανάλυσης στοιχείων και δεδομένων, να μπορεί καταδικάσει ανθρώπους που εναντιώνονται στους νόμους της Βόρειας Κορέας. Κι όμως, ως προς τους τυραννικούς νόμους που λογικά θα έχουν θεσπιστεί σε τέτοιες τυραννίες, η ΑΙ-δικαστής θα ανταποκρίνεται με φοβερή επιτυχία. Κι όλοι οι παραβάτες αυτών των νόμων θα καταδικάζονται από μια απολύτως *νομικά* δίκαιη ΑΙ-δικαστή, ακόμα κι όταν οι νόμοι είναι *πραγματικά* Άδικοι. Μπορούμε όμως να αποδεχτούμε ένα τέτοιο σύστημα ως παράδειγμα πραγματικής Δικαιοσύνης; Αμφιβάλλω.

Μέσω της παραπάνω διαπίστωσης, γρήγορα μπορούμε να δούμε πως μια ΑΙ-δικαστής δεν μπορεί να ταυτιστεί ούτε με το δεύτερο καίριο σημείο της Αριστοτελικής Δικαιοσύνης που υπέδειξα νωρίτερα. Όταν κάποιος ακολουθεί απλά και ρητά το νομικό δίκαιο, ακόμα κι όταν τυχαίνει αυτό το νομικό δίκαιο να αντικατοπτρίζει το πραγματικά Δίκαιο, *δεν* καθίσταται πραγματικά Δίκαιος. Η ΑΙ-δικαστής όμως ξεκάθαρα κάνει ακριβώς και μόνο αυτό: Μιμείται τη Δίκαιη συμπεριφορά, ακολουθώντας ρητά και

ανελαστικά το νομικό δίκαιο. Παρ' όλα αυτά, εύκολα κάποιος μπορεί να υποστηρίξει πως μια AI-δικαστής που έχει παραμετροποιηθεί με πραγματικά Δίκαιους νόμους (άρα νόμους θεσπισμένους από Αριστοτελικά Ενάρετους ανθρώπους) μπορεί κάλλιστα να αντικαταστήσει ανθρώπους δικαστές, κι έτσι να δημιουργηθεί ένα αυτοματοποιημένο και αποδοτικότατο δικαστικό σύστημα. Οι εκδικάσεις της AI-δικαστή θα ήταν, έστω και μόνο φαινομενικά, πραγματικά Δίκαιες, ως έμμεσα αποτελέσματα πράξεων των Ενάρετων.

Ή μήπως όχι;

Η AI-δικαστής δεν μπορεί να είναι Επιεικής εκ Φύσεως

Τώρα λοιπόν είναι που η AI-δικαστής αποδεικνύεται απόλυτα ανεπαρκής ακόμα και στο να προσομοιάσει ένα Αριστοτελικά Δίκαιο δικαστικό σύστημα. Ο λόγος είναι ότι η AI-δικαστής δε μπορεί με κανένα τρόπο να ρυθμιστεί ώστε να δείχνει Επιείκεια (με την ανάγκη για Επιείκεια να αποτελεί το τρίτο και πιο καίριο σημείο της Αριστοτελικής Δικαιοσύνης για το οποίο μίλησα προηγουμένως). Θυμίζω πως ο Αριστοτέλης μίλησε για τη δυνατότητα να εμφανιστούν περιπτώσεις και υποθέσεις μοναδικές στη ζωή, οι οποίες, αν εκδικαστούν με βάση το γενικό νόμο, θα εκδικαστούν Άδικα, ακόμα κι αν ο νόμος είναι γενικά δίκαιος και θεσπισμένος από Ενάρετους ανθρώπους. Αυτές είναι οι περιπτώσεις και υποθέσεις που ξεπερνούν το νόμο (και ίσως τις προβλεπτικές ικανότητες οποιουδήποτε ανθρώπου) και χρειάζονται να διαλευκανθούν κατ' ιδίαν από τους Αριστοτελικά Ενάρετους όταν θα προκύψουν.[15]

Πιστεύω θα ήταν πολύ χρήσιμο να δημιουργήσω ένα φανταστικό παράδειγμα, για να καταλάβουμε και στη πράξη πως μια AI-δικαστής θα αποτύγχανε εκ φύσεως να είναι Επιεικής, άρα, πέραν από ουσιαστικά, θα αποδεικνυόταν *και πρακτικά, και φαινομενικά,* Αριστοτελικά Άδικη.

Ας θεωρήσουμε ότι έχουμε θεσπίσει νόμο στην Ελλάδα που απαγορεύει την κλοπή και τη σωματική βία (μιας και πιστεύω πως τέτοιοι νόμοι φαντάζουν γενικά δίκαιοι σε όλους μας). Ας θεωρήσουμε επίσης ότι έχουμε σχεδιάσει και προγραμματίσει στην εντέλεια μια πανίσχυρη AI-δικαστή

[15] Την συγκεκριμένη αδυναμία της AI να επιδείξει Αριστοτελικά ορισμένη επιείκεια έχει συστηματικά υποδείξει ο Νίκος Α. Παρασκευόπουλος, καθηγητής νομικής του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης, ειδικά στο βιβλίο του *Ο επιεικής αλγόριθμος – Από την Αριστοτελική σκέψη στην τεχνητή νοημοσύνη* (2024).

που ακολουθεί τους νόμους και εκδικάζει υποθέσεις με 99.999% επιτυχία.

Ας φανταστούμε τώρα πως κάποιος άνδρας της φυλής των Ρομά εισέρχεται σε ένα πολυκατάστημα. Τον βλέπουμε να μιλάει με τον υπάλληλο στο ταμείο. Ο υπάλληλος φαίνεται πολύ γρήγορα να θυμώνει με τον άνδρα, και με έντονες και απειλητικές κινήσεις να του δείχνει την πόρτα της εξόδου… και ξαφνικά, ο άνδρας επιτίθεται στον υπάλληλο, τον χτυπάει και τον αφήνει αναίσθητο στο πάτωμα. Μεμιάς ο άνδρας πηδάει πίσω από το ταμείο, ανοίγει ένα ψυγείο με αφεψήματα, αρπάζει ένα χυμό πορτοκάλι, και τρέπεται σε φυγή.

Προφανώς, βάσει της παραστατικότατης εικόνας που μόλις περιέγραψα, θα φωνάξουμε πως ο άνδρας καθαρά αδίκησε: Αδίκησε πρώτα τον υπάλληλο ασκώντας πάνω του σωματική βία, και μετά τον ιδιοκτήτη του πολυκαταστήματος κλέβοντας το προϊόν του. Αντίστοιχα με μας, η AI-δικαστής θα αναλύσει τις πράξεις αυτού του άνδρα (που έτυχε να βιντεοσκοπηθούν από τις κάμερες ασφαλείας του πολυκαταστήματος) και πολύ γρήγορα και απλά θα τον καταδικάσει για κλοπή και άσκηση σωματικής βίας. Ο άνδρας τιμωρείται, και τώρα θα εκτίσει την ποινή που έχει ορίσει ο νόμος (που κατά πάσα πιθανότητα θα είναι μια μακρά ποινή φυλάκισης). Ο νόμος εφαρμόστηκε με επιτυχία.

Τι θα γίνει όμως όταν ανακαλύψουμε και άλλα χαρακτηριστικά μοναδικά σε αυτήν την (φανταστική πάντα) υπόθεση, τα οποία δεν προβλέπει ο νόμος; Πως θα επηρεαστούμε αν ανακαλύψουμε λοιπόν ότι αυτός ο άνθρωπος, λόγω άτυχων συγκυριών, δεν είχε πιει το οτιδήποτε για τρεις μέρες; Αν μαθαίναμε επίσης, ότι έξω από πολυκατάστημα τον περίμενε η πεντάχρονη κόρη του, η οποία επίσης δεν είχε πιει τίποτα τρεις μέρες; Κι αν τέλος, ανακαλύπταμε πως αυτός ο πατέρας είναι η μόνη οικογένεια που έχει απομείνει στον κόσμο για αυτό το άμοιρο πεντάχρονο κορίτσι;

Ο πατέρας λοιπόν, μπαίνοντας στο πολυκατάστημα, εξήγησε στον υπάλληλο την κατάσταση στην οποία αυτός κι η κόρη του βρισκόταν. Πιστεύω πως θα συμφωνούσαμε στο ότι, πατέρας και κόρη βρίσκονταν ξεκάθαρα σε μια εξαθλιωμένη και απελπιστική κατάσταση. Μα για δικούς του λόγους, ο υπάλληλος αρνήθηκε με ασυνήθιστα έντονο ύφος να του προσφέρει το οτιδήποτε για να τους βοηθήσει να σβήσουν τη δίψα τους. Μπορούμε λοιπόν να υποθέσουμε πως ο πατέρας αντέδρασε υπό την επήρεια της απελπισίας και του άγχους του για την υγεία της κόρης του.

Μήπως λοιπόν είναι πρέπων, σε αυτή τη μια και μοναδική υπόθεση, να πάρουμε υπόψιν μας όλους τους δευτερεύοντες λόγους που μόλις φανταστήκαμε, και να δείξουμε Επιείκεια απέναντι στις (πράγματι) Άδικες πράξεις αυτού του άνδρα? Μήπως ο γενικά δίκαιος νόμος θα δημιουργούσε

περαιτέρω Αδικία εις βάρος αυτού του πατέρα και της πεντάχρονης κόρης του αν εφαρμοστεί κατά γράμμα?

Ένα είναι σίγουρο: Η ΑΙ-δικαστής θα παρέμενε απόλυτα καθηλωμένη στα πλαίσιο του νομικού δικαίου και θα καταδίκαζε τον άνδρα χωρίς προστριβές. Διότι μόνο ένας Ενάρετος άνθρωπος θα μπορούσε να θέσει τέτοιες ερωτήσεις στον εαυτό του.

Εν Κατακλείδι

Έχοντας αναλύσει τι ήταν η Αρετή και η Δικαιοσύνη για τον Αριστοτέλη, καθώς και το πως λειτουργεί μια ΑΙ-δικαστής, μπορούμε πλέον να υποθέσουμε με αρκετή σιγουριά για το ποια θα ήταν η άποψη του Αριστοτέλη για μια ΑΙ-δικαστή. Πιστεύω πως, για τον Σταγειρίτη, η ΑΙ γενικότερα θα χαρακτηριζόταν στην ουσία της ως απλά ένα ακόμα δημιούργημα του ανθρώπου: Η ΑΙ είναι ένα αποτέλεσμα ανθρώπινης πράξης, και μιας πράξης προφανώς με απίστευτη επίπτωση πάνω στο σύνολο της ανθρωπότητας, καθώς και απίστευτη επίπτωση στη διατήρηση ή διαστρέβλωση της αρμονίας και της ισότητας στο εσωτερικό της ανθρώπινης κοινωνίας. Αυτό σημαίνει ότι, στην ουσία της, η δημιουργία της ΑΙ θα είναι το αποτέλεσμα μιας Δίκαιης ή Άδικης πράξης, ακριβώς όπως είναι το αποτέλεσμα μιας Δίκαιης ή Άδικης πράξης η θέσπιση νόμων (αλλά και, ουσιαστικά, σχεδόν κάθε ανθρώπινη δημιουργία).

Όμως, όπως υποστηρίζει ο Αριστοτέλης, όλοι οι άνθρωποι έχουν τη δυνατότητα μέσα τους να γίνουν Ενάρετοι. Κι όλοι οι άνθρωποι, είτε είναι υπάλληλοι πολυκαταστημάτων, στρατιώτες, δικαστές ή οτιδήποτε άλλο, έχουν πάντα την δυνατότητα μέσα τους να γίνουν Επιεικείς. Κι ακόμη κι αν όλη τους τη ζωή εφαρμόζουν το νόμο κατά γράμμα (είτε ο νόμος είναι γενικά Δίκαιος ή Άδικος) αυτή η δυνατότητα τους δεν χάνεται πότε. Πάντα η ερώτηση θα πλανάται σε κάποιο επίπεδο της σκέψης τους για το αν αυτό που κρίνουν και πράττουν είναι Δίκαιο ή όχι.

Η ΑΙ-δικαστής θα είναι κατά πάσα πιθανότητα η καλύτερη και πιο αποτελεσματική μορφή εφαρμογής του νομικού δικαίου. Όμως, ελλοχεύει κίνδυνος με διπλή φύση αν αντικαταστήσουμε τους ανθρώπους δικαστές με μια ΑΙ-δικαστή. Πρώτον, να μη παρανοήσουμε τις δυνατότητες της και πιστέψουμε ότι γνωρίζει η ίδια η ΑΙ τι είναι Δίκαιο καλύτερα από τον άνθρωπο. Και δεύτερον, μέσα στην αυταρέσκεια μας για το δημιούργημα μας, να μην εναποθέσουμε την ευθύνη της κρίσης μας ολότελα πάνω της, πνίγοντας τελικά την δυνατότητα μας να γίνουμε οι ίδιοι Ενάρετοι και

Δίκαιοι. Διότι έτσι η δυνατότητα για Αριστοτελική Δικαιοσύνη θα χαθεί οριστικά, καθώς Επιείκεια μέσω AI δε μπορεί να υπάρξει. Πιστεύω πλέον πως είναι κατανοητό γιατί.

Πηγές

Αριστοτέλης. 2006a. *Ηθικά Νικομάχεια Ά–Δ*. Μετάφραση Δημήτριος Λυπουρλής. Θεσσαλονίκη: Εκδόσεις Ζήτρος.

Αριστοτέλης. 2006b. *Ηθικά Νικομάχεια Ε–Κ*. Μετάφραση Δημήτριος Λυπουρλής. Θεσσαλονίκη: Εκδόσεις Ζήτρος.

Bringsjord, Selmer και Naveen Sundar Govindarajulu. 2018. "Artificial Intelligence." *Stanford Encyclopedia of Philosophy*. Stanford University. https://plato.stanford.edu/entries/artificial-intelligence/.

Chen, Stephen. 2021. "Chinese Scientists Develop AI Prosecutor That Can Press Its Own Charges." *South China Morning Post*, 26 Δεκεμβρίου. https://www.scmp.com/news/china/science/article/3160997/chinese-scientists-develop-ai-prosecutor-can-press-its-own.

Γούναρης, Άλκης, και Γιώργος Κωστελέτος. 2024. «Γράφοντας τον αλγόριθμο του Καλού: Η Τεχνητή Νοημοσύνη ως μηχανή απόδοσης Δικαιοσύνης.» *Ηθική. Περιοδικό Φιλοσοφίας* (19). https://doi.org/10.12681/ethiki.39654.

Dennett, Daniel C. 1987. "Cognitive Wheels: The Frame Problem of AI." *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*, 129–151. Norwood, NJ: Ablex Publishing Corporation.

European Parliament and Council (2024). Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). *Official Journal of the European Union*, L series, no. 2024/1689 (12 July 2024). http://data.europa.eu/eli/reg/2024/1689/oj

Newman, Jack. 2021. "China Develops AI Prosecutor That Can Press Charges with 97% Accuracy." *Daily Mail*, 27 Δεκεμβρίου. https://www.dailymail.co.uk/news/article-10346933/China-develops-AI-prosecutor-press-charges-97-accuracy.html.

Παρασκευόπουλος, Α. Νίκος. 2024. *Ο επιεικής αλγόριθμος: Από την Αριστοτελική σκέψη στην τεχνητή νοημοσύνη*. Θεσσαλονίκη: Ίδρυμα Τριανταφυλλίδη.

Rocha, Cinara, and João Alvaro Carvalho. 2022. "Artificial Intelligence in the Judiciary: Uses and Threats." *CEUR Workshop Proceedings*, EGOV-CeDEM-ePart 2022, Linköping University, Sweden, September 6–8. http://ceur-ws.org/Vol-3399/paper17.pdf.

Samson, Carl. 2021. "China Develops AI 'Prosecutor' That Can Charge Citizens with Crimes with '97% Accuracy'." *NextShark*, 27 Δεκεμβρίου. https://nextshark.com/china-ai-prosecutor-97-percent-accuracy.

Searle, John R. 1980. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3 (3): 417–424. https://doi.org/10.1017/s0140525x00005756.

Shanahan, Murray. 2016. "The Frame Problem." *The Stanford Encyclopedia of Philosophy*. Stanford University. https://plato.stanford.edu/entries/frame-problem/.

The Korea Times. 2021. "More than 97% Accuracy: Chinese Scientists Develop AI Prosecutor." *The Korea Times*, 26 Δεκεμβρίου. https://www.koreatimes.co.kr/world/20211226/more-than-97-accuracy-chinese-scientists-develop-ai-prosecutor.

Περίληψη

Σε αυτό το δοκίμιο θα εξετάσω κατά πόσο η Τεχνητή Νοημοσύνη μπορεί να είναι Αριστοτελικά Δίκαιη, και κατά πόσο θα μπορούσε να λειτουργήσει ως Αριστοτελικά Δίκαιη αν αντικαθιστούσε τους ανθρώπους δικαστές στη κοινωνία μας. Αρχικά, θα αναλύσω τι ακριβώς είναι η Δικαιοσύνη στο πλαίσιο της Ηθικής του Αριστοτέλη, και θα υποδείξω πως η Δικαιοσύνη σχετίζεται απόλυτα με την ανθρώπινη Αρετή και πράξη. Επακολούθως, θα ερμηνεύσω την Αριστοτελική Δικαιοσύνη με σκοπό την επαγωγή τριών πορισμάτων που θα έχουν καίρια σημασία για τα τελικά συμπεράσματα της εξέτασής μου. Στη συνέχεια, θα αναφερθώ στους δυο θεμελιώδεις τύπους Τεχνητής Νοημοσύνης που αναγνωρίζονται στη φιλοσοφική κοινότητα, για να ξεδιαλύνουμε πως ακριβώς νοείται η Τεχνητή Νοημοσύνη φιλοσοφικά. Έπειτα θα εξηγήσω πως ακριβώς θα λειτουργούσε η Τεχνητή

Νοημοσύνη, και ποιος θα ήταν ο στόχος αυτής της λειτουργίας, αν έμελλε να πάρει το ρόλο δικαστή. Τέλος, θα συσχετίσω τη λειτουργία και το στόχο της Τεχνητής Νοημοσύνης με την Αριστοτελική Δικαιοσύνη (και συγκεκριμένα με τα τρία προαναφερθέντα πορίσματα) ώστε να αποδείξω πως η Τεχνητή Νοημοσύνη, εκ φύσεως, δε θα μπορούσε ποτέ να είναι Αριστοτελικά Δίκαιη.

Λέξεις Κλειδιά: Αρετή, Αριστοτέλης, Δικαιοσύνη, Δικαστής, Επιείκεια, Νομικό Δίκαιο, Τεχνητή Νοημοσύνη.

Keywords: Virtue, Aristotle, Justice, Judge, Equity, Legal Justice, Artificial Intelligence.

Αλέξανδρος Νούνεσης
Τμήμα Φιλοσοφίας, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών
Email: alexander.nounesis@gmail.com
ORCiD: https://orcid.org/0009-0003-7298-7836