

Ρυθμίζοντας τη Χρήση της Τεχνητής Νοημοσύνης στη Δημόσια Υγεία με στόχο την Πρόληψη των Έμφυλων και Φυλετικών Διακρίσεων

Ένα Ρυθμιστικό Μοντέλο για την Χρήση της
Τεχνητής Νοημοσύνης στη Βιοϊατρική,
Ευθυγραμμισμένο με τη Χάρτα Θεμελιωδών
Δικαιωμάτων

Policy document



Co-funded by
the European Union

Εισαγωγή

Η τεχνητή νοημοσύνη (TN) ενσωματώνεται όλο και περισσότερο στα συστήματα υγειονομικής περίθαλψης σε ολόκληρη την Ευρώπη, υποστηρίζοντας τη διάγνωση, τη λήψη κλινικών αποφάσεων, τον σχεδιασμό θεραπείας και την πρόβλεψη κινδύνου για τους/τις ασθενείς. Αυτές οι τεχνολογίες προσφέρουν σημαντικές ευκαιρίες για τη βελτίωση της αποτελεσματικότητας της υγειονομικής περίθαλψης. Ωστόσο, ολοένα και περισσότερα στοιχεία δείχνουν ότι **τα συστήματα τεχνητής νοημοσύνης (Artificial Intelligence/AI) μπορούν να αναπαράγουν ή να εντείνουν τις υπάρχουσες ανισότητες στον χώρο της υγείας.**¹

Αρκετοί μηχανισμοί μπορούν να οδηγήσουν σε μεροληπτικά αποτελέσματα που αφορούν στη λήψη αποφάσεων οι οποίες υποστηρίζονται από συστήματα TN στον τομέα της υγειονομικής περίθαλψης. Αυτοί περιλαμβάνουν ανισορροπίες στα δεδομένα που χρησιμοποιούνται για την εκπαίδευση των συστημάτων, στα σφάλματα μέτρησης σε κλινικές μεταβλητές, στους ακατάλληλους δείκτες αντικατάστασης (proxy indicators) καθώς και στις διαφοροποιήσεις που εντοπίζονται στα περιβάλλοντα της ανάπτυξης τους και των συνθηκών εφαρμογής τους στον πραγματικό κόσμο.² Προκαταλήψεις μπορεί επίσης να προκύψουν όταν τα συστήματα TN εισάγονται στις εργασίες (σωματικές και πνευματικές) που εκτελούν οι επαγγελματίες υγείας για την παροχή φροντίδας στις/τους ασθενείς (κλινικές ροές εργασίας), όταν δεν υποστηρίζεται από επαρκή εκπαίδευση αυτών ή όταν δεν υπάρχουν ισχυροί μηχανισμοί ανθρώπινης εποπτείας και λογοδοσίας για την παρακολούθηση της απόδοσης των συστημάτων μεταξύ διαφορετικών πληθυσμών ασθενών.³

Επιπλέον, η εμπειρία μέσα από την έρευνα έχει δείξει ότι **τέτοιες προκαταλήψεις μπορεί να έχουν σοβαρές συνέπειες σε κλινικά πλαίσια.** Μελέτες έχουν αναφέρει επαναλαμβανόμενες ανισότητες στη διάγνωση και τη θεραπεία σημαντικών παθήσεων. Για παράδειγμα, οι γυναίκες και οι ασθενείς που ανήκουν σε φυλετικές/εθνοτικές μειονότητες με καρδιαγγειακά νοσήματα, συχνά λαμβάνουν καθυστερημένη διάγνωση ή λιγότερο εντατική φροντίδα σε

¹ Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y., [Η τεχνητή νοημοσύνη στον τομέα της υγειονομικής περίθαλψης: παρελθόν, παρόν και μέλλον](#), *Stroke and Vascular Neurology* 2(4): 230–243, 2017.

² Vokinger, K. N., Feuerriegel, S., & Kesselheim, A. S., [Μείωση της μεροληψίας στη μηχανική μάθηση στον τομέα της ιατρικής](#), *Communications Medicine* 1, 25, 2021.

³ Panesar, S., Kliot, M., Parrish, R., Fernandez-Miranda, J., Cagle, Y., & Britz, G.W., [Οι υποσχέσεις και οι κίνδυνοι της τεχνητής νοημοσύνης στη νευροχειρουργική](#), *Neurosurgery* 87(1): 33–44, 2020.

σύγκριση με τους άνδρες και λευκούς ασθενείς,⁴ ενώ οι φυλετικές και εθνοτικές διαφορές επηρεάζουν τη φροντίδα και τα τελικά αποτελέσματα ως προς την ασθένεια του διαβήτη.⁵ Παράλληλα, ψυχικές παθήσεις όπως η κατάθλιψη υποδιαγιγνώσκονται και υποθεραπεύονται τόσο στις γυναίκες όσο και στους μειονοτικούς πληθυσμούς.⁶

Ένα καλά τεκμηριωμένο παράδειγμα, αποτελεί μια μελέτη στο *Journal of Medical Internet Research*, η οποία διαπίστωσε διαφοροποιήσεις στη διαγνωστική επίδοση- βάσει φύλου σε αλγόριθμους μηχανικής μάθησης που χρησιμοποιούνται για την πρόβλεψη καρδιακών παθήσεων. Πιο συγκεκριμένα, οι γυναίκες ασθενείς υποεκπροσωπούσαν συστηματικά στα σύνολα δεδομένων που χρησιμοποιούνταν για την εκπαίδευση των συστημάτων, ενώ τα μοντέλα εμφάνιζαν υψηλότερα ποσοστά ψευδώς αρνητικών αποτελεσμάτων για τις γυναίκες, αποδεικνύοντας **πώς η ύπαρξη μεροληπτικών δεδομένων μπορεί να συμβάλλει αρνητικά στη διαγνωστική ακρίβεια.**⁷ Όσον αφορά τη φυλετική ισότητα, μια άλλη μελέτη ανέδειξε ότι μοντέλα TN που εκπαιδεύτηκαν για την εκτίμηση της σοβαρότητας της κατάθλιψης μέσω ανάλυσης γλωσσικών δεδομένων από μέσα κοινωνικής δικτύωσης, παρουσίασαν υψηλή ακρίβεια για τους λευκούς συμμετέχοντες, αλλά σημαντικά χαμηλότερη για τους μαύρους συμμετέχοντες. Το εύρημα αυτό αναδεικνύει διαφοροποιήσεις στην απόδοση του συστήματος, οι οποίες φαίνεται να συνδέονται με διαφορές στα πρότυπα γλωσσικής έκφρασης που δεν αποτυπώθηκαν επαρκώς κατά την εκπαίδευση των μοντέλων.⁸

Η αντιμετώπιση των κινδύνων που σχετίζονται με τις προκαταλήψεις **απαιτεί ρυθμιστικά μοντέλα που ενσωματώνουν τις διεθνείς και ευρωπαϊκές ηθικές και νομικές αρχές σε συγκεκριμένους μηχανισμούς διακυβέρνησης**, τους οποίους μπορούν να υιοθετήσουν και να εφαρμόσουν τα ιδρύματα υγειονομικής περίθαλψης. Εντός της Ευρωπαϊκής Ένωσης (ΕΕ), το ρυθμιστικό πλαίσιο έχει εξελιχθεί ραγδαία με την υιοθέτηση του Κανονισμού για την Τεχνητή Νοημοσύνη (AI Act), η οποία θεσπίζει υποχρεώσεις για συστήματα TN υψηλού κινδύνου και

⁴ Balla, S., Gomez, S.E., & Rodriguez, F., [Ανισότητες στην καρδιαγγειακή περίθαλψη και στα αποτελέσματα για τις γυναίκες που προέρχονται από φυλετικές/εθνοτικές μειονότητες](#), *Current Treatment Options in Cardiovascular Medicine* 22(12):75, 2020.

⁵ Young, C. & Myers, A. K., [Φυλετικές και εθνοτικές ανισότητες στην κλινική φροντίδα και διαχείριση του διαβήτη: Μια αφηγηματική ανασκόπηση](#), *Endocrine Practice* 29(4): 295–300, 2023.

⁶ Sorkin, D.H., Ngo-Metzger, Q., Billimek, J., August, K.J., Greenfield, S., & Kaplan, S.H., [Υποδιαγνωσμένη και ανεπαρκώς θεραπευόμενη κατάθλιψη σε ασθενείς διαφορετικής φυλετικής/εθνοτικής καταγωγής με διαβήτη τύπου 2](#), *Diabetes Care* 34(3): 598–600, 2011.

⁷ Straw, I., Rees, G., & Nachev, P., [Διαφορές στην απόδοση ανάλογα με το φύλο σε αλγόριθμους μηχανικής μάθησης για την πρόβλεψη καρδιακών παθήσεων: Μια διερευνητική μελέτη](#), *Journal of Medical Internet Research* 26: e46936, 2024.

⁸ Rai, S., Stadel, E. C., Giorgi, S., Francisco, A., Ungar, L. H., Curtis, B., & Guntuku, S. C., [Οι βασικοί γλωσσικοί δείκτες της κατάθλιψης στα μέσα κοινωνικής δικτύωσης εξαρτώνται από τη φυλή](#), *Proceedings of the National Academy of Sciences* 121(14): e2319837121, 2024.

για όσους τα θέτουν σε λειτουργία, όπως για παράδειγμα τα νοσοκομεία.⁹ Ωστόσο, η πρακτική εφαρμογή αυτών των νομικών απαιτήσεων και διαδικασιών εντός των ιδρυμάτων υγείας, παραμένει μια σημαντική πρόκληση.

Το παρόν Έγγραφο Πολιτικής (*Policy Document*) εισηγείται ένα **ρυθμιστικό μοντέλο για τη διακυβέρνηση βιοϊατρικών συστημάτων ΤΝ, το οποίο ευθυγραμμίζεται με το ευρωπαϊκό κεκτημένο (EU acquis)**, συμπεριλαμβανομένων των αρχών που κατοχυρώνονται στον Χάρτη Θεμελιωδών Δικαιωμάτων της ΕΕ, ενώ είναι σχεδιασμένο για τα δημόσια συστήματα υγειονομικής περίθαλψης και τις οργανώσεις της κοινωνίας των πολιτών (ΚΤΠ). Το προτεινόμενο μοντέλο βασίζεται σε ευρωπαϊκά ρυθμιστικά μέσα, συμπεριλαμβανομένης του [Κανονισμού για την ΤΝ \(AI Act\)](#), του [Κανονισμού για τα Ιατροτεχνολογικά Προϊόντα \(Medical Device Regulation/MDR\)](#), του [Κανονισμού για τα In Vitro Διαγνωστικά Ιατροτεχνολογικά Προϊόντα \(In Vitro Diagnostic Regulation /IVDR\)](#), του [Γενικού Κανονισμού για την Προστασία Δεδομένων \(GDPR\)](#) και του αναδυόμενου Ευρωπαϊκού Χώρου Δεδομένων Υγείας (*European Health Data Space/EHDS*). Μετασχηματίζει έτσι τις αρχές των θεμελιωδών δικαιωμάτων, όπως την ανθρώπινη αξιοπρέπεια, την ισότητα, τη μη διάκριση και το δικαίωμα στην υγειονομική περίθαλψη, σε πρακτικές διασφαλίσεις προστασίας και ελέγχου που εφαρμόζονται σε όλα τα στάδια του κύκλου ζωής των συστημάτων ΤΝ. Αναφορές σε επιλεγμένα παραδείγματα δικαιοδοσιών εκτός ΕΕ παρουσιάζονται στο παρόν έγγραφο μόνο όπου προσφέρουν χρήσιμες γνώσεις, λειτουργώντας συμπληρωματικά προς το ευρωπαϊκό νομικό πλαίσιο καθώς και το πλαίσιο που βασίζεται στα δικαιώματα, τα οποία στηρίζουν το μοντέλο.

Επιπλέον, το προτεινόμενο ρυθμιστικό μοντέλο αξιοποιεί πρακτικές από καθιερωμένες ρυθμιστικές προσεγγίσεις που συναντώνται σε άλλους τομείς υψηλού κινδύνου, όπως η διαχείριση της ασφάλειας στην αεροπορία, η διακυβέρνηση κινδύνου χρηματοοικονομικών μοντέλων, η φαρμακο-επαγρύπνηση για τη ρύθμιση του τομέα των φαρμάκων και οι αξιολογήσεις αλγοριθμικού αντικτύπου στη δημόσια διοίκηση. Αυτές οι προσεγγίσεις παρέχουν πολύτιμες γνώσεις για τη δημιουργία πλαισίων διακυβέρνησης, ικανών να αντιμετωπίσουν περίπλοκους κινδύνους σε όλα τα στάδια του κύκλου ζωής των βιοϊατρικών συστημάτων ΤΝ. Στο παρόν πλαίσιο αυτές οι προσεγγίσεις έχουν προσαρμοστεί ώστε να λαμβάνουν υπόψη και να ενσωματώνουν τις διαστάσεις της ισότητας των φύλων και της φυλετικής ισότητας σε όλα τα στάδια του κύκλου ζωής των συστημάτων ΤΝ.

⁹ Regulation (EU) 2024/1689 of the European Parliament and of the Council, [Θέσπιση εναρμονισμένων κανόνων για την τεχνητή νοημοσύνη \(AI Act\)](#), PE/24/2024/REV/1, Official Journal of the European Union, 13 June 2024.

Το προτεινόμενο μοντέλο παρέχει ένα **ολοκληρωμένο και εφαρμόσιμο πλαίσιο διακυβέρνησης** για τα ιδρύματα υγειονομικής περίθαλψης και τις οργανώσεις της κοινωνίας των πολιτών, καλύπτοντας όλα τα στάδια του κύκλου ζωής των βιοϊατρικών συστημάτων ΤΝ, **από τις προμήθειες και την αξιολόγηση έως την ανάπτυξη, τη συνεχή παρακολούθηση και τη λογοδοσία**. Μέσω της ενσωμάτωσης των αρχών προστασίας των θεμελιωδών δικαιωμάτων στις επιχειρησιακές διαδικασίες, το πλαίσιο στοχεύει στον περιορισμό των κινδύνων που σχετίζονται με έμφυλες και φυλετικές προκαταλήψεις στη λήψη αποφάσεων οι οποίες υποστηρίζονται από συστήματα ΤΝ στον χώρο της υγείας. Παράλληλα στόχος του πλαισίου είναι να συμβάλλει σε πιο δίκαια αποτελέσματα στην υγειονομική περίθαλψη σε ολόκληρη την ΕΕ, ενώ **προσφέρει μια σειρά από πρακτικές πολιτικές, εργαλεία και πρότυπα τεκμηρίωσης (*artifacts*), τα οποία τα νοσοκομεία και οι οργανώσεις της κοινωνίας των πολιτών (ΚΤΠ) μπορούν να προσαρμόσουν σε διαφορετικά εθνικά πλαίσια**, διευκολύνοντας την αποτελεσματική εφαρμογή του μοντέλου στην πράξη.

Το Ευρωπαϊκό Νομικό Πλαίσιο για την ΤΝ στην Υγειονομική Περίθαλψη

Η ΕΕ έχει θεσπίσει ένα πολυεπίπεδο ρυθμιστικό περιβάλλον για τα βιοϊατρικά συστήματα ΤΝ, το οποίο στηρίζεται σε τέσσερις αλληλένδετες νομικές / κανονιστικές διαστάσεις: τον Κανονισμό για την ΤΝ (AI Act), τον Κανονισμό για τα ιατροτεχνολογικά προϊόντα με τις σχετικές κατευθυντήριες γραμμές (MDR), τον Κανονισμό για την προστασία των προσωπικών δεδομένων (GDPR) και τη διακυβέρνηση των δεδομένων που σχετίζονται με τον χώρο της υγείας (EHDS).

Ο Κανονισμός για την Τεχνητή Νοημοσύνη (AI Act) θεσπίζει ένα ρυθμιστικό πλαίσιο για τα συστήματα ΤΝ, το οποίο βασίζεται στην προσέγγιση της αξιολόγησης των κινδύνων. Πολλές εφαρμογές στον τομέα της υγειονομικής περίθαλψης, συμπεριλαμβανομένων των εργαλείων υποστήριξης κλινικών αποφάσεων και των διαγνωστικών αλγορίθμων, ταξινομούνται στην κατηγορία των συστημάτων ΤΝ υψηλού κινδύνου, και ως εκ τούτου οφείλουν να συμμορφώνονται με αυστηρές προδιαγραφές που αφορούν τη διαχείριση κινδύνων, τη διακυβέρνηση των δεδομένων, τη διαφάνεια, την ανθρώπινη εποπτεία και την ακρίβεια των αποτελεσμάτων που παράγουν. Οι ισχύουσες προδιαγραφές περιλαμβάνουν επιπλέον τον εντοπισμό και τον περιορισμό των κινδύνων εμφάνισης μεροληπτικών αποτελεσμάτων μέσω αναλυτικής αξιολόγησης των επιμέρους πληθυσμιακών ομάδων, λαμβάνοντας υπόψη το βιολογικό και κοινωνικό φύλο (sex and gender), την φυλετική προέλευση και άλλα κοινωνικά χαρακτηριστικά ως οριζόντιες μεταβλητές. **Τα ιδρύματα υγειονομικής**

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them. Project code: 101215009 – AEQUITAS – CERV-2024-CHAR-LITI

περίθαλψης που θέτουν σε λειτουργία συστήματα TN υπέχουν επίσης ειδικές υποχρεώσεις βάσει της Πράξης για την TN (AI Act). Σε αυτές περιλαμβάνονται η διασφάλιση ότι τα συστήματα χρησιμοποιούνται σύμφωνα με τις οδηγίες του παρόχου, παρακολουθείται η επίδοσή τους, τηρούνται αρχεία λειτουργίας και αναφέρονται σοβαρά περιστατικά στις αρμόδιες αρχές, ιδίως όταν τα αυτά ενδέχεται να επηρεάζουν δυσανάλογα διαφορετικές ομάδες ασθενών.

Όσον αφορά την εφαρμογή του Κανονισμού για την TN (AI Act), το 2025 τέθηκε ήδη σε ισχύ η απαγόρευση ορισμένων πρακτικών, ενώ άρχισαν να εφαρμόζονται ορισμένοι κανόνες εποπτείας και διακυβέρνησης.¹⁰ Το επόμενο σημαντικό ορόσημο είναι ο Αύγουστος του 2026, όπου θα τεθεί σε ισχύ η πλειονότητα των υποχρεώσεων που προβλέπονται από τον Κανονισμό (συμπεριλαμβανομένων των συστημάτων υψηλού κινδύνου του Παραρτήματος III και των υποχρεώσεων διαφάνειας του Άρθρου 50). Στην Ελλάδα ο Κανονισμός (ΕΕ) 2024/1689 του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου της 13ης Ιουνίου 2024 (Κανονισμός για την τεχνητή νοημοσύνη) είναι σε [δημόσια ηλεκτρονική διαβούλευση](#) μέχρι 6 Ιουλίου 2026.

Τα συστήματα TN υψηλού κινδύνου που ενσωματώνονται σε ρυθμιζόμενα προϊόντα (κατηγορία στην οποία περιλαμβάνονται πολλά συστήματα TN για ιατροτεχνολογικά προϊόντα) προβλέπεται να υπαχθούν στις σχετικές υποχρεώσεις από τον Αύγουστο του 2027.

Για τα νοσοκομεία, οι σημαντικότερες πρακτικά διατάξεις του Κανονισμού για την TN είναι οι εξής:

- **Απαραίτητες προδιαγραφές για συστήματα TN υψηλού κινδύνου:** Οι πάροχοι συστημάτων TN οφείλουν να εφαρμόζουν ένα σύστημα διαχείρισης κινδύνων σε όλα τα στάδια του κύκλου ζωής του συστήματος (άρθρο 9), καθώς και πρακτικές διακυβέρνησης δεδομένων οι οποίες περιλαμβάνουν: σαφείς απαιτήσεις για τον εντοπισμό και τον μετριασμό των προκαταλήψεων (άρθρο 10), συμπεριλαμβανομένης της αξιολόγησης της αντιπροσωπευτικότητας των δεδομένων και των πιθανών αποκλίσεων/ανισοτήτων μεταξύ διαφορετικών πληθυσμιακών ομάδων λαμβάνοντας υπόψη τη χρήση ανάλυσης με βάση το βιολογικό και κοινωνικό φύλο καθώς και την εθνοτική ή φυλετική προέλευση· διασφαλίσεις διαφάνειας και παροχή σχετικών πληροφοριών σε όσους είναι υπεύθυνοι για την ανάπτυξη και λειτουργία των συστημάτων (Άρθρο 13)· μηχανισμούς ανθρώπινης εποπτείας που αποσκοπούν στην πρόληψη της υπερβολικής εξάρτησης από αυτοματοποιημένες αποφάσεις (automation bias) και οι οποίοι θα επιτρέπουν την παρέμβαση ή τη διακοπή

¹⁰ European Commission, [Χρονοδιάγραμμα για την εφαρμογή του κανονισμού της ΕΕ για την τεχνητή νοημοσύνη](#), Brussels: European Commission, 2024.

της λειτουργίας του συστήματος (άρθρο 14) όταν χρειάζεται· ελέγχους ακρίβειας, ανθεκτικότητας και κυβερνοασφάλειας, με ιδιαίτερη προσοχή σε βρόχους ανατροφοδότησης και επιθέσεις ειδικά σχεδιασμένες για συστήματα ΤΝ (άρθρο 15), οι οποίες, εάν δεν αντιμετωπιστούν, ενδέχεται να ενισχύσουν υφιστάμενες ανισότητες.

- **Υποχρεώσεις όσων είναι υπεύθυνοι για την ανάπτυξη και λειτουργία των συστημάτων (deployers):** Τα νοσοκομεία και άλλοι φορείς που χρησιμοποιούν συστήματα ΤΝ υψηλού κινδύνου υποχρεούνται να λειτουργούν τα συστήματα σύμφωνα με τις οδηγίες του παρόχου (άρθρο 26)· να ορίζουν κατάλληλα καταρτισμένο ανθρώπινο προσωπικό για την εποπτεία τους (άρθρο 26)· να διασφαλίζουν ότι τα δεδομένα εισόδου είναι αντιπροσωπευτικά και πλήρη, όταν ο διαχειριστής ελέγχει την εισαγωγή των δεδομένων (άρθρο 26), λαμβάνοντας, όπου είναι εφικτό, υπόψη τη δημογραφική ποικιλομορφία των κλινικών πληθυσμών· να παρακολουθούν τη λειτουργία και την απόδοση των συστημάτων (άρθρο 26)· να διατηρούν αρχεία λειτουργίας που φυλάσσονται για τουλάχιστον έξι μήνες (άρθρο 26)· να αναφέρουν εγκαίρως σοβαρά περιστατικά στους παρόχους και στις αρμόδιες αρχές (άρθρα 26 και 73), ιδίως όταν υπάρχει κίνδυνος άνισων κλινικών αποτελεσμάτων.
- **Εκτίμηση Αντικτύπου / Επιπτώσεων στα Θεμελιώδη Δικαιώματα (Fundamental Rights Impact Assessment /FRIA):** Πριν από την ανάπτυξη ορισμένων συστημάτων ΤΝ υψηλού κινδύνου, οι φορείς δημοσίου δικαίου, καθώς και ιδιωτικοί οργανισμοί που παρέχουν δημόσιες υπηρεσίες, οφείλουν να διενεργούν Εκτίμηση Αντικτύπου στα Θεμελιώδη Δικαιώματα, περιγράφοντας τις διαδικασίες χρήσης του συστήματος, τις ομάδες που ενδέχεται να επηρεαστούν, τους σχετικούς κινδύνους καθώς και τους μηχανισμούς περιορισμού αυτών, τα μέτρα που έχουν ληφθεί για την ανθρώπινη εποπτεία και τους μηχανισμούς υποβολής παραπόνων καθώς και τις κοινοποιήσεις αυτών στην αρμόδια αρχή εποπτείας της αγοράς (άρθρο 28). Η απαίτηση αυτή είναι ιδιαίτερα σημαντική για τα δημόσια νοσοκομεία πολλών κρατών-μελών, καθώς παρέχει έναν μηχανισμό έγκαιρου εντοπισμού και αντιμετώπισης κινδύνων, διακρίσεων και άνισων επιπτώσεων, μέσω της ανάλυσης των αποτελεσμάτων λαμβάνοντας υπόψη το φύλο και τη φυλή, πριν από την ανάπτυξη/εφαρμογή των συστημάτων ΤΝ.

Πέραν του Κανονισμού για την ΤΝ, τα συστήματα **που χρησιμοποιούνται σε ιατρικά πλαίσια υπόκεινται, σε συγκεκριμένους κανονισμούς και κατευθυντήριες οδηγίες**, συμπεριλαμβανομένου του Κανονισμού για τα Ιατροτεχνολογικά Προϊόντα (MDR)¹¹ και του Κανονισμού για τα In Vitro

¹¹ European Union, [Κανονισμός \(ΕΕ\) 2017/745 σχετικά με τα ιατροτεχνολογικά προϊόντα \(MDR\)](#), Official Journal of the European Union, L117, Brussels: EU, 2017.

Διαγνωστικά Ιατροτεχνολογικά Προϊόντα (IVDR),¹² οι οποίοι ρυθμίζουν την ασφάλεια και την αποτελεσματικότητα των ιατρικών τεχνολογιών. Οι κανονισμοί αυτοί επιβάλλουν τη κλινική αξιολόγηση και τη συνεχή παρακολούθηση της απόδοσης των λογισμικών ιατροτεχνολογικών προϊόντων μετά τη διάθεσή τους στην αγορά, συμπεριλαμβανομένων των εργαλείων που ενσωματώνουν τεχνολογίες ΤΝ. Ιδιαίτερη έμφαση, δίνεται στην αξιολόγηση της αποτελεσματικότητας και της ισότιμης λειτουργίας τους μεταξύ διαφορετικών ομάδων ασθενών, όπου υπάρχουν διαθέσιμα επαρκή και κατάλληλα δεδομένα. Επιπλέον, οι Κατευθυντήριες Οδηγίες του Συντονιστικού Οργάνου Ιατροτεχνολογικών Προϊόντων (*Medical Device Coordination Group/ MDCG*) σχετικά με τον Χαρακτηρισμό και την Ταξινόμηση του Λογισμικού βάσει του MDR/IVDR, παρέχουν οδηγίες για τον τρόπο χαρακτηρισμού και ταξινόμησης του λογισμικού.¹³ Εκεί υπογραμμίζεται ότι η ταξινόμηση εξαρτάται από τον προβλεπόμενο σκοπό χρήσης και όχι από τη θέση του λογισμικού (π.χ. αν λειτουργεί στο cloud ή είναι ενσωματωμένο σε συσκευή), ενώ περιλαμβάνονται παραδείγματα από πληροφοριακά συστήματα νοσοκομείων και λογισμικά υποστήριξης κλινικών αποφάσεων, τονίζοντας ιδιαίτερα την ανάγκη αξιολόγησης πιθανών διαφοροποιημένων επιπτώσεων σε διαφορετικούς πληθυσμούς ασθενών. Οι Κατευθυντήριες οδηγίες του MDCG για την κλινική και λειτουργική αξιολόγηση του λογισμικού ιατροτεχνολογικών προϊόντων ενισχύουν περαιτέρω την απαίτηση η τεκμηρίωση να είναι ανάλογη με τον κλινικό ρόλο του λογισμικού, συμπεριλαμβάνοντας στοιχεία που αποδεικνύουν ότι η αποτελεσματικότητα των συστημάτων ΤΝ παραμένει συνεπής μεταξύ πληθυσμιακών ομάδων που διαφοροποιούνται με βάση το βιολογικό ή κοινωνικό φύλο και την εθνοτική ή φυλετική προέλευση.¹⁴

Η σχέση μεταξύ του Κανονισμού για την ΤΝ (AI Act) και της υφιστάμενης νομοθεσίας για τα ιατροτεχνολογικά προϊόντα αποσαφηνίζεται στις Κατευθυντήριες Οδηγίες για την Αλληλεπίδραση μεταξύ του Κανονισμού για την ΤΝ και του MDR/IVDR.¹⁵ Το παρόν έγγραφο εξηγεί ότι ο Κανονισμός για την ΤΝ εφαρμόζεται παράλληλα με τους δύο αυτούς κανονισμούς. Στην πράξη, αυτό σημαίνει ότι **πολλά συστήματα ιατρικής ΤΝ πρέπει να συμμορφώνονται ταυτόχρονα και με τα δύο κανονιστικά πλαίσια**. Οι κατευθυντήριες οδηγίες

¹² European Union, [Κανονισμός \(ΕΕ\) 2017/746 σχετικά με τα ιατροτεχνολογικά προϊόντα in vitro \(IVDR\)](#), Official Journal of the European Union, L117, Brussels: EU, 2017.

¹³ Medical Device Coordination Group, [Κατευθυντήριες οδηγίες σχετικά με τον χαρακτηρισμό και την ταξινόμηση του λογισμικού στον Κανονισμό \(ΕΕ\) 2017/745 \(MDR\) και στον Κανονισμό \(ΕΕ\) 2017/746 \(IVDR\)](#), Brussels: MDCG, 2019.

¹⁴ Medical Device Coordination Group, [Κατευθυντήριες οδηγίες για την κλινική αξιολόγηση \(MDR\) και την αξιολόγηση επιδόσεων \(IVDR\) λογισμικού σε ιατρικές συσκευές](#), Brussels: MDCG, 2020.

¹⁵ Medical Device Coordination Group & European AI Board, [Κατευθυντήριες οδηγίες σχετικά με την αλληλεπίδραση μεταξύ των κανονισμών MDR/IVDR και του νόμου για την τεχνητή νοημοσύνη](#), Brussels: European Commission, 2025.

διευκρινίζουν επίσης πότε ένα ιατρικό σύστημα TN θεωρείται υψηλού κινδύνου σύμφωνα με τον Κανονισμό για την TN, κάτι που συνήθως συμβαίνει όταν το ίδιο το σύστημα είναι ιατροτεχνολογικό προϊόν ή στοιχείο ασφαλείας ενός τέτοιου προϊόντος και επομένως υπόκειται σε αξιολόγηση συμμόρφωσης από ανεξάρτητο διαπιστευμένο φορέα. Για τις διαδικασίες προμηθειών των νοσοκομείων, αυτό έχει σημαντικές πρακτικές συνέπειες. Τα νοσοκομεία θα πρέπει να απαιτούν από τους προμηθευτές να τεκμηριώνουν το κανονιστικό καθεστώς των συστημάτων τους, συμπεριλαμβανομένης της ταξινόμησης MDR/IVDR, την αναφορά διαπιστευμένου φορέα για αξιολόγηση όπου αυτό απαιτείται, τη συμμόρφωση με τις απαιτήσεις υψηλού κινδύνου του Κανονισμού για την TN, καθώς και την τεκμηρίωση της απόδοσης ανά πληθυσμιακή ομάδα και των μέτρων περιορισμού των προκαταλήψεων. Οι συμβάσεις προμηθειών θα πρέπει επίσης να απαιτούν από τους αναδόχους να παρέχουν την απαραίτητη τεκμηρίωση, πληροφορίες διαφάνειας και μέτρα εποπτείας που χρειάζονται όσοι είναι υπεύθυνοι για την ανάπτυξη και λειτουργία των συστημάτων, σύμφωνα με τον Κανονισμό για την TN, συμπεριλαμβανομένων προβλέψεων για τη διασφάλιση δίκαιων αποτελεσμάτων για όλες/όλους τις/τους ασθενείς.

Επιπλέον του Κανονισμού για την TN (AI Act), αρκετά ευρωπαϊκά εργαλεία έχουν σημασία για τη διακυβέρνηση των συστημάτων TN στον τομέα της υγιονομικής περίθαλψης. Ο Γενικός Κανονισμός για την Προστασία Δεδομένων (GDPR) εξακολουθεί να αποτελεί έναν βασικό περιορισμό για τη βιοϊατρική και τα συστήματα TN, καθώς τα δεδομένα υγείας αποτελούν ειδική κατηγορία προσωπικών δεδομένων και η ανάπτυξη συστημάτων σε νοσοκομειακό περιβάλλον συχνά απαιτεί τη διενέργεια Εκτίμησης Αντικτύπου σχετικά με την Προστασία Δεδομένων (Data Protection Impact Assessments/ DPIA).

Ο Κανονισμός για την TN (AI Act) (άρθρα 26(9) και 27(4)) συνδέει ρητά τις πρακτικές των υπεύθυνων ανάπτυξης και λειτουργίας των συστημάτων με τις υποχρεώσεις για Εκτίμηση Αντικτύπου σχετικά με την Προστασία Δεδομένων (DPIA) και θεωρεί την **Εκτίμηση Αντικτύπου / Επιπτώσεων στα Θεμελιώδη Δικαιώματα** (FRIA) συμπληρωματική της πρώτης, η οποία μπορεί να περιλαμβάνει την αξιολόγηση διαφορετικών επιπτώσεων μεταξύ πληθυσμιακών ομάδων σε αναλύσεις που λαμβάνουν υπόψη το φύλο και τη φυλή. Επιπλέον, ο Ευρωπαϊκός Χώρος Δεδομένων Υγείας (*European Health Data Space/EHDS*) αποσκοπεί στη διευκόλυνση της ασφαλούς πρόσβασης σε δεδομένα υγείας τόσο για την παροχή υπηρεσιών υγείας όσο και για την έρευνα, διατηρώντας παράλληλα ισχυρές εγγυήσεις για την προστασία των δεδομένων.¹⁶ Ο EHDS αναμένεται να διαδραματίσει σημαντικό ρόλο στη δημιουργία πιο αντιπροσωπευτικών συνόλων δεδομένων για την ανάπτυξη συστημάτων TN στη βιοϊατρική, ενισχύοντας τα με αξιολογήσεις που θα έχουν στο επίκεντρο την

¹⁶ European Union, [Κανονισμός για τον Ευρωπαϊκό Χώρο Δεδομένων Υγείας \(EHDS\)](#), Brussels: European Union, 2024.

ισότητα, ενώ θα μειώνουν τον κίνδυνο της υποεκπροσώπησης ομάδων ασθενών με συγκεκριμένα δημογραφικά χαρακτηριστικά.¹⁷

Τέλος, όσον αφορά την κυβερνοασφάλεια, ο Οργανισμός της Ευρωπαϊκής Ένωσης για την Κυβερνοασφάλεια (*European Union Agency for Cybersecurity*) έχει αναπτύξει κατευθυντήριες οδηγίες σχετιζόμενες με πρακτικές κυβερνοασφάλειας στην ΤΝ. Οι οδηγίες αυτές υποστηρίζουν ότι υφιστάμενα οργανωτικά πρότυπα (όπως πρότυπα διαχείρισης ασφάλειας και ποιότητας ISO) μπορούν να προσαρμοστούν στις ανάγκες των συστημάτων ΤΝ.¹⁸ Οι κατευθυντήριες οδηγίες του Συντονιστικού Οργάνου Ιατροτεχνολογικών Προϊόντων (MDCG) για την κυβερνοασφάλεια των ιατροτεχνολογικών προϊόντων καθορίζουν το πλαίσιο συμμόρφωσης των κατασκευαστών με τις σχετικές απαιτήσεις των Κανονισμών MDR και IVDR, δίνοντας έμφαση στην εφαρμογή διαδικασιών διαχείρισης κινδύνου και στην υιοθέτηση βασικών μέτρων κυβερνοασφάλειας.¹⁹ Η ισχυρή διακυβέρνηση της κυβερνοασφάλειας συνιστά θεμελιώδη παράγοντα για τη διασφάλιση της δίκαιης και αξιόπιστης λειτουργίας των συστημάτων ΤΝ. Μέσω της προστασίας της ακεραιότητας των δεδομένων, της ασφαλούς πρόσβασης σε αυτά και της αξιοπιστίας των συστημάτων, μειώνεται ο κίνδυνος ακούσιας διαιώνισης ή ενίσχυσης υφιστάμενων έμφυλων, φυλετικών ή εθνοτικών ανισοτήτων.

Συνολικά, τα παραπάνω νομικά εργαλεία δημιουργούν ένα σύνθετο ρυθμιστικό περιβάλλον εντός του οποίου καλούνται να λειτουργήσουν τα συστήματα ΤΝ στον τομέα της υγειονομικής περίθαλψης. Το προτεινόμενο μοντέλο ρύθμισης συστημάτων ΤΝ για τα ευρωπαϊκά δημόσια νοσοκομεία έχει σχεδιαστεί με βάση αυτές τις τέσσερις βασικές νομικές «διαστάσεις» που διέπουν την πρακτική εφαρμογή της ΤΝ στον τομέα της υγείας. Αντίστοιχα, το μοντέλο προβλέπει μηχανισμούς για την παρακολούθηση και τον περιορισμό έμφυλων και φυλετικών ανισοτήτων σε όλα τα στάδια του κύκλου ζωής των συστημάτων Τεχνητής Νοημοσύνης.

Ευρωπαϊκά και Διεθνή Πλαίσια Διακυβέρνησης που εφαρμόζονται στην Υπεύθυνη Τεχνητή Νοημοσύνη στην Υγεία

¹⁷ European Commission, [Ευρωπαϊκός Χώρος Δεδομένων Υγείας \(EHDS\)](#), Brussels: European Commission, 2025.

¹⁸ European Union Agency for Cybersecurity (ENISA), [Κυβερνοασφάλεια της ΤΝ και τυποποίηση](#), Athens: ENISA, 2023.

¹⁹ Medical Device Coordination Group, [MDCG 2019 16 Αναθ. 1 – Κατευθυντήριες οδηγίες για την κυβερνοασφάλεια των ιατρικών συσκευών](#), Brussels: European Commission, 2020.

Είναι σημαντικό να σημειωθεί ότι τα ευρωπαϊκά ρυθμιστικά και κανονιστικά πλαίσια που αναλύθηκαν παραπάνω αποτελούν το πρωταρχικό νομικό θεμέλιο για τη διακυβέρνηση της τεχνητής νοημοσύνης στον τομέα της βιοϊατρικής εντός της Ευρωπαϊκής Ένωσης. Πέραν του ευρωπαϊκού ρυθμιστικού περιβάλλοντος για τη ΤΝ στη βιοϊατρική, διάφορα μη δεσμευτικά διεθνή εργαλεία παρέχουν κατευθύνσεις για τη διαχείριση των κινδύνων που συνδέονται με την ΤΝ. **Τα πρόσθετα πλαίσια διακυβέρνησης που εξετάζονται στην παρούσα ενότητα παρουσιάζονται για να αναδείξουν πρακτικές και μεθόδους εφαρμογής που μπορούν να αξιοποιηθούν στο ευρωπαϊκό πλαίσιο, λειτουργώντας πάντοτε συμπληρωματικά στις ευρωπαϊκές νομικές απαιτήσεις και στις αρχές των θεμελιωδών δικαιωμάτων.** Η αξιοποίησή τους περιορίζεται σε περιπτώσεις που προσφέρουν χρήσιμα παραδείγματα από άλλες δικαιοδοσίες, στις οποίες έχουν ήδη εφαρμοστεί αντίστοιχοι μηχανισμοί διακυβέρνησης, ιδίως στους τομείς των μοντέλων διαχείρισης κινδύνου, του αλγοριθμικού ελέγχου και της αξιολόγησης θεσμικών επιπτώσεων.

Για παράδειγμα, το Πλαίσιο Διαχείρισης Κινδύνων Τεχνητής Νοημοσύνης (*AI Risk Management Framework/AI RMF*) του Εθνικού Ινστιτούτου Προτύπων και Τεχνολογίας των Ηνωμένων Πολιτειών (*National Institute of Standards and Technology/ NIST*) προσφέρει μια ολοκληρωμένη προσέγγιση για τη διαχείριση κινδύνων σε συστήματα ΤΝ, που εφαρμόζεται σε όλα τα στάδια του κύκλου ζωής αυτών και οργανώνεται γύρω από τη σωστή διακυβέρνηση, τη χαρτογράφηση, την αξιολόγηση και τη διαχείριση των κινδύνων.²⁰ Το πλαίσιο αυτό αναδεικνύει ρητά τη δικαιοσύνη και την αντιμετώπιση επιβλαβών μορφών που προκύπτουν από προκαταλήψεις ως κεντρικές διαστάσεις για την αξιόπιστη χρήση της ΤΝ. Επιπλέον, στο πλαίσιο της διαδικασίας εντοπισμού και αξιολόγησης κινδύνων, ενθαρρύνονται οι οργανισμοί να αναγνωρίσουν και να αναλύσουν πιθανές διαφορετικές επιπτώσεις μεταξύ επιμέρους πληθυσμιακών ομάδων, λαμβάνοντας υπόψη παράγοντες όπως το βιολογικό και κοινωνικό φύλο, καθώς και τη φυλετική και εθνοτική προέλευση. Σκοπός είναι η ενσωμάτωση μηχανισμών διασφάλισης της αξιοπιστίας του συστήματος και ελέγχου των κινδύνων σε όλα τα στάδια σχεδιασμού, ανάπτυξης και παρακολούθησής του, συνδυάζοντας τεχνικές αξιολογήσεις με οργανωτικές διαδικασίες λογοδοσίας. Για τα νοσοκομεία, ιδιαίτερη αξία του Πλαισίου Διαχείρισης Κινδύνων Τεχνητής Νοημοσύνης (*AI RMF*), αποτελεί η δυνατότητα μετασχηματισμού του σε συγκεκριμένες πρακτικές, **όπως ο σαφής καθορισμός αρμοδιοτήτων για την εποπτεία ζητημάτων ισότητας, η συστηματική αξιολόγηση της απόδοσης των συστημάτων ΤΝ σε διαφορετικές πληθυσμιακές ομάδες και οι τεκμηριωμένες στρατηγικές περιορισμού κινδύνων,** στο πλαίσιο μιας ευρύτερης προσέγγισης που βασίζεται στα δικαιώματα.

²⁰ National Institute of Standards and Technology, [Πλαίσιο Διαχείρισης Κινδύνων Τεχνητής Νοημοσύνης \(AI RMF 1.0\)](#), Gaithersburg, MD: NIST, 2023.

Σε επίπεδο διαχείρισης, το πρότυπο ISO/IEC 42001 του Διεθνούς Οργανισμού Τυποποίησης (ISO), καθορίζει προϋποθέσεις για ένα οργανωτικό σύστημα διαχείρισης συστημάτων ΤΝ, εστιάζοντας σε πολιτικές, στόχους, διαδικασίες και μηχανισμούς συνεχούς βελτίωσης με στόχο την υπεύθυνη χρήση αυτών.²¹ Αν και δεν καθορίζει συγκεκριμένους τύπους βλαβών, υποστηρίζει την ενσωμάτωση των αρχών της δικαιοσύνης και της μη διάκρισης στις πολιτικές που αφορούν την οργάνωση και τους μηχανισμούς ελέγχου κινδύνων, επιτρέποντας στα ιδρύματα να αντιμετωπίζουν συντονισμένα πιθανές προκαταλήψεις ως προς το φύλο και τη φυλή. Συμπληρωματικά, το πρότυπο ISO/IEC 23894 παρέχει καθοδήγηση για την ενσωμάτωση της διαχείρισης κινδύνων στην ανάπτυξη και εφαρμογή συστημάτων ΤΝ, συμπεριλαμβανομένου του εντοπισμού και της αντιμετώπισής τους, οι οποίοι σχετίζονται με ακούσιες ή δυσανάλογες επιπτώσεις σε διαφορετικές πληθυσμιακές ομάδες.²² Ειδικότερα στον τομέα της κλινικής ΤΝ, το πρότυπο ISO 14971 παρέχει ένα δομημένο πλαίσιο αρχών διαχείρισης κινδύνου, το οποίο χρησιμοποιείται ευρέως για τη ρύθμιση των ιατροτεχνολογικών προϊόντων. Οι αρχές αυτές περιλαμβάνουν τη συστηματική αναγνώριση κινδύνων, την αξιολόγηση και τον έλεγχό τους, καθώς και τη συνεχή παρακολούθηση σε όλα τα στάδια του κύκλου ζωής του προϊόντος.²³ Οι διαδικασίες αυτές μπορούν να επεκταθούν ώστε να περιλαμβάνουν την αξιολόγηση διαφοροποιήσεων στην αποτελεσματικότητα των συστημάτων - ως μορφή δυνητικής βλάβης- ανάμεσα σε διαφορετικές πληθυσμιακές ομάδες. Τα παραπάνω πρότυπα είναι ιδιαίτερα χρήσιμα, καθώς **προτείνουν μια δομημένη αρχιτεκτονική ενός συστήματος διαχείρισης που μπορεί να υιοθετηθεί από τα νοσοκομεία**. Παράλληλα, παρέχουν μια κοινή γλώσσα για την ενσωμάτωση απαραίτητων προϋποθέσεων που σχετίζονται με ζητήματα ισότητας, όπως η αξιολόγηση απόδοσης αποτελεσμάτων ανά πληθυσμιακή ομάδα και η αντιμετώπιση των προκαταλήψεων στις διαδικασίες προμηθειών και στις συμβάσεις με προμηθευτές.

Επίσης, τα πλαίσια ηθικής διακυβέρνησης παρέχουν μια σημαντική κανονιστική καθοδήγηση. Οι Κατευθυντήριες Οδηγίες της Ευρωπαϊκής Επιτροπής για Αξιόπιστη Τεχνητή Νοημοσύνη (*Ethics Guidelines for Trustworthy AI*) αποτελούν ένα χρήσιμο, αν και μη δεσμευτικό, εργαλείο.²⁴ Οι οδηγίες αυτές ορίζουν επτά (7) απαραίτητες προϋποθέσεις για υπεύθυνα συστήματα ΤΝ, μεταξύ των οποίων περιλαμβάνονται η ποικιλομορφία, η μη διάκριση και η δικαιοσύνη, οι οποίες

²¹ International Organization for Standardization, [ISO/IEC 42001:2023 – Σύστημα διαχείρισης τεχνητής νοημοσύνης](#), Geneva: ISO, 2023.

²² International Organization for Standardization, [ISO/IEC 23894:2023 – Διαχείριση κινδύνων τεχνητής νοημοσύνης](#), Geneva: ISO, 2023

²³ International Organization for Standardization, [ISO 14971:2019 – Ιατρικές συσκευές: Εφαρμογή της διαχείρισης κινδύνων στις ιατρικές συσκευές](#), Geneva: ISO, 2019

²⁴ European Commission, [Κατευθυντήριες γραμμές δεοντολογίας για μια αξιόπιστη τεχνητή νοημοσύνη](#), Brussels: European Commission, 2019.

ουσιαστικά υπογραμμίζουν την ανάγκη πρόληψης, αντιμετώπισης και αποφυγής μορφών προκαταλήψεων καθώς και την αξιολόγηση των επιπτώσεων αυτών σε πληθυσμούς με διαφορετικά δημογραφικά χαρακτηριστικά, όπως το φύλο και η φυλετική προέλευση. Η Λίστα Αξιολόγησης για Αξιόπιστη Τεχνητή Νοημοσύνη (*Assessment List for Trustworthy Artificial Intelligence/ ALTAI*), η οποία αναπτύχθηκε από την Ομάδα Εμπειρογνομόνων της Ευρωπαϊκής Επιτροπής για την ΤΝ, μετατρέπει τις παραπάνω οδηγίες σε ένα εργαλείο αυτοαξιολόγησης.²⁵ Στο πλαίσιο ενός μοντέλου υπεύθυνης διακυβέρνησης της ΤΝ στον τομέα της υγειονομικής περίθαλψης, η ιδιαίτερη αξία του ALTAI έγκειται στη μορφή του ως **μια λίστα ερωτημάτων, που μπορεί να ενσωματωθεί σε πρακτικές ελέγχου για διαδικασίες προμηθειών και ανάπτυξης συστημάτων**. Με τον τρόπο αυτό, ενθαρρύνονται οι εμπλεκόμενοι φορείς να εξετάζουν ζητήματα όπως η αντιπροσωπευτικότητα των δεδομένων εκπαίδευσης των συστημάτων ΤΝ, ο κίνδυνος εμφάνισης άνισων κλινικών αποτελεσμάτων και η επάρκεια των μέτρων προστασίας έναντι επιπτώσεων που προκύπτουν από διακρίσεις, ενσωματώνοντας έτσι τη διάσταση ζητημάτων ισότητας στις καθημερινές αποφάσεις κλινικών, τεχνικών στελεχών και διοικήσεων νοσοκομείων.

Από την οπτική των ανθρωπίνων δικαιωμάτων, η Σύμβαση-πλαίσιο του Συμβουλίου της Ευρώπης σχετικά με την τεχνητή νοημοσύνη και τα ανθρώπινα δικαιώματα, τη δημοκρατία και το κράτος δικαίου ([Framework Convention on AI](#)) αποτελεί ένα ιδιαίτερα σημαντικό νομικό εργαλείο, καθώς στηρίζεται στις αρχές της ανθρώπινης αξιοπρέπειας, της διαφάνειας, της λογοδοσίας, της ισότητας, της μη διάκρισης και της προστασίας της ιδιωτικότητας, παρέχοντας ένα σαφές νομικό πλαίσιο για την αντιμετώπιση των επιπτώσεων που προκύπτουν από τις διακρίσεις, συμπεριλαμβανομένων εκείνων που σχετίζονται με έμφυλες ή φυλετικές μορφές προκατάληψης στα συστήματα ΤΝ.²⁶ Αντίστοιχα, οι Αρχές για την Τεχνητή Νοημοσύνη του Οργανισμού Οικονομικής Συνεργασίας και Ανάπτυξης,²⁷ και η Σύσταση της UNESCO για τη Δεοντολογία της Τεχνητής Νοημοσύνης,²⁸ παρέχουν διεθνή κανονιστική καθοδήγηση, ενθαρρύνοντας τον εντοπισμό και τον μετριασμό δυσανάλογων επιπτώσεων μεταξύ διαφορετικών πληθυσμιακών ομάδων και ενισχύοντας την πεποίθηση ότι τα συστήματα ΤΝ δεν πρέπει να αναπαράγουν ή να ενισχύουν δομικές κοινωνικές ανισότητες. Ειδικότερα, όσον αφορά την ηθική διακυβέρνηση της ΤΝ στον τομέα της υγείας, ο Παγκόσμιος Οργανισμός Υγείας (ΠΟΥ) έχει διατυπώσει έξι (6) θεμελιώδεις αρχές: την αυτονομία, την ευημερία, την επεξηγησιμότητα, την ευθύνη, την ισότητα και

²⁵ European Commission, [Κατάλογος Αξιολόγησης για Αξιόπιστη Τεχνητή Νοημοσύνη \(ALTAI\) – Αυτοαξιολόγηση](#), Brussels: European Commission, 2020.

²⁶ Council of Europe, [Σύμβαση-Πλαίσιο για την Τεχνητή Νοημοσύνη \(CAI\)](#), Strasbourg: Council of Europe, 2023.

²⁷ Organisation for Economic Co-operation and Development, [OECD Αρχές του ΟΟΣΑ για την Τεχνητή Νοημοσύνη](#), Paris: OECD, 2019

²⁸ UNESCO, [Σύσταση σχετικά με τη δεοντολογία της τεχνητής νοημοσύνης](#), Paris: UNESCO, 2021.

την ανταποκρισιμότητα.²⁹ Η αρχή της ισότητας, επιβάλλει την αξιολόγηση πιθανών διαφοροποιήσεων σχετικά με την πρόσβαση, την απόδοση και τα αποτελέσματα των συστημάτων ΤΝ μεταξύ διαφορετικών ομάδων ασθενών, λαμβάνοντας υπόψη παράγοντες όπως το βιολογικό φύλο, το κοινωνικό φύλο και τη φυλετική προέλευση.

Τα διεθνή αυτά πλαίσια υπεύθυνης διακυβέρνησης συνδέονται άμεσα με τις απαιτήσεις του Χάρτη Θεμελιωδών Δικαιωμάτων της Ευρωπαϊκής Ένωσης σχετικά με την ανθρώπινη αξιοπρέπεια, την ισότητα, τη μη διάκριση και το δικαίωμα στην υγεία. Παράλληλα, μοιράζονται μια κοινή έμφαση στη διακυβέρνηση καθ' όλο τον κύκλο ζωής των συστημάτων, αρχές οι οποίες μπορούν να μεταφραστούν άμεσα σε πρακτικούς μηχανισμούς ελέγχου και λογοδοσίας για τα ιδρύματα υγειονομικής περίθαλψης.

Διδάγματα από διατομεακά ρυθμιστικά μοντέλα

Τομείς με υψηλά επίπεδα τεχνολογικών κινδύνων, όπως η αεροπορία, ο τραπεζικός κλάδος και η φαρμακευτική βιομηχανία, έχουν αναπτύξει εδώ και πολλά χρόνια εξελιγμένα συστήματα διακυβέρνησης για τη διαχείριση τους. Οι τομείς αυτοί προσφέρουν πολύτιμες πρακτικές ρύθμισης, οι οποίες μπορούν να προσαρμοστούν στη διακυβέρνηση των συστημάτων ΤΝ στον χώρο της υγείας και στην αντιμετώπιση συναφών προκλήσεων που σχετίζονται με προκαταλήψεις λόγω φύλου και φυλετικής προέλευσης.

Παρότι η έμφυλη ή φυλετική προκατάληψη δεν αποτελεί τεχνολογικό κίνδυνο με τη στενή έννοια του όρου, αναδύεται ως κοινωνικο-τεχνικός κίνδυνος όταν τα δεδομένα, οι αλγόριθμοι και οι πρακτικές οργάνωσης των συστημάτων, αλληλοεπιδρούν με δομικές ανισότητες στον χώρο της υγειονομικής περίθαλψης. Στο πλαίσιο αυτό, ο όρος «κίνδυνος» αναφέρεται στην εκδήλωση της προκατάληψης σε επίπεδο συστήματος, όπως αυτή αποτυπώνεται στα αποτελέσματα, στις αποφάσεις και στις κλινικές εκβάσεις, παρά στην αρχική αιτία δημιουργίας της. Ο ίδιος ο τομέας της υγειονομικής περίθαλψης διαθέτει εκτενή δεδομένα που καταδεικνύουν την ύπαρξη αντίστοιχων φαινομένων, όπως οι τεκμηριωμένες διαφοροποιήσεις στη διάγνωση, στη θεραπευτική αντιμετώπιση και στην πρόβλεψη κινδύνου μεταξύ πληθυσμιακών ομάδων που διακρίνονται με βάση το φύλο και τη φυλετική / εθνοτική προέλευση. Συνεπώς, τα παραδείγματα από άλλους τομείς λειτουργούν συμπληρωματικά προς τα εμπειρικά δεδομένα του χώρου της υγειονομικής περίθαλψης, καθώς αναδεικνύουν δοκιμασμένους μηχανισμούς διακυβέρνησης για τη διαχείριση σύνθετων συστημάτων υψηλού αντικτύπου, ενώ τα δεδομένα από τον χώρο της υγειονομικής περίθαλψης τεκμηριώνουν την ανάγκη εφαρμογής αντίστοιχων ρυθμιστικών μηχανισμών.

²⁹ World Health Organization, [Ηθική και Διακυβέρνηση της Τεχνητής Νοημοσύνης στον Τομέα της Υγείας](#), Geneva: WHO, 2021.

Ένα σχετικό παράδειγμα αποτελεί το Σύστημα Διαχείρισης Ασφάλειας (*Safety Management System – SMS*) που εφαρμόζεται στην αεροπορία. Οι διεθνείς κανονισμοί πολιτικής αεροπορίας αντιμετωπίζουν την ασφάλεια όχι ως μια μεμονωμένη διαδικασία πιστοποίησης, αλλά ως μια συνεχώς εξελισσόμενη οργανωσιακή πρακτική που βασίζεται στον εντοπισμό κινδύνων και τη διαχείρισή τους, στη διασφάλιση της ασφάλειας μέσω μηχανισμών παρακολούθησης και ελέγχου, στη διαχείριση επιχειρησιακών αλλαγών και στη συνεχή βελτίωση των διαδικασιών.³⁰ Η εφαρμογή αυτής της προσέγγισης στη διακυβέρνηση των συστημάτων ΤΝ στον χώρο της υγειονομικής περίθαλψης **θα μπορούσε να οδηγήσει στη δημιουργία ενός «Συστήματος Διαχείρισης Ασφάλειας ΤΝ» (*AI Safety Management System*)**, το οποίο θα ενσωματώνει παραμέτρους ισότητας, όπως:

- (1) τη δημιουργία μηχανισμού αναφοράς κινδύνων και περιστατικών που ωστόσο δεν οδήγησαν σε βλάβη (*near misses*), μέσω του οποίου οι κλινικοί επαγγελματίες θα μπορούν να καταγράφουν πιθανές περιπτώσεις έμφυλων ή φυλετικών προκαταλήψεων·
- (2) τη διατήρηση επίσημου μητρώου κινδύνων για κάθε σύστημα ΤΝ, στο οποίο θα καταγράφονται οι διαφοροποιήσεις στα αποτελέσματα μεταξύ διαφορετικών πληθυσμιακών ομάδων ασθενών·
- (3) την εφαρμογή διαδικασιών ελέγχου αλλαγών, μέσω των οποίων θα αξιολογείται ο αντίκτυπος των αναβαθμίσεων των μοντέλων, των μεταβολών στα δεδομένα ή των τροποποιήσεων στις κλινικές διαδικασίες και στα αποτελέσματα όσον αφορά την ισότητα και
- (4) τη διενέργεια περιοδικών αξιολογήσεων ασφάλειας με διορθωτικές ενέργειες που αντιμετωπίζουν τόσο τεχνικούς κινδύνους, όσο και εντοπισμένες έμφυλες και φυλετικές ανισότητες. Η προσέγγιση αυτή βρίσκεται σε συμφωνία με τον Κανονισμό για την ΤΝ (*AI Act*) (άρθρα 9 και 72), που αφορά στη διαχείριση κινδύνων σε όλα τα στάδια του κύκλου ζωής των συστημάτων ΤΝ και στην παρακολούθησή τους μετά τη διάθεση στην αγορά.

Μια δεύτερη χρήσιμη πρακτική προέρχεται από τον χρηματοπιστωτικό τομέα, όπου σύνθετα ποσοτικά εργαλεία αντιμετωπίζονται ως «μοντέλα» που απαιτούν αυστηρές διαδικασίες ανάπτυξης, ανεξάρτητης επικύρωσης και εποπτείας της διακυβέρνησης. Στο πλαίσιο της τραπεζικής εποπτείας, αντίστοιχα μοντέλα διαχείρισης κινδύνου έχουν εφαρμοστεί σε διάφορες δικαιοδοσίες, όπως στις Ηνωμένες Πολιτείες (π.χ. οι Εποπτικές Κατευθυντήριες Οδηγίες της Ομοσπονδιακής Τράπεζας των ΗΠΑ σχετικά με το Πρότυπο/ Μοντέλο

³⁰ International Civil Aviation Organization, [Εγχειρίδιο Διαχείρισης Ασφάλειας \(Εγγράφο 9859\)](#), 4th ed., Montréal: ICAO, 2018.

Διαχείρισης Κινδύνου), παρέχοντας ένα παράδειγμα για την εφαρμογή διαδικασιών επικύρωσης, ελέγχου και ανεξάρτητης αξιολόγησης μοντέλων.³¹

Αντίστοιχες πρακτικές μπορούν να εφαρμοστούν σε συστήματα ΤΝ στη βιοϊατρική, ώστε να διασφαλίζεται ότι τα εργαλεία υποστήριξης κλινικών αποφάσεων υποβάλλονται σε ανεξάρτητη επικύρωση (εσωτερικά ή εξωτερικά) πριν από την εισαγωγή τους στην κλινική πρακτική, με ειδική αξιολόγηση της απόδοσής τους σε διαφορετικές πληθυσμιακές ομάδες με διαφορετικά χαρακτηριστικά ως προς το φύλο και την φυλετική ή εθνοτική προέλευση. Τα βασικά τεκμήρια επικύρωσης θα πρέπει να περιλαμβάνουν την περιγραφή του σκοπού χρήσης του συστήματος, των ορίων λειτουργίας του, των γνωστών περιορισμών του, καθώς και της απόδοσής του ανά πληθυσμιακή ομάδα, συμπεριλαμβανομένων τυχόν διαφορών καθώς και μέτρων μετριασμού. Η επανεπικύρωση θα πρέπει να ενεργοποιείται όταν το σύστημα εφαρμόζεται σε νέο οργανισμό υγειονομικής περίθαλψης, σε διαφορετικό πληθυσμό ασθενών, όταν συνδέεται με νέο εξοπλισμό ή πρωτόκολλα, όταν τροποποιείται από τον πάροχο ή παρουσιάζει ενδείξεις υποβάθμισης της απόδοσής του. Οι πρακτικές αυτές συνάδουν με τον Κανονισμό για την Τεχνητή Νοημοσύνη (άρθρα 13, 14 και 60), ο οποίος επιβάλλει τη διεξαγωγή δοκιμών, τη διαφάνεια έναντι των φορέων εφαρμογής και την ανθρώπινη εποπτεία με σκοπό τον εντοπισμό και την πρόληψη της υπερβολικής εξάρτησης από αυτοματοποιημένες αποφάσεις (*automation bias*), διασφαλίζοντας ότι η φροντίδα που υποστηρίζεται από την τεχνητή νοημοσύνη είναι ασφαλής και δίκαιη για όλες τις διαφορετικές πληθυσμιακές ομάδες.

Όσον αφορά στη διακυβέρνηση των συστημάτων μετά την ανάπτυξή τους, το μοντέλο ρύθμισης στον χώρο των φαρμάκων παρέχει μια ιδιαίτερα χρήσιμη πρακτική, αυτή της «φαρμακοεπαγρύπνησης» (*pharmacovigilance*). Στο πλαίσιο αυτό, οι ανεπιθύμητες ενέργειες και τα σήματα ασφάλειας συλλέγονται, αξιολογούνται και αναλύονται συστηματικά για τον εντοπισμό νέων κινδύνων. Συγκεκριμένα, το Μέρος VI των Ορθών Πρακτικών Φαρμακοεπαγρύπνησης (*Good Pharmacovigilance Practices/GVP*) του Ευρωπαϊκού Οργανισμού Φαρμάκων (*European Medicines Agency*) καθορίζει λεπτομερείς απαιτήσεις για τη συλλογή, τη διαχείριση και την υποβολή αναφορών σχετικά με υποψίες ανεπιθύμητων ενεργειών, αντανακλώντας μια ολιστική προσέγγιση στην αναφορά περιστατικών και την παρακολούθησή τους.³² Μια αντίστοιχη προσέγγιση «επαγρύπνησης για την ΤΝ» (*AI-vigilance*) θα μπορούσε να επιτρέψει στα νοσοκομεία την καταγραφή, αξιολόγηση και ανάλυση περιστατικών που σχετίζονται με

³¹ Board of Governors of the Federal Reserve System and Office of the Comptroller of the Currency, [Εποπτικές κατευθυντήριες οδηγίες για τη διαχείριση του κινδύνου μοντέλων \(SR 11 7\)](#), Washington, D.C., 4 April 2011.

³² European Medicines Agency, [Ενότητα VI του GVP: Συλλογή και υποβολή αναφορών για ύποπτες ανεπιθύμητες ενέργειες](#), London: EMA, 2017

κλινικές αποφάσεις έπειτα από τη χρήση συστημάτων TN, με ιδιαίτερη έμφαση στον εντοπισμό πιθανών επιπτώσεων που προκύπτουν από προκαταλήψεις καθώς και ζητημάτων ισότητας.

Ένα τέτοιο σύστημα θα μπορούσε να περιλαμβάνει:

- (1) συστηματική υποβολή εκθέσεων που καταγράφει τυχόν παρατηρούμενες ανισότητες στη φροντίδα ή στα αποτελέσματα ανάλογα με το βιολογικό και το κοινωνικό φύλο ή τις φυλετικές/εθνοτικές διαφορές,
- (2) υποχρεωτική καταγραφή σοβαρών περιστατικών στα οποία οι προκαταλήψεις ενδέχεται να συνέβαλαν στη διαφοροποιημένη μεταχείριση των ασθενών,
- (3) υποχρεώσεις γνωστοποίησης στον πάροχο και στις αρμόδιες αρχές των προβλημάτων που σχετίζονται με ζητήματα ισότητας (ενημέρωση των αρχών εάν οι προκαταλήψεις οδηγούν σε σοβαρή βλάβη, όπως εσφαλμένη διάγνωση ή άνιση μεταχείριση, ενώ οι συνήθεις ανισότητες παραμένουν στο πλαίσιο της εσωτερικής παρακολούθησης της ποιότητας), και
- (4) μηχανισμούς ανίχνευσης τάσεων και ενδείξεων που εντοπίζουν αναδυόμενα μοτίβα προκαταλήψεων σε τμήματα, εγκαταστάσεις ή στο ευρύτερο δίκτυο υγειονομικής περίθαλψης.

Μια τέτοια διαδικασία παρακολούθησης θα επέτρεπε την εφαρμογή προληπτικών μέτρων περιορισμού κινδύνων και τη συνεχή βελτίωση, διασφαλίζοντας ότι η περίθαλψη που υποστηρίζεται από την τεχνητή νοημοσύνη δεν είναι μόνο ασφαλής, αλλά και ισότιμη.

Τέλος, εργαλεία δημόσιας διακυβέρνησης, όπως οι αξιολογήσεις αλγοριθμικού αντικτύπου και οι έλεγχοι για εντοπισμό προκαταλήψεων, προσφέρουν πρακτικές μεθόδους για την αξιολόγηση των κοινωνικών επιπτώσεων των αυτοματοποιημένων συστημάτων λήψης αποφάσεων, συμπεριλαμβανομένων πιθανών ανισοτήτων μεταξύ των διαφορετικών φύλων ή φυλών. Τα εργαλεία αυτά μετατρέπουν τις ηθικές αρχές και τα πρότυπα ανθρωπίνων δικαιωμάτων σε εφαρμόσιμες διαδικασίες, μέσω των οποίων οι οργανισμοί μπορούν να εντοπίζουν και να περιορίζουν πιθανές ανισότητες πριν από την ανάπτυξη των συστημάτων. Για παράδειγμα, η ομοσπονδιακή Αξιολόγηση Αλγοριθμικού Αντικτύπου του Καναδά αξιολογεί τους κινδύνους των συστημάτων TN και τους συνδέει με συγκεκριμένα μέτρα μετριασμού,³³ αντιμετωπίζοντας την ανάπτυξη της TN ως ζήτημα διαχείρισης επιπτώσεων. Αντίστοιχα, η προσέγγιση της πόλης της Νέας Υόρκης για εργαλεία αυτοματοποιημένης λήψης αποφάσεων στην

³³ Treasury Board of Canada Secretariat, [Αξιολόγηση των επιπτώσεων των αλγορίθμων](#), Ottawa: Government of Canada, 2026.

απασχόληση απαιτεί ελέγχους μεροληψίας και υποχρεώσεις δημοσιοποίησης πληροφοριών.³⁴ Σε αυτό το πλαίσιο, τα νοσοκομεία και οι οργανώσεις της κοινωνίας των πολιτών στον χώρο της υγείας **θα μπορούσαν να εφαρμόσουν ένα ολοκληρωμένο «Πακέτο Αξιολόγησης και Ελέγχου Αντικτύπου ΤΝ»**, το οποίο θα συνδυάζει ένα δομημένο ερωτηματολόγιο εστιασμένο στην ισότητα με περιοδικούς ανεξάρτητους ελέγχους δικαιοσύνης, παρέχοντας τεκμηριωμένη εικόνα της απόδοσης των συστημάτων σε διαφορετικές πληθυσμιακές ομάδες και των μέτρων αντιμετώπισης που εφαρμόζονται. Η εφαρμογή ενός τέτοιου πλαισίου θα ενίσχυε τη διαφάνεια, τη δημόσια λογοδοσία και την εμπιστοσύνη των ασθενών στη φροντίδα υγείας που υποστηρίζεται από ΤΝ. Η ενσωμάτωση κανονιστικών μοντέλων από τομείς όπως η ασφάλεια της αεροπορίας, η διαχείριση κινδύνου χρηματοπιστωτικών μοντέλων, η φαρμακοεπαγρύπνηση και η αλγοριθμική διακυβέρνηση, έχει ως στόχο τη σκόπιμη μεταφορά καθιερωμένων αρχών διακυβέρνησης. Οι τομείς αυτοί επιλέχθηκαν επειδή εφαρμόζουν στην πράξη συστηματικές προσεγγίσεις εντοπισμού, παρακολούθησης και μετριασμού κινδύνων σε σύνθετα συστήματα, όπου ενδέχεται να προκύψουν διαφοροποιημένες επιπτώσεις για διαφορετικές ομάδες ανθρώπων. Η σημασία τους έγκειται τόσο στην εννοιολογική τους συνάφεια όσο και στην αποδεδειγμένη εφαρμογή μηχανισμών διακυβέρνησης σε όλα τα στάδια του κύκλου ζωής, υποχρεωτικής αναφοράς περιστατικών και ανεξάρτητης εποπτείας, οι οποίοι μπορούν να προσαρμοστούν για την αντιμετώπιση προκαταλήψεων που σχετίζονται με ζητήματα φύλου και φυλετικής/εθνοτικής προέλευσης στα βιοϊατρικά συστήματα ΤΝ.

³⁴ New York City Department of Consumer and Worker Protection, [Εργαλεία αυτοματοποιημένης λήψης αποφάσεων σχετικά με την απασχόληση](#), New York: City of New York, 2023.

Σχέδιο για ένα κανονιστικό μοντέλο Τεχνητής Νοημοσύνης ευθυγραμμισμένο με τον Χάρτη Θεμελιωδών Δικαιωμάτων για δημόσια νοσοκομεία και οργανώσεις της κοινωνίας των πολιτών

Η παρούσα ενότητα παρουσιάζει την αρχιτεκτονική ενός κανονιστικού μοντέλου που βασίζεται στις αρχές του [Χάρτη Θεμελιωδών Δικαιωμάτων της ΕΕ](#), αλλά διατυπώνεται ως ένα σύνολο πρακτικών διαδικασιών που μπορούν να εφαρμόσουν στην πράξη τα νοσοκομεία και οι οργανώσεις της κοινωνίας των πολιτών. Ο Χάρτης συνιστά το θεμέλιο για τη δημιουργία ενός τέτοιου πλαισίου που αφορά στη διακυβέρνηση των τεχνολογιών υγείας στην Ευρώπη, με ιδιαίτερη έμφαση στις αρχές της ανθρώπινης αξιοπρέπειας, της ισότητας, της απαγόρευσης των διακρίσεων και του δικαιώματος στην υγειονομική περίθαλψη.³⁵ Τα δικαιώματα αυτά αποκτούν ιδιαίτερη σημασία όταν αυτοματοποιημένα συστήματα ΤΝ επηρεάζουν κλινικές αποφάσεις που αφορούν τη φροντίδα των ασθενών, καθώς ενδέχεται να ενισχύουν ανισότητες μεταξύ πληθυσμιακών ομάδων που διαφοροποιούνται με βάση το φύλο και τη φυλετική ή εθνοτική καταγωγή.

Σε θεμελιώδες επίπεδο, το άρθρο 1 του Χάρτη, το οποίο ορίζει ότι η ανθρώπινη αξιοπρέπεια είναι απαραβίαστη, αποτελεί τη βάση για την υιοθέτηση απαραίτητων προϋποθέσεων διακυβέρνησης που διασφαλίζουν ότι τα συστήματα ΤΝ χρησιμοποιούνται με τρόπο που διατηρείται η ανθρώπινη κρίση και ο σεβασμός στην αυτονομία του/της ασθενούς. Στην πράξη, η αρχή αυτή απαιτεί την υιοθέτηση δικλίδων ασφαλείας έναντι της εξάρτησης από αυτοματοποιημένες συστάσεις, τη θέσπιση μηχανισμών αμφισβήτησης αποφάσεων που υποστηρίζονται από συστήματα ΤΝ, καθώς και τη διασφάλιση αποτελεσματικής κλινικής εποπτείας. Τα μέτρα αυτά αποσκοπούν στην αποτροπή της υπερβολικής εξάρτησης από αυτοματοποιημένες αποφάσεις (*automation bias*), δηλαδή της τάση των ανθρώπων που λαμβάνουν αποφάσεις να εμπιστεύονται υπερβολικά τα αλγοριθμικά αποτελέσματα, γεγονός που μπορεί να οδηγήσει στην ενίσχυση υφιστάμενων ανισοτήτων ιδιαίτερα αν οι συστάσεις των συστημάτων ΤΝ συστηματικά κάνουν διακρίσεις σε βάρος συγκεκριμένων πληθυσμιακών ομάδων.

Αντίστοιχα, τα άρθρα 20 και 21, τα οποία κατοχυρώνουν την ισότητα ενώπιον του νόμου και απαγορεύουν τις διακρίσεις, θεμελιώνουν τη νομική βάση για την επιβολή ρητών υποχρεώσεων δικαιοσύνης και τον περιορισμό των προκαταλήψεων στα συστήματα ΤΝ στη βιοϊατρική. Τα νοσοκομεία που

³⁵ European Union, [Χάρτης των Θεμελιωδών Δικαιωμάτων της Ευρωπαϊκής Ένωσης](#), 2012/C 326/02, Strasbourg: EU.

αναπτύσσουν τέτοια συστήματα οφείλουν να εφαρμόζουν διαδικασίες συστηματικής παρακολούθησης της απόδοσής τους σε διαφορετικούς πληθυσμούς ασθενών, λαμβάνοντας υπόψη χαρακτηριστικά όπως το φύλο, τη φυλή ή την εθνοτική καταγωγή, καθώς και τη διαθεματικότητα και τις διασταυρούμενες ανισότητες. Όταν εντοπίζονται ανισότητες στην απόδοση ή στα αποτελέσματα των συστημάτων ΤΝ, θα πρέπει να λαμβάνονται διορθωτικά μέτρα μέσω προσαρμογών του μοντέλου, αλλαγών στις κλινικές διαδικασίες ή εναλλακτικών διαδρομών λήψης αποφάσεων, ώστε η χρήση των συστημάτων ΤΝ να μην παγιώνει διαρθρωτικές ανισότητες.

Επιπλέον, το άρθρο 35 του Χάρτη εγγυάται το δικαίωμα σε υψηλό επίπεδο προστασίας της ανθρώπινης υγείας και, ως εκ τούτου, ενισχύει την αρχή ότι τα συστήματα ΤΝ επιτρέπεται να αναπτύσσονται μόνο όταν αποδεδειγμένα συμβάλλουν σε ασφαλή, αποτελεσματική και δίκαιη υγειονομική περίθαλψη. Οι τεχνολογίες ΤΝ δεν θα πρέπει να εισάγουν νέα εμπόδια στην πρόσβαση και την περίθαλψη, ούτε να επιδεινώνουν υφιστάμενες ανισότητες που επηρεάζουν ιστορικά υποεξυπηρετούμενους πληθυσμούς ασθενών, ενώ ο αντίκτυπός τους στις ομάδες αυτές θα πρέπει να υπόκειται σε συνεχή αξιολόγηση.

Για τη διασφάλιση αυτών των δικαιωμάτων στην πράξη, **το κανονιστικό μοντέλο του Εγγράφου Πολιτικής (Policy Document) υιοθετεί μια προσέγγιση διακυβέρνησης που αφορά όλα τα στάδια του κύκλου ζωής των συστημάτων ΤΝ, η οποία οργανώνει την εποπτεία τους γύρω από μια σειρά υποχρεωτικών σημείων λήψης αποφάσεων ή αλλιώς «πυλών» (gates).** Οι πύλες αυτές εκφράζουν τη λογική των ρυθμιστικών διαδικασιών που ενσωματώνεται στον Κανονισμό για την ΤΝ (AI Act) και της συναφούς νομοθεσίας στον τομέα της υγείας, παρέχοντας στα νοσοκομεία τη δυνατότητα μιας σαφούς και ελέγξιμης διαδρομής διακυβέρνησης. Το μοντέλο λαμβάνει ως δεδομένο ότι ένα νοσοκομείο κατά κανόνα λειτουργεί κατά κανόνα ως φορέας ανάπτυξης/χρήσης (deployer) συστημάτων ΤΝ (σύμφωνα με όσα ορίζονται στον Κανονισμό για την ΤΝ) και ως πάροχος υπηρεσιών υγείας βάσει της εθνικής νομοθεσίας. Παράλληλα, το μοντέλο καλύπτει την περίπτωση κατά την οποία ένα νοσοκομείο καθίσταται πάροχος ή κατασκευαστής (provider/manufacturer) (για συστήματα ΤΝ που αναπτύσσονται εσωτερικά).

Τα νοσοκομεία αποτελούν τους κύριους φορείς ανάπτυξης και ελέγχου εφαρμογής των συστημάτων ΤΝ, ενώ οι οργανώσεις της κοινωνίας των πολιτών λειτουργούν ως ανεξάρτητοι φορείς εποπτείας, ελέγχου και παρακολούθησης με γνώμονα τα δικαιώματα, χωρίς άμεση εμπλοκή στην καθημερινή λήψη κλινικών αποφάσεων σχετικά με την ανάπτυξη και χρήση των συστημάτων.

Η εφαρμογή αυτού του κανονιστικού μοντέλου βασίζεται στην αρχή της αναλογικότητας, αναγνωρίζοντας ότι στα διαφορετικά κράτη μέλη και στα συστήματα υγείας **τα νοσοκομεία και οι οργανώσεις της κοινωνίας των**

πολιτών ενδέχεται να διαθέτουν διαφορετικά επίπεδα τεχνικής, οργανωτικής και αναλυτικής ικανότητας. Οι ελάχιστες υποχρεωτικές απαιτήσεις εφαρμόζονται ενιαία ώστε να διασφαλίζεται ένα βασικό επίπεδο προστασίας ως προς την ασφάλεια, τη διαφάνεια και την ισότητα.

Παράλληλα, πιο προηγμένες πρακτικές, όπως η εκτεταμένη ανάλυση πληθυσμιακών υποομάδων, οι εξωτερικοί έλεγχοι και τα εξελιγμένα εργαλεία παρακολούθησης προκαταλήψεων, μπορούν να εφαρμόζονται σταδιακά ανάλογα με τις δυνατότητες κάθε οργανισμού. Το τεχνικό παράρτημα, επίσης ακολουθεί αυτή την προσέγγιση, διακρίνοντας μεταξύ ελάχιστων και συνιστώμενων απαιτήσεων.

1. Πύλη πεδίου εφαρμογής και ταξινόμησης: Πριν από την προμήθεια ή την ανάπτυξη ενός συστήματος ΤΝ, οι οργανισμοί οφείλουν να καθορίζουν το κανονιστικό τους καθεστώς. Η διαδικασία αυτή περιλαμβάνει την αξιολόγηση του κατά πόσον το σύστημα συνιστά λογισμικό ιατροτεχνολογικού προϊόντος, σύμφωνα με τον Κανονισμό για τα Ιατροτεχνολογικά Προϊόντα (MDR) ή τον Κανονισμό για τα In Vitro Διαγνωστικά Ιατροτεχνολογικά Προϊόντα (IVDR), εάν αποτελεί σύστημα ΤΝ υψηλού κινδύνου κατά τον Κανονισμό για την ΤΝ ή εάν ανήκει σε κατηγορία χαμηλότερου κινδύνου που ωστόσο εξακολουθεί να απαιτεί επίβλεψη της διακυβέρνησης. Σε αυτό το πρώιμο στάδιο, τα νοσοκομεία μπορούν να επισημαίνουν συστήματα ΤΝ των οποίων η προβλεπόμενη χρήση ενδέχεται να έχει αυξημένες επιπτώσεις σε ζητήματα ισότητας ή να επηρεάζουν διαφορετικούς πληθυσμούς ασθενών, αναγνωρίζοντας ωστόσο ότι αυτά θα υποβληθούν σε πιο λεπτομερή αξιολόγηση ως προς τις προκαταλήψεις σε σχέση με το φύλο και τη φυλή σε επόμενα στάδια της διαδικασίας. Οι κατευθυντήριες οδηγίες του Συντονιστικού Οργάνου Ιατροτεχνολογικών Προϊόντων (MDCG) σχετικά με την πιστοποίηση και ταξινόμηση λογισμικού, παρέχουν πρακτικά παραδείγματα που συμβάλλουν στην αποφυγή εσφαλμένης ταξινόμησης, ενώ δημιουργούν τις βάσεις για μεταγενέστερες αξιολογήσεις ως προς τις προκαταλήψεις και τα ζητήματα ισότητας.³⁶ Σε αυτό το στάδιο, τα νοσοκομεία ηγούνται της διαδικασίας ταξινόμησης και του κανονιστικού προσδιορισμού, ενώ οι οργανώσεις της κοινωνίας των πολιτών παρέχουν ανεξάρτητες συμβουλευτικές γνωμοδοτήσεις σχετικά με κινδύνους που συνδέονται με ζητήματα ισότητας, χωρίς ωστόσο να διαθέτουν αρμοδιότητα για αποφάσεις σε αυτό το στάδιο.

2. Πύλη προμηθειών και συμβατικών υποχρεώσεων: Κατά το στάδιο των προμηθειών, τα νοσοκομεία ασκούν συχνά τη μεγαλύτερη επιρροή τους στη διακυβέρνηση των συστημάτων ΤΝ. Σε αυτό το πλαίσιο, θα πρέπει να

³⁶ Medical Device Coordination Group, [Κατευθυντήριες οδηγίες σχετικά με τον χαρακτηρισμό και την ταξινόμηση του λογισμικού στον Κανονισμό \(ΕΕ\) 2017/745 \(MDR\) και στον Κανονισμό \(ΕΕ\) 2017/746 \(IVDR\)](#), Brussels: MDCG, 2019.

διασφαλίζουν ότι οι προμηθευτές δεσμεύονται συμβατικά για την αντιμετώπιση κινδύνων που αφορούν ζητήματα ανισότητας, συμπεριλαμβανομένων των έμφυλων και φυλετικών προκαταλήψεων, σε όλα τα στάδια του κύκλου ζωής των συστημάτων. Επιπλέον, θα πρέπει να υποχρεούνται να παρέχουν τεκμηρίωση συμμόρφωσης με το ισχύον κανονιστικό πλαίσιο, όπως ενδεικτικά: αποδεικτικά στοιχεία αξιολόγησης της συμμόρφωσης, αποτελέσματα επικύρωσης απόδοσης και κάθε αναγκαία πληροφορία ώστε οι φορείς ανάπτυξης/ χρήσης (deployers) να εκπληρώνουν τις υποχρεώσεις τους βάσει του Κανονισμού για την ΤΝ. Οι συμβάσεις προμηθειών θα πρέπει να προσδιορίζουν ρητά την κατανομή ευθυνών μεταξύ παρόχου και φορέα ανάπτυξης/ χρήσης, συμπεριλαμβανομένων των απαιτήσεων σχετικά με τη διαφάνεια, την τεκμηρίωση, την καταγραφή συμβάντων (logging), την ανθρώπινη εποπτεία, καθώς και την τεχνική υποστήριξη μετά την ανάπτυξη. Στόχος της πύλης αυτής είναι η ενσωμάτωση διαδικασιών λογοδοσίας και ευαισθητοποίησης σχετικά με τις προκαταλήψεις που εντοπίζονται στο συμβατικό πλαίσιο, διασφαλίζοντας δίκαιη και ασφαλή ανάπτυξη. Σε αυτό το στάδιο, τα νοσοκομεία είναι υπεύθυνα για τις αποφάσεις προμηθειών και των συμβατικών τους ρυθμίσεων, ενώ οι οργανώσεις της κοινωνίας των πολιτών μπορούν να παρέχουν συμβουλευτική για ζητήματα ισότητας, χωρίς ωστόσο αυτή να είναι δεσμευτική.

3. Πύλη επικύρωσης: Το τρίτο στάδιο είναι η πύλη επικύρωσης, η οποία διασφαλίζει ότι τα συστήματα ΤΝ υπόκεινται σε ανεξάρτητη αξιολόγηση πριν ενταχθούν στις κλινικές ροές εργασίας. Η διαδικασία της επικύρωσης θα πρέπει να αξιολογεί τόσο τη συνολική απόδοση, όσο και τα αποτελέσματα σε διακριτούς πληθυσμούς ασθενών, με ιδιαίτερη έμφαση στο φύλο και στη φυλετική ή εθνοτική καταγωγή. Το στάδιο αυτό είναι καθοριστικής σημασίας, καθώς τα συστήματα ΤΝ στη βιοϊατρική μπορεί να παρουσιάζουν διαφοροποιημένη απόδοση μεταξύ πληθυσμιακών ομάδων όταν τα δεδομένα εκπαίδευσης δεν είναι αντιπροσωπευτικά ή όταν δεν λαμβάνονται υπόψη οι διαφοροποιήσεις μεταξύ του περιβάλλοντος ανάπτυξης και του περιβάλλοντος εφαρμογής των συστημάτων.³⁷ Παράλληλα, η διαδικασία αξιολόγησης θα πρέπει να τεκμηριώνει τα μέτρα μετριασμού που έχουν εφαρμοστεί και να αναδεικνύει τους πιθανούς εναπομείναντες κινδύνους αναφορικά με ζητήματα ανισότητας, παρέχοντας στα νοσοκομεία τα αναγκαία στοιχεία για τεκμηριωμένες αποφάσεις ανάπτυξης. Σε αυτό το στάδιο, τα νοσοκομεία είναι υπεύθυνα για τη διενέργεια ή ανάθεση της διαδικασίας επικύρωσης και για την ερμηνεία των αποτελεσμάτων στο πλαίσιο των αποφάσεων για την ανάπτυξη των συστημάτων ΤΝ, ενώ οι οργανώσεις της κοινωνίας των πολιτών παρέχουν ανεξάρτητη αξιολόγηση των δεικτών απόδοσης

³⁷ Rajkomar, A., Hardt, M., Howell, M. D., Corrado, G., & Chin, M. H., [Διασφάλιση της δικαιοσύνης στη μηχανική μάθηση για την προώθηση της ισότητας στην υγεία](#), *Annals of Internal Medicine*, 169(12), 866-872, 2018.

/ αποτελεσμάτων και των κινδύνων ως προς τυχόν προκαταλήψεις, χωρίς ωστόσο να έχουν έλεγχο επί των αποτελεσμάτων της διαδικασίας επικύρωσης.

4. Πύλη ανάπτυξης: Μετά την επιτυχή ολοκλήρωση του σταδίου της επικύρωσης, ακολουθεί η πύλη της διαδικασίας ανάπτυξης των συστημάτων TN. Σε αυτό το στάδιο, τα νοσοκομεία οφείλουν να διασφαλίζουν ότι οι συνθήκες λειτουργίας τους επιτρέπουν την ασφαλή, δίκαιη και υπεύθυνη χρήση τους. Αυτό περιλαμβάνει τον ορισμό κατάλληλα εκπαιδευμένου προσωπικού εποπτείας, την εφαρμογή μηχανισμών που επιτρέπουν στους/στις κλινικούς ιατρούς να παρακάμπτουν ή να απενεργοποιούν τις αλγοριθμικές συστάσεις, καθώς και τη διασφάλιση ότι οι χρήστριες/ χρήστες κατανοούν τον σκοπό, τους περιορισμούς και τυχόν διαφοροποιήσεις απόδοσης των συστημάτων μεταξύ διαφορετικών πληθυσμιακών ομάδων. Οι κλινικοί γιατροί θα πρέπει να ενημερώνονται για ενδεχόμενες αδικαιολόγητες διαφοροποιήσεις στα αποτελέσματα που σχετίζονται με το βιολογικό και το κοινωνικό φύλο ή τη φυλετική και εθνοτική καταγωγή, ενώ οι ασθενείς πρέπει να ενημερώνονται ότι για τη φροντίδα τους χρησιμοποιούνται εργαλεία υποστηριζόμενα από την τεχνητή νοημοσύνη. Σε αυτό το στάδιο, τα νοσοκομεία φέρουν την πλήρη ευθύνη για τις αποφάσεις ανάπτυξης, την κλινική ενσωμάτωση και την εποπτεία της εφαρμογής των συστημάτων TN, ενώ οι οργανώσεις της κοινωνίας των πολιτών δεν συμμετέχουν στη διαδικασία ανάπτυξης, ωστόσο μπορούν να αξιολογούν την ετοιμότητα ανάπτυξης υπό το πρίσμα ζητημάτων ισότητας και δικαιωμάτων.

5. Πύλη παρακολούθησης, ελέγχου και λογοδοσίας : Η τελική αυτή πύλη ρυθμίζει τη μετά-αναπτυξιακή φάση των συστημάτων TN. Τα συστήματα οφείλουν να υπόκεινται σε συνεχή παρακολούθηση ως προς τη συνολική τους απόδοση, τις διαφοροποιήσεις μεταξύ διακριτών πληθυσμών και τυχόν συμβάντα που σχετίζονται με την ασφάλεια ή με ζητήματα ισότητας. Τα νοσοκομεία θα πρέπει να τηρούν αρχεία λειτουργίας του συστήματος, να εφαρμόζουν τυποποιημένες διαδικασίες αναφοράς περιστατικών που καταγράφουν πιθανές προκαταλήψεις λόγω φύλου ή φυλής και να προβαίνουν σε περιοδική επανεκτίμηση της απόδοσης μέσω διαδικασιών επανεπικύρωσης. Σε περιπτώσεις όπου εντοπίζονται ανισότητες ή κίνδυνοι προκαταλήψεων που δεν μπορούν να αντιμετωπιστούν αποτελεσματικά, το σύστημα θα πρέπει να τροποποιείται, να αναστέλλεται ή και να αποσύρεται. Τα πορίσματα θα πρέπει να τροφοδοτούν τόσο την εσωτερική διακυβέρνηση, όσο και τη συλλογική μάθηση σε επίπεδο δικτύου – όπου ενδείκνυται - με στόχο την αποτροπή επανεμφάνισης παρόμοιων μορφών προκαταλήψεων. Σε αυτό το στάδιο, τα νοσοκομεία φέρουν την ευθύνη για τη συνεχή παρακολούθηση, την καταγραφή, την αναφορά περιστατικών και την εφαρμογή διορθωτικών μέτρων, ενώ οι οργανώσεις της κοινωνίας των πολιτών διενεργούν ανεξάρτητους ελέγχους, αξιολογούν συστημικούς κινδύνους ζητημάτων ισότητας μεταξύ διαφορετικών φορέων και υποστηρίζουν την

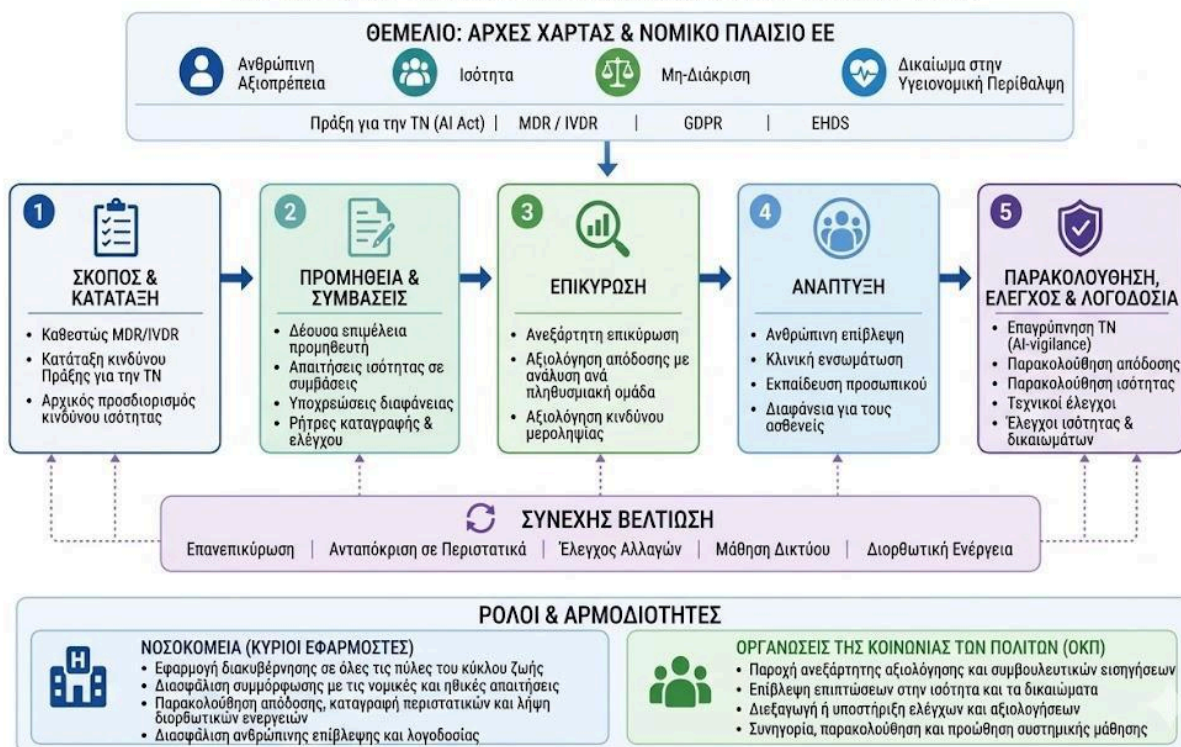
κλιμάκωση ζητημάτων που αφορούν σε επίμονες ανισότητες, χωρίς να ασκούν επιχειρησιακό έλεγχο επί των τροποποιήσεων των συστημάτων.

Συνολικά, οι εν λόγω πύλες μετασχηματίζουν τις αρχές των θεμελιωδών δικαιωμάτων σε συγκεκριμένες διαδικασίες διακυβέρνησης, οι οποίες είναι ικανές να αντιμετωπίσουν κινδύνους προκαταλήψεων που σχετίζονται με το φύλο και τη φυλετική καταγωγή στα κλινικά συστήματα τεχνητής νοημοσύνης.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them. Project code: 101215009 – AEQUITAS – CERV-2024-CHAR-LITI

Σχήμα 1: Πορεία διακυβέρνησης ΤΝ ευθυγραμμισμένη με την Χάρτα ΑΕQUITAS

**ΔΙΑΔΡΟΜΗ ΔΙΑΚΥΒΕΡΝΗΣΗΣ ΤΕΧΝΗΤΗΣ ΝΟΗΜΟΣΥΝΗΣ ΕΥΘΥΓΡΑΜΜΙΣΜΕΝΗ ΜΕ ΤΗ ΧΑΡΤΑ ΑΕQUITAS
ΓΙΑ ΤΗ ΒΙΟΙΑΤΡΙΚΗ ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΣΤΗ ΔΗΜΟΣΙΑ ΥΓΕΙΑ**



Πρακτικές πολιτικές, επιχειρησιακά εργαλεία και εργαλεία διακυβέρνησης για επαγγελματίες νοσοκομείων και οργανώσεις της κοινωνίας των πολιτών

Προκειμένου να διασφαλιστεί η αποτελεσματική λειτουργία του σε πραγματικά περιβάλλοντα υγειονομικής περίθαλψης, το προτεινόμενο κανονιστικό μοντέλο που περιλαμβάνεται στο Έγγραφο Πολιτικής προβλέπει την ανάπτυξη πρακτικών εργαλείων, τα οποία μπορούν να επαναχρησιμοποιούνται από διαφορετικά ιδρύματα και να προσαρμόζονται σε ποικίλα εθνικά συστήματα. Ιδιαίτερη έμφαση δίνεται στην ανάγκη για τυποποιημένες διαδικασίες, σαφούς τεκμηρίωσης και κοινών προτύπων, τα οποία επιτρέπουν τόσο στα νοσοκομεία όσο και στις οργανώσεις της κοινωνίας των πολιτών να διαχειρίζονται τα συστήματα τεχνητής νοημοσύνης με συνεπή τρόπο. Το σύνολο των εργαλείων που ακολουθεί ευθυγραμμίζεται με τις απαιτήσεις του Κανονισμού για την ΤΝ (AI Act), τις

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them. Project code: 101215009 – AEQUITAS – CERV-2024-CHAR-LITI

κατευθυντήριες οδηγίες του Κανονισμού για τα Ιατροτεχνολογικά Προϊόντα (*Medical Device Regulation/MDR*) και του Κανονισμού για τα In Vitro Διαγνωστικά Ιατροτεχνολογικά Προϊόντα (*In Vitro Diagnostic Regulation /IVDR*), καθώς και με ευρύτερες διατομεακές πρακτικές διακυβέρνησης.

Στάδιο προμηθειών και συμβάσεων

- **Πακέτο προμηθειών και συμβάσεων:** Η κοινότητα προμηθειών TN της ΕΕ έχει δημοσιεύσει πρότυπες συμβατικές ρήτρες (σε εκδόσεις για συστήματα υψηλού και μη υψηλού κινδύνου) και τις έχει επικαιροποιήσει ώστε να ευθυγραμμίζονται με τον Κανονισμό για την TN (AI Act), ενθαρρύνοντας τις δημόσιες αρχές να τις αξιοποιούν, και παρέχοντας επεξηγηματικές οδηγίες για την εφαρμογή τους.³⁸ Στο πλαίσιο της TN στην υγειονομική περίθαλψη και με στόχο την αξιολόγηση και τον μετριασμό των έμφυλων και φυλετικών προκαταλήψεων, οι εν λόγω ρήτρες μπορούν να προσαρμοστούν σε ένα εξειδικευμένο παράρτημα για νοσοκομεία, το οποίο θα είναι ένα αυτόνομο εργαλείο, το οποίο κατ' ελάχιστον απαιτεί τα εξής:
 - (1) Τυποποιημένη καταγραφή της σύνθεσης των δεδομένων εκπαίδευσης των συστημάτων, συμπεριλαμβανόμενης της εκπροσώπησης κατά φύλο και φυλετική/ εθνοτική ομάδα, προκειμένου να καθίσταται δυνατή η αξιολόγηση της αντιπροσωπευτικότητας των δεδομένων και ο εντοπισμός του κινδύνου μεροληψίας.
 - (2) Γνωστοποίηση της προέλευσης των δεδομένων, η οποία περιλαμβάνει πληροφορίες σχετικά με το εάν τα χρησιμοποιούμενα σύνολα δεδομένων είναι ιδιοκτησιακά ή ανοικτά, τους όρους πρόσβασης και χρήσης τους, καθώς και τις συνέπειες που ενδέχεται να έχουν για τη διαφάνεια του συστήματος, τη δυνατότητα ανεξάρτητης επικύρωσης και την αξιολόγηση πιθανών φαινομένων έμφυλων ή φυλετικών μεροληψιών.
 - (3) Γνωστοποίηση σχετικά με την απόδοση του συστήματος σε διαφορετικές πληθυσμιακές ομάδες, συνοδευόμενης από την απαραίτητη τεκμηρίωση επικύρωσης, συμπεριλαμβανομένων αποτελεσμάτων που παρουσιάζονται ανά φύλο και φυλετική ή εθνοτική καταγωγή. Η τεκμηρίωση αυτή θα πρέπει να περιλαμβάνει καταγραφή τυχόν γνωστών αποκλίσεων στην απόδοση, περιγραφή των μέτρων που έχουν ληφθεί για τον περιορισμό των προκαταλήψεων κατά τη διάρκεια της ανάπτυξης των συστημάτων, καθώς και τα σχετικά αποτελέσματα δοκιμών και αξιολογήσεων που τεκμηριώνουν την

³⁸ European Commission, [Οι επικαιροποιημένες πρότυπες συμβατικές ρήτρες της ΕΕ για την τεχνητή νοημοσύνη είναι πλέον διαθέσιμες](#), Brussels: European Commission, 2025.

αποτελεσματικότητά τους (συμπεριλαμβανομένων των ειδικών για κάθε τοποθεσία παραμέτρων που σχετίζονται με την ισότητα και τη δικαιοσύνη).

- (4) Καθορισμό των προϋποθέσεων για την καταγραφή (logging) της λειτουργίας του συστήματος και τη δυνατότητα ελέγχου του, ώστε να παράγονται κατάλληλα αρχεία χρήσης που να επιτρέπουν τη διενέργεια ανεξάρτητων ελέγχων για τον εντοπισμό πιθανών προκαταλήψεων ή άνισων επιπτώσεων σε διαφορετικές πληθυσμιακές ομάδες.
- (5) Υποχρεώσεις διαχείρισης των αλλαγών και των ενημερώσεων που προκύπτουν στα συστήματα ΤΝ, σύμφωνα με τις οποίες ο πάροχος οφείλει να ενημερώνει το νοσοκομείο για τυχόν αναβαθμίσεις, επανεκπαίδευση του μοντέλου ή μεταβολές στη ροή εργασίας και τις πιθανές επιπτώσεις αυτών στη λειτουργία των συστημάτων σε σχέση με διαφορετικές πληθυσμιακές ομάδες.
- (6) Δεσμεύσεις για την αντιμετώπιση περιστατικών που σχετίζονται με προκαταλήψεις, οι οποίες υποχρεώνουν τον πάροχο να αναφέρει τα σφάλματα που έχουν εντοπιστεί ή τις ανισότητες που επηρεάζουν ομάδες ασθενών, οι οποίες διακρίνονται με βάση το βιολογικό και κοινωνικό φύλο ή τη φυλετική/εθνοτική καταγωγή τους και να περιγράφει τα μέτρα για την αποκατάστασή τους.
- (7) Σαφής κατανομή αρμοδιοτήτων, η οποία προσδιορίζει τις υποχρεώσεις του παρόχου και του διαχειριστή όσον αφορά τη συμμόρφωση με τις απαιτήσεις του Κανονισμού για την Τεχνητή Νοημοσύνη και τη διαχείριση των κινδύνων που σχετίζονται με τα ζητήματα ισότητας και προκαταλήψεις.

Στάδια επικύρωσης και προ-ανάπτυξης

- **Πολιτική διακυβέρνησης δεδομένων υγείας και αντιπροσωπευτικότητας** : Για τον μετριασμό των έμφυλων και φυλετικών προκαταλήψεων, όλα τα νοσοκομεία και οι οργανώσεις της κοινωνίας των πολιτών που αναπτύσσουν συστήματα ΤΝ πρέπει να υιοθετήσουν μία πολιτική που θα διασφαλίζει ότι τα δεδομένα εκπαίδευσης, ρύθμισης ή εφαρμογής είναι αντιπροσωπευτικά, πλήρη και ακριβή για τους πληθυσμούς που εξυπηρετούν. Αυτό περιλαμβάνει ανάλυση δημογραφικής σύνθεσης, καταγραφή υποεκπροσωπούμενων πληθυσμών και καθορισμό μέτρων μετριασμού κινδύνων, όπου αυτό απαιτείται. Η πολιτική θα πρέπει να καλύπτει τόσο εσωτερικά αναπτυγμένα όσο και εξωτερικά προερχόμενα σύνολα δεδομένων, συμπεριλαμβανομένων ιδιοκτησιακών και ανοικτής πρόσβασης, και να απαιτεί την επαλήθευση της απόδοσής τους σε διακριτές πληθυσμιακές ομάδες, καθώς και τις στρατηγικές μετριασμού των προκαταλήψεων που εφαρμόζονται από τους

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them. Project code: 101215009 – AEQUITAS – CERV-2024-CHAR-LITI

παρόχους. Η εν λόγω διακυβέρνηση ευθυγραμμίζεται με τον Κανονισμό για την TN (άρθρο 10), ο οποίος απαιτεί τον έλεγχο και τον περιορισμό των κινδύνων που αφορούν μεροληψίες, ενώ αντανάκλα βέλτιστες πρακτικές στη βιβλιογραφία για τον μετριασμό τους στα ιατρικά συστήματα TN, καλύπτοντας όλα τα στάδια του κύκλου ζωής από τη συλλογή δεδομένων έως και την παρακολούθηση μετά την ανάπτυξη του συστήματος.³⁹ Η συγκεκριμένη πολιτική δεν αποτελεί αφ' εαυτής ένα μεμονωμένο τεκμήριο, αλλά ένα επίπεδο πολιτικής που καθοδηγεί άλλα τεκμήρια. Λειτουργεί ως αρχή προ-ανάπτυξης και τροφοδοτεί άμεσα τη συνεχή παρακολούθηση, τον έλεγχο και τους μηχανισμούς λογοδοσίας.

- **Βασικό επίπεδο κυβερνοασφάλειας και μοντέλο απειλών για την Τεχνητή Νοημοσύνη:** Τα συστήματα TN στα νοσοκομεία μπορεί να είναι ευάλωτα όχι μόνο σε συμβατικές επιθέσεις, όπως η «δηλητηρίαση δεδομένων» (*data poisoning*), αντιπαραθετικά παραδείγματα (*adversarial examples*) ή παράκαμψη του μοντέλου (*model evasion*), αλλά και σε χειρισμούς ή σφάλματα που επιδεινώνουν έμφυλες ή φυλετικές προκαταλήψεις, για παράδειγμα επηρεάζοντας δυσανάλογα υποεκπροσωπούμενες δημογραφικές ομάδες στα δεδομένα εκπαίδευσης ή υλοποίησης. Ο Κανονισμός για την TN (άρθρο 15) απαιτεί ανθεκτικότητα έναντι μη εξουσιοδοτημένων αλλοιώσεων των αποτελεσμάτων ή της απόδοσης, αναδεικνύοντας την ανάγκη για ειδικές δικλίδες ασφαλείας στα συστήματα TN. Ως εκ τούτου, τα νοσοκομεία θα πρέπει να θεσπίσουν μια βασική γραμμή κυβερνοασφάλειας και ένα μοντέλο απειλών ειδικό για την TN, το οποίο θα λαμβάνει ρητά υπόψη κινδύνους που σχετίζονται με προκαταλήψεις, παράλληλα με τις τεχνικές απειλές. Αυτό το «ζωντανό» έγγραφο θα λειτουργεί τόσο ως πολιτική όσο και ως πρακτικό εργαλείο, χρησιμεύοντας ως λίστα ελέγχου πριν από την ανάπτυξη (*pre-deployment checklist*), ώστε να διασφαλίζεται η εφαρμογή ελέγχων κυβερνοασφάλειας ευθυγραμμισμένων με τους Κανονισμούς MDR και IVDR, με την οργανωσιακή διακυβέρνηση ασφάλειας πληροφοριών (π.χ. ISO/IEC 27001), καθώς και με μοντέλα απειλών και σχέδια δοκιμών που είναι ευαίσθητα στις προκαταλήψεις.⁴⁰ Επίσης, υποστηρίζει τη συνεχή παρακολούθηση μετά την ανάπτυξη του συστήματος, ώστε να εντοπίζονται αναδυόμενοι τεχνικοί και σχετικοί με τις προκαταλήψεις κίνδυνοι και να προβλεφθεί η προληπτική αντιμετώπισή τους. Με την εφαρμογή αυτής της πολιτικής, τα νοσοκομεία μπορούν να ενισχύσουν την ανθεκτικότητα των συστημάτων TN που χρησιμοποιούν, να

³⁹ Char, D.S., Shah, N.H., & Magnus, D., [Εφαρμογή της μηχανικής μάθησης στον τομέα της υγειονομικής περίθαλψης – Αντιμετώπιση ηθικών προκλήσεων](#), *New England Journal of Medicine* 378: 981-983 (2018).

⁴⁰ Medical Device Coordination Group, [MDCG 2019 16 Αναθ. 1 - Κατευθυντήριες οδηγίες για την κυβερνοασφάλεια των ιατρικών συσκευών](#), Brussels: European Commission, 2020.

- προστατεύσουν την ισότητα στη λήψη αποφάσεων με τη βοήθεια της ΤΝ και να διασφαλίσουν τη συμμόρφωση με τον Κανονισμό για την ΤΝ (άρθρα 9–15) και τις προϋποθέσεις των MDR/IVDR.
- **Ένας «φάκελος συστήματος ΤΝ» ανά εργαλείο:** Αυτό το τεκμήριο αποτελεί ένα τυποποιημένο έγγραφο που τηρείται για κάθε σύστημα ΤΝ πριν από την ανάπτυξή του σε έναν οργανισμό υγειονομικής περίθαλψης. Ο «φάκελος» θα πρέπει να συνοψίζει τον προβλεπόμενο σκοπό του συστήματος, την ενσωμάτωσή του στην κλινική ροή εργασίας, τα αποτελέσματα επικύρωσης και τους δείκτες απόδοσης, με ιδιαίτερη έμφαση στην απόδοση σε πληθυσμιακές ομάδες ασθενών που διαφοροποιούνται ως προς το βιολογικό και το κοινωνικό φύλο ή τη φυλετική/εθνοτική καταγωγή, καθώς και σε άλλες τυχόν γνωστές ανισότητες. Θα πρέπει να τεκμηριώνει τις διαδικασίες της ανθρώπινης εποπτείας, να διευκρινίζει ποιες αποφάσεις παραμένουν αποκλειστικά υπό τον έλεγχο του/της κλινικού/κλινικής ιατρού και να καταγράφει τυχόν αντιφάσεις σχετικές με την ασφάλεια και την ισότητα των ασθενών. Παρέχοντας μια τυποποιημένη, ευαισθητοποιημένη ως προς την πρόληψη για τις προκαταλήψεις επισκόπηση του συστήματος ΤΝ, ο «φάκελος» υποστηρίζει τη διαφάνεια, την εσωτερική διακυβέρνηση και τον ρυθμιστικό έλεγχο, διασφαλίζοντας ότι οι αποφάσεις πριν από την ανάπτυξη των συστημάτων λαμβάνουν σοβαρά υπόψη τους κινδύνους να παραχθούν άνισα αποτελέσματα.
 - **Ένα συνδυαστικό πακέτο για την Εκτίμηση των Επιπτώσεων στα Θεμελιώδη Δικαιώματα (*Fundamental Rights Impact Assessment /FRIA*) και την Εκτίμηση Αντικτύπου σχετικά με την Προστασία Δεδομένων (*Data Protection Impact Assessments/ DPIA*):** Τα δημόσια νοσοκομεία που αναπτύσσουν συστήματα ΤΝ υψηλού κινδύνου καλούνται ολοένα και περισσότερο να αξιολογούν τόσο τους κινδύνους που αφορούν τα θεμελιώδη δικαιώματα, όσο και αυτούς για την προστασία των προσωπικών δεδομένων. Η ενοποίηση αυτών των αξιολογήσεων σε ένα ενιαίο πρότυπο που συμμορφώνεται με τους Κανονισμούς για την ΤΝ (AI Act, άρθρο 27), τα προσωπικά δεδομένα (GDPR) και την Εκτίμηση Αντικτύπου σχετικά με την Προστασία Δεδομένων (DPIA), μπορεί να μειώσει το διοικητικό βάρος, διασφαλίζοντας παράλληλα ότι βασικά ερωτήματα σχετικά με τους πληθυσμούς που επηρεάζονται, τις πιθανές βλάβες που προκαλούνται, τις στρατηγικές μετριασμού που υιοθετούνται και τους μηχανισμούς καταγγελιών αντιμετωπίζονται συστηματικά, με ιδιαίτερη προσοχή σε πιθανές ανισότητες που προκύπτουν λόγω έμφυλων και φυλετικών προκαταλήψεων. Το «συνδυαστικό πακέτο FRIA και DPIA» λειτουργεί κυρίως ως ένα εκ των προτέρων (*ex-ante*) και περιοδικά επικαιροποιούμενο τεκμήριο ελέγχου (*audit artefact*), κατά το στάδιο πριν από την ανάπτυξη του συστήματος και στο επίπεδο του

ελέγχου για ζητήματα ισότητας και δικαιωμάτων. Οι αξιολογήσεις αυτές αποτελούν οργανωμένες, τεκμηριωμένες εκτιμήσεις κινδύνων για τα θεμελιώδη δικαιώματα και την προστασία των δεδομένων, λειτουργώντας ως συγκεκριμένα σημεία ελέγχου διακυβέρνησης και όχι ως εργαλεία συνεχούς παρακολούθησης. Παρότι θα πρέπει να επικαιροποιούνται, όταν αλλάζουν οι κίνδυνοι ή οι συνθήκες του συστήματος, ο βασικός τους ρόλος είναι να παρέχουν μια συστηματική και συνεχή αξιολόγηση πιθανών επιπτώσεων πριν και κατά την ανάπτυξη των συστημάτων, υποστηρίζοντας τις συμμορφώσεις τόσο με τον Κανονισμό για την ΤΝ όσο και με αυτόν για τα προσωπικά δεδομένα (GDPR) βοηθώντας στον εντοπισμό και τον μετριασμό έμφυλων και φυλετικών προκαταλήψεων. Για την ενίσχυση της διάστασης των ανθρωπίνων δικαιωμάτων, θα πρέπει κατά τον σχεδιασμό των εκτιμήσεων αντικτύπου να προβλέπεται ουσιαστική διαβούλευση με τις οργανώσεις της κοινωνίας των πολιτών, ώστε να ενσωματώνεται σε αυτές η εξειδικευμένη γνώση τους σε θέματα δικαιωμάτων. Για παράδειγμα, το Ευρωπαϊκό Δίκτυο Εθνικών Οργανισμών για τα Ανθρώπινα Δικαιώματα (*European Network of National Human Rights Institutions/ ENNHRI*) καλεί την Ευρωπαϊκή Επιτροπή να συμπληρώσει τα πρότυπα με κατευθυντήριες οδηγίες και δημόσια διαβούλευση, ώστε να διασφαλιστούν αποτελεσματικές εκτιμήσεις αντικτύπου (FRIAs), οι οποίες λαμβάνουν υπόψη την ισότητα ως προς το φύλο και τις φυλετικές/εθνοτικές διαστάσεις.⁴¹

- **Πολιτική ελέγχου μεταβολών:** Αυτή η πολιτική διέπει τις ενημερώσεις των συστημάτων ΤΝ, συμπεριλαμβανομένης της επανεκπαίδευσης μοντέλων, των αναβαθμίσεων λογισμικού ή των τροποποιήσεων στις κλινικές ροές εργασίας. Οι οργανισμοί θα πρέπει να ορίζουν με σαφήνεια τι συνιστά «ουσιώδη μεταβολή» που απαιτεί επαναξιολόγηση ή έγκριση, διασφαλίζοντας ότι οι ενημερώσεις δεν εισάγουν απρόβλεπτους κινδύνους που θα επηρεάσουν τη φροντίδα των ασθενών. Ειδικότερα, οι μεταβολές πρέπει να αξιολογούνται ως προς τις πιθανές επιπτώσεις που ενδέχεται να έχουν σε ζητήματα έμφυλης και φυλετικής ισότητας, συμπεριλαμβανομένων μεταβολών στην αποδοτικότητα του συστήματος ανά πληθυσμιακή υποομάδα ή διαφοροποιήσεων στα αποτελέσματα μεταξύ πληθυσμών που διακρίνονται βάσει άλλων δημογραφικών χαρακτηριστικών. Τα νοσοκομεία θα πρέπει επίσης να διασφαλίζουν ότι οι κλινικοί γιατροί και το σχετικό προσωπικό είναι ενήμεροι για τροποποιήσεις που ενδέχεται να επηρεάσουν τα αποτελέσματα της φροντίδας για διαφορετικούς πληθυσμούς ασθενών, ώστε να μπορούν να εφαρμοστούν κατάλληλα μέτρα εποπτείας και μετριασμού κινδύνων. Με

⁴¹ European Network of National Human Rights Institutions (ENNHRI), [Δήλωση σχετικά με τη διασφάλιση της αποτελεσματικότητας των εκτιμήσεων επιπτώσεων στα θεμελιώδη δικαιώματα \(FRIAs\) στο πλαίσιο του κανονισμού της ΕΕ για την τεχνητή νοημοσύνη](#), Brussels: ENNHRI, 2025.

την ενσωμάτωση παραμέτρων που αφορούν ζητήματα ισότητας στη διαδικασία διαχείρισης αλλαγών, τα νοσοκομεία διατηρούν τα επίπεδα ασφάλειας και δικαιοσύνης στη λήψη των αποφάσεων που υποστηρίζονται από συστήματα Τεχνητής Νοημοσύνης στην κλινική πρακτική.

- **Σύστημα Διαχείρισης Ποιότητας (Quality Management System - QMS) για φορείς ανάπτυξης συστημάτων ΤΝ που είναι ταυτοχρόνως και πάροχοι:** Μια ακόμη σημαντική διάσταση του προτεινόμενου μοντέλου αφορά τα συστήματα Τεχνητής Νοημοσύνης που αναπτύσσονται ενδογενώς από νοσοκομεία και ερευνητικά ιδρύματα. Η τάση των δημόσιων οργανισμών υγείας να αναπτύσσουν ή να προσαρμόζουν ιδιόκτητα συστήματα Τεχνητής Νοημοσύνης ενισχύεται διαρκώς, ιδίως στο πλαίσιο των ακαδημαϊκών ιατρικών κέντρων. Στις περιπτώσεις αυτές, το ίδιο το ίδρυμα αναλαμβάνει ουσιαστικά τον ρόλο του κατασκευαστή ή του παρόχου του συστήματος, με αποτέλεσμα να υπέχει πρόσθετες κανονιστικές και ρυθμιστικές υποχρεώσεις. Οι ευρωπαϊκές κατευθυντήριες οδηγίες για τα ιατροτεχνολογικά προϊόντα που αναπτύσσονται και χρησιμοποιούνται εντός των ίδιων των υγειονομικών ιδρυμάτων (*in-house medical devices*) διευκρινίζουν ότι τα ιδρύματα υγείας οφείλουν να εφαρμόζουν συστήματα διαχείρισης ποιότητας και να διαθέτουν τεχνική τεκμηρίωση, η οποία να αποδεικνύει ότι τα εσωτερικά αναπτυγμένα εργαλεία ανταποκρίνονται σε συγκεκριμένες ανάγκες των ασθενών που δεν μπορούν να καλυφθούν από διαθέσιμα εμπορικά προϊόντα.⁴² Το Σύστημα Διαχείρισης Ποιότητας θα πρέπει να διασφαλίζει ότι τα αποτελέσματα και οι επιδόσεις των συστημάτων Τεχνητής Νοημοσύνης αξιολογούνται ως προς τυχόν διαφοροποιήσεις μεταξύ πληθυσμιακών ομάδων που διακρίνονται βάσει βιολογικού και κοινωνικού φύλου και φυλετικής ή εθνοτικής καταγωγής, ότι οι στρατηγικές μετριασμού των σχετικών κινδύνων τεκμηριώνονται επαρκώς και ότι οι μηχανισμοί ανθρώπινης εποπτείας έχουν σχεδιαστεί κατά τρόπο που να διασφαλίζουν ρητά τόσο την ασφάλεια των ασθενών όσο και τη δίκαιη μεταχείρισή τους. Ακόμη και όταν εφαρμόζονται κανονιστικές εξαιρέσεις για τα εσωτερικά (*in-house*) ιατροτεχνολογικά προϊόντα, οι υποχρεώσεις που απορρέουν από τα θεμελιώδη δικαιώματα και τα ευρύτερα πλαίσια διακυβέρνησης της Τεχνητής Νοημοσύνης εξακολουθούν να ισχύουν πλήρως. Ως εκ τούτου, το Σύστημα Διαχείρισης Ποιότητας αποτελεί κρίσιμο μηχανισμό για την ενσωμάτωση της αρχής της ισότητας και της μη διάκρισης στον σχεδιασμό, την επικύρωση και την ανάπτυξη βιοϊατρικών συστημάτων Τεχνητής Νοημοσύνης.

⁴² Medical Device Coordination Group, [MDCG 2023 1 – Κατευθυντήριες οδηγίες σχετικά με την εξαίρεση των ιδρυμάτων υγείας \(MDR/IVDR\)](#), Brussels: European Commission, 2023.

Στάδιο παρακολούθησης, ελέγχου και λογοδοσίας

- **Έκθεση απόδοσης και ισότητας με ανάλυση ανά πληθυσμιακή ομάδα:** Δεδομένου ότι οι διαφοροποιήσεις στην επίδοση των αλγορίθμων μπορούν να συμβάλλουν σε άνισα αποτελέσματα σε σχέση με θέματα υγείας, τα νοσοκομεία δεν οφείλουν μόνο να απαιτούν από τους προμηθευτές να παρέχουν σχετικές πληροφορίες πριν από την προμήθεια ενός συστήματος, αλλά και να εκπονούν εσωτερικές αναλύσεις που να δείχνουν πώς τα συστήματα αποδίδουν σε ασθενείς με διαφορετικά δημογραφικά χαρακτηριστικά. Οι αναλύσεις αυτές θα πρέπει να εστιάζουν στη διαστρωματωμένη αξιολόγηση της αποτελεσματικότητας ως προς το βιολογικό φύλο και την ταυτότητα φύλου, καθώς και στις σχετικές φυλετικές ή εθνοτικές διαφοροποιήσεις, όπου αυτό είναι νόμιμο και εφικτό, παρέχοντας παράλληλα αιτιολόγηση για τυχόν αναπάνευκτα κενά στα αποτελέσματα. Όπου εντοπίζονται διαφορές στην απόδοση των αποτελεσμάτων, οι οργανισμοί οφείλουν να τεκμηριώνουν τις στρατηγικές μετριασμού των διαφορών και να παρακολουθούν διαχρονικά την αποτελεσματικότητα των μέτρων αυτών. Η προσέγγιση αυτή αντανακλά τις προϋποθέσεις που τίθενται στο πλαίσιο του ευρωπαϊκού κανονιστικού πλαισίου για την Τεχνητή Νοημοσύνη, σύμφωνα με τις οποίες οι οργανισμοί πρέπει να αξιολογούν ενεργά και να μετριάσουν τους κινδύνους προκατάληψης τόσο στα σύνολα των δεδομένων που χρησιμοποιούνται, όσο και στα αποτελέσματα των μοντέλων. Ο Κανονισμός για την ΤΝ (άρθρο 10) προβλέπει ρητά την αξιολόγηση και τον περιορισμό των κινδύνων προκατάληψης στα σύνολα δεδομένων, ενώ παράλληλα επιτρέπει, υπό αυστηρές εγγυήσεις, την περιορισμένη επεξεργασία δεδομένων ειδικών κατηγοριών με σκοπό τον εντοπισμό πιθανών προκαταλήψεων.
- Η συγκεκριμένη αναφορά λειτουργεί πρωτίστως ως εργαλείο συνεχούς παρακολούθησης εντός του πλαισίου διακυβέρνησης. Απαιτεί συνεχή συλλογή δεδομένων, τακτική παρακολούθηση της επίδοσης των συστημάτων σε διαφορετικές δημογραφικές ομάδες και συνεχή επικαιροποίηση των μέτρων μετριασμού κινδύνων όταν εντοπίζονται διαφοροποιημένα αποτελέσματα. Με τον τρόπο αυτό, οι φορείς υγειονομικής περίθαλψης μπορούν να εντοπίζουν και να αντιμετωπίζουν τους κινδύνους προκατάληψης σε βάθος χρόνου. Παράλληλα, τα βασικά ευρήματα της αναφοράς, όπως οι διαπιστωμένες ανισότητες και η αποτελεσματικότητα των στρατηγικών μετριασμού κινδύνων, θα πρέπει να συνοψίζονται και να ενσωματώνονται στον «έλεγχο για ζητήματα ισότητας και δικαιωμάτων» (*equity and rights audit*), διασφαλίζοντας τη διαφάνεια και τη δημόσια λογοδοσία.

- **Σύστημα ελέγχου δύο επιπέδων, το οποίο περιλαμβάνει έναν τεχνικό έλεγχο και έναν έλεγχο για ζητήματα ισότητας και δικαιωμάτων:** Ένα εφαρμόσιμο πλαίσιο ελέγχου (*audit regime*) για τα νοσοκομεία θα πρέπει να συνδυάζει την αυστηρότητα με την πρακτικότητα και ταυτόχρονα να είναι κατανοητό τόσο από το προσωπικό όσο και από το ευρύ κοινό. Στο πλαίσιο αυτό, ο έλεγχος (*audit*) νοείται ως μια τυποποιημένη και περιοδική διαδικασία αξιολόγησης της συμμόρφωσης, της επίδοσης και των επιπτώσεων ενός συστήματος ΤΝ στα δικαιώματα των ατόμων, η οποία συμπληρώνει τη συνεχή παρακολούθηση χωρίς να την υποκαθιστά.
- **Τεχνικός έλεγχος (*Technical audit*):** Το επίπεδο αυτό συνιστά μια εμπιστευτική επιχειρησιακή αξιολόγηση και περιλαμβάνει λεπτομερείς ελέγχους που παραμένουν εσωτερικοί στον οργανισμό, όπως δοκιμές ασφάλειας, αξιολογήσεις ανθεκτικότητας, επικύρωση της κλινικής απόδοσης καθώς και έλεγχο συμμόρφωσης με τις απαιτήσεις διακυβέρνησης δεδομένων. Ιδιαίτερη έμφαση θα πρέπει να δίνεται στην αξιολόγηση της επίδοσης του συστήματος λαμβάνοντας υπόψη το βιολογικό φύλο και την ταυτότητα φύλου, καθώς και τις φυλετικές ή εθνοτικές διαφοροποιήσεις, εξετάζοντας κατά πόσο τα αποτελέσματα του μοντέλου ενδέχεται να θέτουν συστηματικά σε μειονεκτική θέση συγκεκριμένες πληθυσμιακές ομάδες. Οι δραστηριότητες αυτές θα πρέπει να εναρμονίζονται με τις σχετικές διατάξεις του Κανονισμού για την ΤΝ (AI Act) (άρθρα 9–15), καθώς και με τον Κανονισμό για τα Ιατροτεχνολογικά Προϊόντα (MDR) και τον Κανονισμό για τα In Vitro Διαγνωστικά Ιατροτεχνολογικά Προϊόντα (IVDR), όπου απαιτείται η εφαρμογή τους.
- **Έλεγχος για ζητήματα ισότητας και δικαιωμάτων (*Equity and rights audit*):** Το δεύτερο αυτό επίπεδο ελέγχου επικεντρώνεται στη διαφάνεια, τη λογοδοσία και τον μετριασμό των προκαταλήψεων. Θα πρέπει να τεκμηριώνει τις ανισότητες στα αποτελέσματα, τις πληθυσμιακές ομάδες που συμπεριλήφθηκαν στις δοκιμές, τους γνωστούς περιορισμούς του συστήματος και τους διαθέσιμους μηχανισμούς υποβολής παραπόνων, περιλαμβάνοντας παράλληλα ρητή ανάλυση των επιπτώσεων που εμφανίζονται σε σχέση με το φύλο και τη φυλή. Επιπλέον, θα πρέπει να καταγράφονται στοιχεία που αποδεικνύουν ότι η ανθρώπινη εποπτεία είναι ουσιαστική και ικανή να εντοπίζει και να διορθώνει αποτελέσματα του συστήματος που προκύπτουν από προκαταλήψεις. Με βάση την προσέγγιση των «ελέγχων και γνωστοποιήσεων προκατάληψης» (*bias audit & disclosure*), η οποία εφαρμόζεται στον τομέα της εργασιακής ρύθμισης,⁴³ μπορεί να συνταχθεί μια συνοπτική έκθεση προσανατολισμένη στο ευρύ κοινό η οποία θα παρουσιάζει τα βασικά ευρήματα, τις στρατηγικές αντιμετώπισης των προκαταλήψεων και τις

⁴³ New York City Department of Consumer and Worker Protection, [Αυτοματοποιημένα εργαλεία λήψης αποφάσεων σχετικά με την απασχόληση \(AEDT\)](#), New York: City of New York, 2025.

βελτιώσεις που έχουν εφαρμοστεί, χωρίς να αποκαλύπτονται εμπιστευτικές ή ιδιοκτησιακές πληροφορίες. Παράλληλα, «το συνδυαστικό πακέτο FRIA και DPIA», το οποίο αναφέρεται παραπάνω, εντάσσεται στο επίπεδο του ελέγχου για ζητήματα ισότητας και δικαιωμάτων, ενισχύοντας τη συστηματική αξιολόγηση των θεμελιωδών δικαιωμάτων, της προστασίας δεδομένων και της δημογραφικής ισότητας.

- **Μητρώο “επαγρύπνησης” για την TN (AI vigilance log).** Η παρακολούθηση μετά την εφαρμογή ενός συστήματος TN υποστηρίζεται επίσης μέσω της δημιουργίας ενός αρχείου/μητρώου «επαγρύπνησης» για την TN, εμπνευσμένου από τα συστήματα φαρμακοεπαγρύπνησης που χρησιμοποιούνται στη κανονιστική ρύθμιση του τομέα των φαρμάκων.⁴⁴ Οι επαγγελματίες υγείας θα πρέπει να έχουν τη δυνατότητα να αναφέρουν περιστατικά, παρ’ ολίγο ατυχήματα (*near-misses*) ή απρόβλεπτα αποτελέσματα που σχετίζονται με αποφάσεις υποστηριζόμενες από συστήματα TN, με σαφή πεδία καταγραφής διαφορών που σχετίζονται με το φύλο και την ταυτότητα φύλου ή τις φυλετικές/εθνοτικές ομάδες. Οι αναφορές αυτές θα πρέπει να αναλύονται συστηματικά για την ανίχνευση αναδυόμενων προτύπων, συμπεριλαμβανομένων διαφοροποιημένων επιπτώσεων σε συγκεκριμένες δημογραφικές ομάδες, ώστε οι οργανισμοί υγειονομικής περίθαλψης να μπορούν να εντοπίζουν θέματα ασφάλειας, ζητήματα ισότητας ή τάσεις προκατάληψης. Όπου είναι απαραίτητο, τα περιστατικά, ιδίως εκείνα που υποδεικνύουν άνιση μεταχείριση ή κινδύνους για προστατευόμενες ομάδες, θα πρέπει να κοινοποιούνται άμεσα στους παρόχους των συστημάτων, στις αρμόδιες επιτροπές διακυβέρνησης των νοσοκομείων και στις ρυθμιστικές αρχές. Το μητρώο αυτό θα πρέπει επίσης να τροφοδοτεί περιοδικούς ελέγχους για ζητήματα ισότητας και δικαιωμάτων, υποστηρίζοντας τη συνεχή βελτίωση και τον περιορισμό των προκαταλήψεων στην κλινική εφαρμογή της Τεχνητής Νοημοσύνης.
- **Συνεργασία σε επίπεδο δικτύου και ανταλλαγή πληροφοριών σχετικά με περιστατικά.** Το ευρωπαϊκό δίκτυο των οργανώσεων της κοινωνίας των πολιτών και δημόσιων νοσοκομείων που θα αναπτυχθεί στο πλαίσιο του AEQUITAS λειτουργεί ως ένα κοινό επίπεδο διακυβέρνησης, παρέχοντας ενιαίο λεξιλόγιο και κοινά πρότυπα αναφοράς για περιστατικά και ευρήματα που σχετίζονται με έμφυλες και φυλετικές προκαταλήψεις στα συστήματα Τεχνητής Νοημοσύνης. Το δίκτυο αυτό υποστηρίζει τη διατοπική μάθηση και τη διάχυση βέλτιστων πρακτικών για τον εντοπισμό, τον περιορισμό και την αποκατάσταση των ανισοτήτων, κατά τρόπο ανάλογο με προσεγγίσεις που εφαρμόζονται σε τομείς υψηλής ασφάλειας και στη φαρμακοεπαγρύπνηση. Μέσω του διαμοιρασμού

⁴⁴ European Medicines Agency, [Ενότητα VI του GVP – Συλλογή και υποβολή αναφορών για ύποπτες ανεπιθύμητες ενέργειες \(Αναθ. 2\)](#), London: EMA, 2017.

δεδομένων που αφορούν σε καταγεγραμμένα περιστατικά προκαταλήψεων, σε διαστρωματωμένες αναλύσεις για την επίδοση των συστημάτων, καθώς και σε εφαρμοσμένες στρατηγικές περιορισμού των κινδύνων, το δίκτυο επιτρέπει σε νοσοκομεία και οργανώσεις της κοινωνίας των πολιτών να εντοπίζουν έγκαιρα συστημικά πρότυπα ανισότητας και να συντονίζουν στοχευμένες παρεμβάσεις. Εφόσον ο Ευρωπαϊκός Χώρος Δεδομένων Υγείας (*European Health Data Space/ EHDS*), στο πλαίσιο μιας ρυθμιστικής διαδικασίας, διευκολύνει τόσο την πρωτογενή χρήση (παροχή φροντίδας) όσο και τη δευτερογενή χρήση (έρευνα και χάραξη πολιτικής), επιτρέπει στο δίκτυο να ευθυγραμμίσει τις πρακτικές ανταλλαγής δεδομένων και μάθησης με τα χρονοδιαγράμματα και τις εκτελεστικές πράξεις του EHDS καθώς αυτά εξελίσσονται, υποστηρίζοντας μια πιο δίκαιη και διαφανή εφαρμογή της Τεχνητής Νοημοσύνης σε όλα τα ιδρύματα υγειονομικής περίθαλψης.

- **Εκπαιδευτική ενότητα/Εκπαιδευτικό σεμινάριο για κλινικούς ιατρούς και εμπειρογνώμονες των οργανώσεων της κοινωνίας των πολιτών.** Προκειμένου να περιοριστεί η πιθανότητα της «προκατάληψης αυτοματισμού» (*automation bias*) και των ανισοτήτων στην κλινική Τεχνητή Νοημοσύνη, το ρυθμιστικό μοντέλο AEQUITAS περιλαμβάνει υποχρεωτική εκπαίδευση, σχεδιασμένη ώστε να εφοδιάζει τις/τους επαγγελματίες υγείας με δεξιότητες κατανόησης των περιορισμών των συστημάτων ΤΝ, κριτικής ερμηνείας των αποτελεσμάτων τους και δυνατότητας παράκαμψης ή διακοπής της λειτουργίας τους όταν τα αποτελέσματα παράγουν προκατειλημμένα ή άνισα αποτελέσματα. Η εκπαίδευση δίνει έμφαση στην ευαισθητοποίηση σχετικά με τους κινδύνους εμφάνισης έμφυλων και φυλετικών προκαταλήψεων, στις αρχές των θεμελιωδών δικαιωμάτων και σε ηθικά ζητήματα, συμπεριλαμβανομένης της ισότητας, της μη διάκρισης και της συμπερίληψης των ασθενών. Η ενότητα αυτή μπορεί να θεωρηθεί ως ένα υποστηρικτικό εργαλείο, το οποίο ενισχύει ενεργά την εφαρμογή του ρυθμιστικού μοντέλου, διασφαλίζοντας ότι το προσωπικό που είναι υπεύθυνο για την εποπτεία συστημάτων ΤΝ μπορεί να εφαρμόζει με επάρκεια διαδικασίες παρακολούθησης, επικύρωσης και λογοδοσίας μέσα από το πρίσμα της ισότητας. Η εκπαίδευση περιλαμβάνει επίσης ασκήσεις βασισμένες σε συγκεκριμένες μελέτες περίπτωσης για τον εντοπισμό και τον μετριασμό ανισοτήτων με βάση το βιολογικό φύλο, την ταυτότητα φύλου και τις φυλετικές/εθνοτικές διαφορές.

Συμπερασματικά

Η Τεχνητή Νοημοσύνη προσφέρει σημαντικές δυνατότητες για τη βελτίωση των αποτελεσμάτων στην υγειονομική περίθαλψη. Ωστόσο, χωρίς προσεκτική διακυβέρνηση μπορεί να αναπαράγει ή να επιτείνει ανισότητες λόγω φύλου και ταυτότητας φύλου, καθώς και φυλετικές ή εθνοτικές ανισότητες, αλλά και άλλες

μορφές συστημικών προκαταλήψεων. Το ρυθμιστικό μοντέλο που παρουσιάζεται στο παρόν Έγγραφο Πολιτικής επιχειρεί να μετασχηματίσει τις αρχές των θεμελιωδών δικαιωμάτων σε συγκεκριμένες δικλείδες ασφαλείας που θα πρέπει να ισχύουν σε όλα τα στάδια του κύκλου ζωής των συστημάτων ΤΝ, από την προμήθεια και την επικύρωση έως την ανάπτυξη, την παρακολούθηση και τη λογοδοσία, με ρητή έμφαση στον εντοπισμό και τον περιορισμό των διαφορών στα αποτελέσματα που εμφανίζονται μεταξύ διαφορετικών πληθυσμιακών ομάδων ασθενών. Μέσω της εφαρμογής ενός συστήματος διακυβέρνησης συστημάτων ΤΝ που στηρίζεται σε ένα δομημένο σύνολο σταδίων κύκλου ζωής, πολιτικών, πρακτικών εργαλείων και τεκμηρίων, τα νοσοκομεία και οι οργανώσεις της κοινωνίας των πολιτών μπορούν να αξιολογούν συστηματικά, να μετριάσουν και να τεκμηριώνουν κινδύνους, συμπεριλαμβανομένων των άνισων επιπτώσεων σε κοινωνικά αποκλεισμένες ομάδες ασθενών.

Με την ενσωμάτωση διδαγμάτων από διαφορετικούς τομείς, των νομικών απαραίτητων προϋποθέσεων της ΕΕ και διεθνών πλαισίων διακυβέρνησης, το ρυθμιστικό μοντέλο υποστηρίζει τις/ τους επαγγελματίες υγείας στην υπεύθυνη χρήση συστημάτων ΤΝ και διασφαλίζει δίκαιη, ασφαλή και διαφανή υγειονομική περίθαλψη για όλες/όλους τις/τους ασθενείς, ανεξαρτήτως φύλου, ταυτότητας φύλου, φυλετικής/εθνοτικής καταγωγής ή άλλων δημογραφικών χαρακτηριστικών.