

ΑΝΕΞΑΡΤΗΤΗ

ΟΜΑΔΑ ΕΜΠΕΙΡΟΓΝΩΜΟΝΩΝ ΥΨΗΛΟΥ

ΕΠΙΠΕΔΟΥ

ΓΙΑ ΤΗΝ ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ

ΣΥΣΤΑΘΕΙΣΑ ΑΠΟ ΤΗΝ ΕΥΡΩΠΑΪΚΗ ΕΠΙΤΡΟΠΗ ΤΟΝ ΙΟΥΝΙΟ ΤΟΥ 2018



ΚΑΤΕΥΘΥΝΤΗΡΙΕΣ ΓΡΑΜΜΕΣ

ΔΕΟΝΤΟΛΟΓΙΑΣ

ΓΙΑ ΑΞΙΟΠΙΣΤΗ ΤΕΧΝΗΤΗ

ΝΟΗΜΟΣΥΝΗ

# ΚΑΤΕΥΘΥΝΤΗΡΙΕΣ ΓΡΑΜΜΕΣ ΔΕΟΝΤΟΛΟΓΙΑΣ ΓΙΑ ΑΞΙΟΠΙΣΤΗ ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ

Ομάδα εμπειρογνομόνων υψηλού επιπέδου για την τεχνητή νοημοσύνη

Το παρόν έγγραφο εκπονήθηκε από την ομάδα εμπειρογνομόνων υψηλού επιπέδου για την ΤΝ (ΟΕΥΕ για την ΤΝ). Τα μέλη της ΟΕΥΕ για την ΤΝ που κατονομάζονται στο παρόν έγγραφο στηρίζουν το συνολικό πλαίσιο για αξιόπιστη ΤΝ που προτείνεται στις παρούσες κατευθυντήριες γραμμές, χωρίς αυτό να σημαίνει κατ' ανάγκη ότι συμφωνούν με κάθε δήλωση που διατυπώνεται στο έγγραφο.

Ο κατάλογος αξιολόγησης της αξιοπιστίας της ΤΝ που παρουσιάζεται στο κεφάλαιο III του παρόντος εγγράφου θα εφαρμοστεί πιλοτικά από ενδιαφερόμενα μέρη με στόχο τη συλλογή πρακτικής αναπληροφόρησης. Αναθεωρημένη έκδοση του καταλόγου αξιοπιστίας, στην οποία θα λαμβάνεται υπόψη η αναπληροφόρηση που θα συλλεχθεί στο πλαίσιο της πιλοτικής φάσης, θα υποβληθεί στην Ευρωπαϊκή Επιτροπή στις αρχές του 2020.

Η ομάδα εμπειρογνομόνων υψηλού επιπέδου για την τεχνητή νοημοσύνη (ΟΕΥΕ για την ΤΝ) είναι μια ανεξάρτητη ομάδα εμπειρογνομόνων που συστάθηκε από την Ευρωπαϊκή Επιτροπή τον Ιούνιο του 2018.

Επικοινωνία Nathalie Smuha - Συντονίστρια της ΟΕΥΕ για την ΤΝ  
E-mail CNECT-HLG-AI@ec.europa.eu

European Commission  
B-1049 Brussels

Το έγγραφο δημοσιεύτηκε στις 8 Απριλίου 2019.

Το πρώτο σχέδιο του παρόντος εγγράφου δημοσιεύτηκε στις 18 Δεκεμβρίου 2018 και τέθηκε σε ανοιχτή διαβούλευση στο πλαίσιο της οποίας παρείχαν αναπληροφόρηση περισσότεροι από 500 συμμετέχοντες. Θα θέλαμε να ευχαριστήσουμε ιδιαίτερα και θερμώς όσους διατύπωσαν τις απόψεις τους επί του πρώτου σχεδίου του εγγράφου, οι οποίες ελήφθησαν υπόψη κατά την εκπόνηση της παρούσας αναθεωρημένης έκδοσης.

Ούτε η Ευρωπαϊκή Επιτροπή ούτε οποιοδήποτε άλλο πρόσωπο ενεργεί εξ ονόματός της φέρει ευθύνη για την ενδεχόμενη χρήση των κάτωθι πληροφοριών. Για τα περιεχόμενα του παρόντος εγγράφου εργασίας αποκλειστικά υπεύθυνη είναι η ομάδα εμπειρογνομόνων υψηλού επιπέδου για την τεχνητή νοημοσύνη (ΟΕΥΕ για την ΤΝ). Παρόλο που μέλη του προσωπικού της Επιτροπής διευκόλυναν την εκπόνηση των κατευθυντήριων γραμμών, οι απόψεις που διατυπώνονται

στο παρόν έγγραφο εκφράζουν την άποψη της ΟΕΥΕ για την ΤΝ και δεν μπορεί σε καμία περίπτωση να εκληφθούν ως επίσημη θέση της Ευρωπαϊκής Επιτροπής.

Περισσότερες πληροφορίες σχετικά με την ομάδα εμπειρογνομόνων υψηλού επιπέδου για την τεχνητή νοημοσύνη διατίθενται στο διαδίκτυο (<https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>). Η πολιτική για την περαιτέρω χρήση εγγράφων της Ευρωπαϊκής Επιτροπής καθορίζεται στην απόφαση 2011/833/ΕΕ (ΕΕ L 330 της 14.12.2011, σ. 39). Για κάθε χρήση ή αναπαραγωγή φωτογραφιών ή άλλου υλικού που δεν υπόκειται στους κανόνες της ΕΕ για τα δικαιώματα πνευματικής ιδιοκτησίας πρέπει να ζητείται άδεια απευθείας από τους κατόχους δικαιωμάτων πνευματικής ιδιοκτησίας.

## ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

<b>ΣΥΝΟΠΤΙΚΗ ΠΑΡΟΥΣΙΑΣΗ</b>	<b>2</b>			
<b>A.</b>	<b>ΕΙΣΑΓΩΓΗ</b>			
	<b>5</b>			
<b>Β.ΈΝΑ</b>	<b>ΠΛΑΙΣΙΟ</b>	<b>ΓΙΑ</b>	<b>ΑΞΙΟΠΙΣΤΗ</b>	<b>ΤΝ</b>
				<b>8</b>
<b>I. Κεφάλαιο I: Βάσεις της αξιόπιστης ΤΝ</b>				<b>11</b>
1. Τα θεμελιώδη δικαιώματα ως ηθικά και νομικά δικαιώματα				12
2. Από τα θεμελιώδη δικαιώματα στις δεοντολογικές αρχές				13
<b>II. Κεφάλαιο II: Πραγμάτωση της αξιόπιστης ΤΝ</b>				<b>17</b>
1. Απαιτήσεις της αξιόπιστης ΤΝ				17
2. Τεχνικές και μη τεχνικές μέθοδοι για την πραγμάτωση της αξιόπιστης ΤΝ				26
<b>III. Κεφάλαιο III: Αξιολόγηση της αξιοπιστίας της τεχνητής νοημοσύνης</b>				<b>31</b>
<b>Γ.ΠΑΡΑΔΕΙΓΜΑΤΑ ΕΥΚΑΙΡΙΩΝ ΚΑΙ ΚΑΙΡΙΩΝ ΠΡΟΒΛΗΜΑΤΙΣΜΩΝ ΠΟΥ ΕΓΕΙΡΕΙ Η ΤΝ</b>				<b>44</b>
<b>Δ.</b>				<b>ΣΥΜΠΕΡΑΣΜΑ</b>
				<b>48</b>
<b>ΓΛΩΣΣΑΡΙΟ</b>				<b>50</b>

## **ΣΥΝΟΠΤΙΚΗ ΠΑΡΟΥΣΙΑΣΗ**

- (1) Στόχος των κατευθυντήριων γραμμών είναι η προαγωγή της αξιοπιστίας της τεχνητής νοημοσύνης (TN). Η αξιόπιστη TN έχει **τρεις συνιστώσες**, οι οποίες θα πρέπει να πληρούνται σε ολόκληρο τον κύκλο ζωής του συστήματος: α) θα πρέπει να είναι **σύννομη**, ήτοι να τηρεί όλες τις εφαρμοστέες νομοθετικές και κανονιστικές διατάξεις, β) θα πρέπει να είναι **δεοντολογική**, ήτοι να διασφαλίζει τη συμμόρφωση με δεοντολογικές αρχές και αξίες και γ) θα πρέπει να είναι **στιβαρή**, τόσο από τεχνικής όσο και από κοινωνικής άποψης, διότι, ακόμη κι όταν υπάρχει καλή πρόθεση, τα συστήματα TN μπορούν να προκαλέσουν ακούσια βλάβη. Κάθε συνιστώσα από μόνη της αποτελεί αναγκαία αλλά όχι ικανή συνθήκη για την επίτευξη αξιόπιστης TN. Υπό ιδανικές συνθήκες, και οι τρεις συνιστώσες συνεργάζονται αρμονικά και αλληλεπικαλύπτονται στη λειτουργία τους. Εάν, στην πράξη, προκύψουν τριβές μεταξύ αυτών των συνιστωσών, η κοινωνία οφείλει να επιδιώκει την ευθυγράμμισή τους.
- (2) Με τις παρούσες κατευθυντήριες γραμμές θεσπίζεται ένα **πλαίσιο για την επίτευξη αξιόπιστης TN**. Το πλαίσιο δεν πραγματεύεται ρητά την πρώτη συνιστώσα της αξιόπιστης TN (σύννομη TN)<sup>1</sup>. Στόχος του είναι, αντ' αυτού, να παρέχει κατευθύνσεις όσον αφορά την προαγωγή και τη διασφάλιση δεοντολογικής και στιβαρής TN (της δεύτερης και τρίτης συνιστώσας). Οι παρούσες κατευθυντήριες γραμμές, οι οποίες απευθύνονται σε όλα τα ενδιαφερόμενα μέρη, δεν περιορίζονται σε έναν κατάλογο δεοντολογικών αρχών, αλλά παρέχουν κατευθύνσεις σχετικά με τους τρόπους με τους οποίους οι αρχές αυτές είναι δυνατό να υλοποιηθούν επιχειρησιακά στο πλαίσιο κοινωνικοτεχνικών συστημάτων. Οι κατευθύνσεις παρέχονται σε τρία επίπεδα, από τα πιο αφηρημένα στο κεφάλαιο I έως τα πιο συγκεκριμένα στο κεφάλαιο III, και ολοκληρώνονται με παραδείγματα ευκαιριών και καίριων προβληματισμών που εγείρουν τα συστήματα TN.
- I. Βάσει προσέγγισης που εδράζεται στα ανθρώπινα δικαιώματα, στο κεφάλαιο I προσδιορίζονται οι **δεοντολογικές αρχές** και οι συνυφασμένες με αυτές αξίες που πρέπει να τηρούνται κατά την ανάπτυξη, την εγκατάσταση και τη χρήση συστημάτων TN.

### **Βασικές κατευθύνσεις που προκύπτουν από το κεφάλαιο I:**

- ✓ Ανάπτυξη, εγκατάσταση και χρήση των συστημάτων TN με τρόπους ώστε να τηρούνται οι ακόλουθες δεοντολογικές αρχές: *σεβασμός της ανθρώπινης αυτονομίας, πρόληψη βλάβης, δικαιοσύνη και επεξηγησιμότητα*. Αναγνώριση και αντιμετώπιση των πιθανών τριβών μεταξύ αυτών των αρχών.
- ✓ Απόδοση ιδιαίτερης προσοχής σε καταστάσεις που αφορούν πιο ευάλωτες ομάδες όπως τα παιδιά, τα άτομα με αναπηρίες και άλλες ομάδες που βρίσκονται ανέκαθεν σε μειονεκτική θέση, ή αντιμετωπίζουν κίνδυνο αποκλεισμού, καθώς και σε καταστάσεις που χαρακτηρίζονται από ασυμμετρίες εξουσίας ή πληροφόρησης, όπως μεταξύ εργοδοτών και εργαζομένων ή επιχειρήσεων και καταναλωτών<sup>2</sup>.
- ✓ Αναγνώριση και συνεκτίμηση του γεγονότος ότι τα συστήματα TN, αν και αποφέρουν ουσιαστικά οφέλη για τους πολίτες και την κοινωνία, ενέχουν επίσης ορισμένους κινδύνους και μπορεί να έχουν αρνητικές επιπτώσεις, συμπεριλαμβανομένων επιπτώσεων που μπορεί να είναι δύσκολο να προβλεφθούν, να προσδιοριστούν ή να μετρηθούν (π.χ. στη δημοκρατία, στο κράτος δικαίου και στη διανεμητική δικαιοσύνη ή και στον ίδιο τον ανθρώπινο νου.) Λήψη επαρκών μέτρων για τον μετριασμό των εν λόγω κινδύνων, όταν κρίνεται σκόπιμο, και ανάλογα με το μέγεθος του κινδύνου.

<sup>1</sup> Στόχος όλων των κανονιστικών διατυπώσεων που περιλαμβάνονται στο παρόν έγγραφο είναι η παρουσίαση κατευθύνσεων για την επίτευξη της δεύτερης και της τρίτης συνιστώσας της αξιόπιστης TN (δεοντολογική και στιβαρή TN). Ως εκ τούτου, οι εν λόγω διατυπώσεις δεν έχουν στόχο την παροχή νομικών συμβουλών ή κατευθύνσεων σχετικά με τη συμμόρφωση με εφαρμοστέες νομοθετικές διατάξεις, αν και αναγνωρίζεται ότι πολλές από τις διατυπώσεις αυτές αποτυπώνονται ήδη σε ορισμένο βαθμό σε ισχύουσες νομοθετικές διατάξεις. Βλ. σχετικά παράγραφο 21 και εξής.

<sup>2</sup> Βλ. άρθρα 24 έως 27 του Χάρτη των Θεμελιωδών Δικαιωμάτων της ΕΕ (Χάρτης της ΕΕ), σχετικά με τα δικαιώματα του παιδιού και των ηλικιωμένων, την ένταξη των ατόμων με αναπηρία και τα δικαιώματα των εργαζομένων. Βλ. επίσης άρθρο 38 σχετικά με την προστασία του καταναλωτή.

- II. Με βάση το κεφάλαιο I, στο κεφάλαιο II παρέχονται κατευθύνσεις σχετικά με τους τρόπους με τους οποίους η αξιόπιστη ΤΝ μπορεί να πραγματοποιηθεί, μέσω της παράθεσης **επτά απαιτήσεων** τις οποίες θα πρέπει να τηρούν τα συστήματα ΤΝ. Για την εφαρμογή τους μπορούν να χρησιμοποιούνται τόσο τεχνικές όσο και μη τεχνικές μέθοδοι.

**Βασικές κατευθύνσεις που προκύπτουν από το κεφάλαιο II:**

- ✓ Διασφάλιση ότι η ανάπτυξη, εγκατάσταση και χρήση των συστημάτων ΤΝ πληροί τις απαιτήσεις για αξιόπιστη ΤΝ: 1) ανθρώπινη παρέμβαση και εποπτεία, 2) τεχνική στιβαρότητα και ασφάλεια, 3) ιδιωτική ζωή και διακυβέρνηση των δεδομένων, 4) διαφάνεια, 5) πολυμορφία, απαγόρευση των διακρίσεων και δικαιοσύνη, 6) περιβαλλοντική και κοινωνιακή ευημερία και 7) λογοδοσία.
- ✓ Εξέταση τεχνικών και μη τεχνικών μεθόδων για την εξασφάλιση της εφαρμογής των εν λόγω απαιτήσεων.
- ✓ Προώθηση της έρευνας και της καινοτομίας ώστε να συμβάλλουν στην αξιολόγηση των συστημάτων ΤΝ και να προάγουν την επίτευξη των απαιτήσεων· διάδοση των αποτελεσμάτων και άνοιγμα των ερωτημάτων στο ευρύτερο κοινό, καθώς και συστηματική εκπαίδευση μιας νέας γενιάς εμπειρογνομόνων στη δεοντολογία της ΤΝ.
- ✓ Ενημέρωση, με σαφή και ενεργό τρόπο, των ενδιαφερόμενων μερών σχετικά με τις δυνατότητες και τους περιορισμούς του συστήματος ΤΝ, ώστε να δημιουργούνται ρεαλιστικές προσδοκίες, καθώς και σχετικά με τον τρόπο με τον οποίο εφαρμόζονται οι απαιτήσεις. Διαφάνεια έναντι των εν λόγω ενδιαφερόμενων μερών ως προς το γεγονός ότι έχουν να κάνουν με ένα σύστημα ΤΝ.
- ✓ Διευκόλυνση της ιχνηλασιμότητας και της ελεγχιμότητας των συστημάτων ΤΝ, ιδίως σε πλαίσια ή καταστάσεις κρίσιμης σημασίας.
- ✓ Συμμετοχή των ενδιαφερόμενων μερών καθόλη τη διάρκεια του κύκλου ζωής του συστήματος ΤΝ. Ενίσχυση της κατάρτισης και της εκπαίδευσης, έτσι ώστε όλα τα ενδιαφερόμενα μέρη να είναι κατατοπισμένα και εκπαιδευμένα σε σχέση με την αξιόπιστη ΤΝ.
- ✓ Επίγνωση του γεγονότος ότι μπορεί να υπάρχουν θεμελιώδεις τριβές μεταξύ διαφόρων αρχών και απαιτήσεων. Διαρκής προσδιορισμός, αξιολόγηση, τεκμηρίωση και γνωστοποίηση των εν λόγω τριβών και των λύσεών τους με τη μορφή αντισταθμιστικών ρυθμίσεων.

- III. Στο κεφάλαιο III παρατίθεται συγκεκριμένος και μη εξαντλητικός κατάλογος αξιολόγησης της αξιοπιστίας της ΤΝ, στόχος του οποίου είναι η επιχειρησιακή υλοποίηση των απαιτήσεων που εκτίθενται στο κεφάλαιο II. Ο εν λόγω **κατάλογος αξιολόγησης** θα πρέπει να προσαρμόζεται στη συγκεκριμένη περίπτωση χρήσης του συστήματος ΤΝ<sup>3</sup>.

**Βασικές κατευθύνσεις που προκύπτουν από το κεφάλαιο III:**

- ✓ Έκδοση καταλόγου αξιολόγησης της αξιοπιστίας της ΤΝ κατά την ανάπτυξη, εγκατάσταση και χρήση συστημάτων ΤΝ, και προσαρμογή του εν λόγω καταλόγου στη συγκεκριμένη περίπτωση χρήσης του συστήματος.
- ✓ Σημειωτέον ότι ο κατάλογος αξιολόγησης δεν θα είναι ποτέ εξαντλητικός. Η εξασφάλιση αξιόπιστης ΤΝ δεν συνίσταται σε συμπλήρωση τετραγωνιδίων, αλλά σε διαρκή διαδικασία προσδιορισμού και εφαρμογής απαιτήσεων, αξιολόγησης λύσεων και διασφάλισης βελτιωμένων αποτελεσμάτων σε

<sup>3</sup> Σύμφωνα με το πεδίο εφαρμογής του πλαισίου που ορίζεται στην παράγραφο 2, ο εν λόγω κατάλογος αξιολόγησης δεν παρέχει συμβουλές για τη διασφάλιση νομικής συμμόρφωσης (σύννομη ΤΝ) αλλά περιορίζεται στην παροχή κατευθύνσεων σχετικά με την εκπλήρωση της δεύτερης και της τρίτης συνιστώσας της αξιόπιστης ΤΝ (δεοντολογική και στιβαρή ΤΝ).

ολόκληρο τον κύκλο ζωής του συστήματος TN, καθώς και εξασφάλισης της συμμετοχής των ενδιαφερόμενων μερών στην όλη διαδικασία.

- (3) Στόχος της τελευταίας ενότητας του εγγράφου είναι η εξειδίκευση μερικών από τα ζητήματα που τίγονται σε διάφορα σημεία του πλαισίου, μέσω της παρουσίασης παραδειγμάτων χρήσιμων ευκαιριών, οι οποίες θα πρέπει να αξιοποιηθούν, καθώς και καίριων προβληματισμών που εγείρουν τα συστήματα TN, οι οποίοι θα πρέπει να εξεταστούν προσεκτικά.
- (4) Παρότι στόχος αυτών των κατευθυντήριων γραμμών είναι η παροχή κατευθύνσεων για τις εφαρμογές TN εν γένει, ώστε να δημιουργηθεί μια οριζόντια βάση για την επίτευξη αξιόπιστης TN, διαφορετικές καταστάσεις παρουσιάζουν διαφορετικές προκλήσεις. Θα πρέπει επομένως να διερευνηθεί κατά πόσο, επιπρόσθετα σε αυτό το οριζόντιο πλαίσιο, απαιτείται μια τομεακή προσέγγιση, δεδομένων των ιδιαιτεροτήτων των συστημάτων TN ανάλογα με το εκάστοτε πλαίσιο.
- (5) Οι παρούσες κατευθυντήριες γραμμές δεν έχουν σκοπό να υποκαταστήσουν οποιαδήποτε μορφή υφιστάμενης ή μελλοντικής διαδικασίας χάραξης πολιτικής ή κανονιστικής ρύθμισης, ούτε έχουν στόχο να αποτρέψουν τη θέσπισή τους. Θα πρέπει να αντιμετωπίζονται ως ένα «ζωντανό» κείμενο, το οποίο θα αναθεωρείται και θα επιθεωρείται με την πάροδο του χρόνου ώστε να διασφαλίζεται ότι θα παραμένει διαρκώς συναφές, καθώς η τεχνολογία, τα κοινωνικά μας περιβάλλοντα και οι γνώσεις μας θα εξελίσσονται. Το παρόν έγγραφο σχεδιάστηκε ως εφελτήριο για την έναρξη του διαλόγου σχετικά με την «Αξιόπιστη TN για την Ευρώπη»<sup>4</sup>. Οι κατευθυντήριες γραμμές αποσκοπούν επίσης στη δημιουργία ευνοϊκών συνθηκών για την έρευνα, την ανάπτυξη προβληματισμού και τη διεξαγωγή διαλόγου με αντικείμενο ένα πλαίσιο δεοντολογίας για τα συστήματα TN, πέρα από την Ευρώπη, σε παγκόσμιο επίπεδο.

---

<sup>4</sup> Το ιδεώδες αυτό προορίζεται να εφαρμόζεται σε συστήματα TN που αναπτύσσονται, εγκαθίστανται και χρησιμοποιούνται στα κράτη μέλη της ΕΕ, καθώς και σε συστήματα που αναπτύσσονται ή παράγονται αλλού, αλλά εγκαθίστανται και χρησιμοποιούνται στην ΕΕ. Όταν στο παρόν έγγραφο γίνεται αναφορά στην «Ευρώπη», η αναφορά περιλαμβάνει τα κράτη μέλη της ΕΕ. Ωστόσο, φιλοδοξία αυτών των κατευθυντήριων γραμμών είναι να έχουν συνάφεια και εκτός ΕΕ. Στο πλαίσιο αυτό, επισημαίνεται ότι τόσο η Νορβηγία όσο και η Ελβετία συμμετέχουν στο συντονισμένο σχέδιο για την TN, το οποίο συμφωνήθηκε και δημοσιεύτηκε τον Δεκέμβριο του 2018 από την Επιτροπή και τα κράτη μέλη.

## A. ΕΙΣΑΓΩΓΗ

- (6) Η Ευρωπαϊκή Επιτροπή («η Επιτροπή»), στις ανακοινώσεις που εξέδωσε στις 25 Απριλίου 2018 και 7 Δεκεμβρίου 2018, εκθέτει το όραμά της για την Τεχνητή Νοημοσύνη (TN), το οποίο υποστηρίζει μια «δεοντολογική, ασφαλή και προηγμένη TN με τη σφραγίδα της Ευρώπης»<sup>5</sup>. Το όραμα της Επιτροπής βασίζεται σε τρεις πυλώνες: i) αύξηση των δημόσιων και ιδιωτικών επενδύσεων στον τομέα της TN με στόχο να δοθεί ώθηση στη διεύρυσή της, ii) προετοιμασία για τις κοινωνικοοικονομικές αλλαγές και iii) διασφάλιση κατάλληλου δεοντολογικού και νομικού πλαισίου με στόχο την ενίσχυση των ευρωπαϊκών αρχών.
- (7) Προς υποστήριξη της υλοποίησης αυτού του οράματος, η Επιτροπή συνέστησε την ομάδα εμπειρογνομόνων υψηλού επιπέδου για την τεχνητή νοημοσύνη (ΟΕΥΕ για την TN), μια ανεξάρτητη ομάδα επιφορτισμένη με την εκπόνηση δύο παραδοτέων: 1) των κατευθυντήριων γραμμών δεοντολογίας για την TN και 2) συστάσεων πολιτικής και επενδύσεων.
- (8) Το παρόν έγγραφο περιέχει τις κατευθυντήριες γραμμές δεοντολογίας για την TN, οι οποίες αναθεωρήθηκαν κατόπιν περαιτέρω επεξεργασίας από την ομάδα μας, λαμβανομένης υπόψη της αναπληροφόρησης που ελήφθη στο πλαίσιο της δημόσιας διαβούλευσης επί του σχεδίου που δημοσιεύτηκε στις 18 Δεκεμβρίου 2018. Βασίζεται στο έργο της ευρωπαϊκής ομάδας για τη δεοντολογία της επιστήμης και των νέων τεχνολογιών<sup>6</sup> και αντλεί έμπνευση από άλλες παρόμοιες προσπάθειες<sup>7</sup>.
- (9) Τους τελευταίους μήνες, τα 52 μέλη της ομάδας μας συναντήθηκαν, συζήτησαν και αντάλλαξαν απόψεις, όπως προτάσσει το σύνθημα της Ευρωπαϊκής Ένωσης: ενωμένοι στην πολυμορφία. Πιστεύουμε ότι η TN έχει τις δυνατότητες να μεταμορφώσει ριζικά την κοινωνία. Η TN δεν αποτελεί αυτοσκοπό, αλλά ένα πολλά υποσχόμενο μέσο για την ανθρώπινη ευημερία, που μπορεί να προάγει την ατομική και κοινωνιακή ευζωία και το κοινό καλό, καθώς και να συντελέσει στην πρόοδο και την καινοτομία. Ειδικότερα, τα συστήματα TN μπορούν να συμβάλλουν στη διευκόλυνση της επίτευξης των στόχων βιώσιμης ανάπτυξης των Ηνωμένων Εθνών, όπως είναι η προαγωγή της ισόρροπης εκπροσώπησης των φύλων και η αντιμετώπιση της κλιματικής αλλαγής, ο εξορθολογισμός της χρήσης των φυσικών πόρων, η βελτίωση των διεργασιών μας όσον αφορά την υγεία, την κινητικότητα και την παραγωγή, και η ενίσχυση της παρακολούθησης της προόδου έναντι των δεικτών βιωσιμότητας και κοινωνικής συνοχής.
- (10) Για να συμβεί αυτό, τα συστήματα TN<sup>8</sup> πρέπει να είναι **ανθρωποκεντρικά**, εδραζόμενα στη δέσμευση ότι η χρήση τους τίθεται στην υπηρεσία της ανθρωπότητας και του κοινού καλού, με στόχο τη βελτίωση της ευημερίας και της ελευθερίας του ανθρώπου. Παράλληλα με τις μεγάλες ευκαιρίες που προσφέρουν, τα συστήματα TN ενέχουν και ορισμένους κινδύνους που θα πρέπει να αντιμετωπιστούν με κατάλληλο και αναλογικό τρόπο. Σήμερα μας δίνεται μια σημαντική ευκαιρία να διαμορφώσουμε την ανάπτυξή τους. Θέλουμε να διασφαλίσουμε ότι μπορούμε να εμπιστευόμαστε τα κοινωνικοτεχνικά περιβάλλοντα στα οποία είναι ενσωματωμένα και θέλουμε οι κατασκευαστές των συστημάτων TN να εξασφαλίζουν ανταγωνιστικό πλεονέκτημα όταν ενσωματώνουν αξιόπιστη TN στα προϊόντα και τις υπηρεσίες τους. Προϋπόθεση για να συμβεί αυτό είναι να επιδιώκεται, αφενός, **ημεγιστοποίηση των οφελών των συστημάτων TN** και, αφετέρου,

<sup>5</sup> COM(2018)237 και COM(2018)795. Επισημαίνεται ότι ο όρος «με τη σφραγίδα της Ευρώπης» χρησιμοποιείται σε ολόκληρο το κείμενο της ανακοίνωσης της Επιτροπής. Ωστόσο, οι παρούσες κατευθύνσεις επιδιώκουν να καλύπτουν όχι μόνο τα συστήματα TN «με τη σφραγίδα της Ευρώπης» αλλά και τα συστήματα που αναπτύσσονται αλλού και εγκαθίστανται ή χρησιμοποιούνται στην Ευρώπη. Ως εκ τούτου, σε ολόκληρο το παρόν κείμενο στόχος μας είναι η προαγωγή αξιόπιστης TN «για» την Ευρώπη.

<sup>6</sup> Η ευρωπαϊκή ομάδα για τη δεοντολογία της επιστήμης και των νέων τεχνολογιών είναι συμβουλευτική ομάδα της Επιτροπής.

<sup>7</sup> Βλ. τμήμα 3.3 της ανακοίνωσης COM(2018)237.

<sup>8</sup> Το γλωσσάριο στο τέλος του παρόντος εγγράφου περιλαμβάνει ορισμό των συστημάτων TN για τους σκοπούς του παρόντος εγγράφου. Ο ορισμός αυτός εξειδικεύεται περαιτέρω σε ειδικό έγγραφο που έχει συντάξει η ΟΕΥΕ για την TN, το οποίο συνοδεύει τις παρούσες κατευθυντήριες γραμμές, με τίτλο «Ορισμός της τεχνητής νοημοσύνης: κύριες δυνατότητες και επιστημονικοί κλάδοι».



## η πρόληψη και ελαχιστοποίηση των κινδύνων τους.

- (11) Σε ένα περιβάλλον ραγδαίας τεχνολογικής εξέλιξης, πιστεύουμε ότι είναι αναγκαίο η εμπιστοσύνη να εξακολουθήσει να αποτελεί τη συγκολλητική ουσία των κοινωνιών, των κοινοτήτων, των οικονομιών και της βιώσιμης ανάπτυξης. Ως εκ τούτου, θεωρούμε την **αξιόπιστη TN θεμελιώδη φιλοδοξία μας**, καθώς, για να μπορούν άνθρωποι και κοινότητες να έχουν εμπιστοσύνη στην εξέλιξη της τεχνολογίας και στις εφαρμογές της, είναι απαραίτητο να υπάρχει ένα σαφές και ολοκληρωμένο πλαίσιο για την επίτευξη της αξιοπιστίας της.
- (12) Αυτή είναι η πορεία που θεωρούμε ότι θα πρέπει να ακολουθήσει η Ευρώπη ώστε να τεθεί στο επίκεντρο και να ηγηθεί του τομέα της προηγμένης και δεοντολογικής TN. Μέσω της αξιόπιστης TN, εμείς, ως Ευρωπαίοι πολίτες, θα επιδιώκουμε να αποκομίζουμε τα οφέλη της με τρόπο που συνάδει με τις θεμελιώδεις μας αξίες του σεβασμού των δικαιωμάτων του ανθρώπου, της δημοκρατίας και του κράτους δικαίου.

### Αξιόπιστη TN

- (13) Η αξιοπιστία αποτελεί προϋπόθεση για την ανάπτυξη, εγκατάσταση και χρήση συστημάτων TN από τον άνθρωπο και τις κοινωνίες. Εάν τα συστήματα TN —και οι άνθρωποι πίσω από αυτά— δεν είναι σε θέση να αποδείξουν ότι είναι άξια εμπιστοσύνης, μπορεί να προκύψουν ανεπιθύμητες συνέπειες και η διείσδυσή τους ενδέχεται να παρεμποδιστεί, με αποτέλεσμα να μην καταστεί δυνατή η αξιοποίηση των δυνητικά τεράστιων κοινωνικών και οικονομικών οφελών που επιφέρουν τα συστήματα TN. Για να μπορέσει η Ευρώπη να αξιοποιήσει αυτά τα οφέλη, όραμά μας είναι να αποτελέσει η δεοντολογία βασικό πυλώνα που θα διασφαλίζει και θα κλιμακώνει την αξιόπιστη TN.
- (14) Η εμπιστοσύνη στην ανάπτυξη, εγκατάσταση και χρήση των συστημάτων TN δεν αφορά μόνο τα εγγενή χαρακτηριστικά της τεχνολογίας, αλλά και τις ιδιότητες των κοινωνικοτεχνικών συστημάτων που περιλαμβάνουν εφαρμογές της TN<sup>9</sup>. Όπως αναλόγως συμβαίνει με τα ζητήματα (απώλειας) εμπιστοσύνης στον τομέα της αεροπορίας, της πυρηνικής ενέργειας ή της ασφάλειας τροφίμων, εμπιστοσύνη μπορεί να εμπνέει ή να μην εμπνέει το σύστημα TN υπό τη συνολικότερη έννοια του και όχι απλώς οι συνιστώσες του. Συνεπώς, η προσπάθεια για την επίτευξη αξιόπιστης TN δεν αφορά μόνο την αξιοπιστία του ίδιου του συστήματος TN, αλλά χρήζει ολιστικής και συστημικής προσέγγισης που ενσωματώνει την αξιοπιστία όλων των παραγόντων και των διεργασιών που αποτελούν μέρος του κοινωνικοτεχνικού περιβάλλοντος του συστήματος καθόλη τη διάρκεια του κύκλου ζωής του.
- (15) Η αξιόπιστη TN έχει **τρεις συνιστώσες**, οι οποίες θα πρέπει να πληρούνται σε ολόκληρο τον κύκλο ζωής του συστήματος:
1. Θα πρέπει να είναι **σύννομη**, ήτοι να διασφαλίζει την τήρηση όλων των εφαρμοστέων νομοθετικών και κανονιστικών διατάξεων.
  2. Θα πρέπει να είναι **δεοντολογική**, ήτοι να διασφαλίζει τη συμμόρφωση με δεοντολογικές αρχές και αξίες και
  3. Θα πρέπει να είναι **στιβαρή**, τόσο από τεχνικής όσο και από κοινωνικής άποψης, διότι, ακόμη κι όταν υπάρχει καλή πρόθεση, τα συστήματα TN μπορούν να προκαλέσουν ακούσια βλάβη.
- (16) Καθεμία από τις τρεις αυτές συνιστώσες αποτελεί από μόνη της αναγκαία αλλά όχι ικανή συνθήκη για την επίτευξη αξιόπιστης TN<sup>10</sup>. Υπό ιδανικές συνθήκες, και οι τρεις συνεργάζονται αρμονικά και αλληλεπικαλύπτονται στη λειτουργία τους. Στην πράξη, ωστόσο, μπορεί να υπάρξουν τριβές μεταξύ αυτών των στοιχείων (π.χ. ενίοτε το πεδίο εφαρμογής και το περιεχόμενο της κείμενης νομοθεσίας μπορεί να μην ευθυγραμμίζεται με τους κανόνες δεοντολογίας). Αποτελεί ατομική και συλλογική ευθύνη μας ως κοινωνίας

<sup>9</sup> Τα συστήματα αυτά περιλαμβάνουν ανθρώπους, δημόσιους φορείς, εταιρείες, υποδομές, λογισμικό, πρωτόκολλα, πρότυπα, διακυβέρνηση, ισχύουσες νομοθετικές διατάξεις, μηχανισμούς εποπτείας, δομές κινήτρων, διαδικασίες ελέγχου, αναφορά βέλτιστων πρακτικών κ.ά.

<sup>10</sup> Αυτό δεν αποκλείει το ενδεχόμενο να είναι ή να καταστούν αναγκαίες και άλλες, πρόσθετες προϋποθέσεις.

να εργαζόμαστε προς την κατεύθυνση της διασφάλισης ότι και οι τρεις συνιστώσες συμβάλλουν στην επίτευξη αξιόπιστης ΤΝ<sup>11</sup>.

- (17) Η αξιόπιστη προσέγγιση συνιστά βασικό παράγοντα για την «υπεύθυνη ανταγωνιστικότητα», καθώς παρέχει τις βάσεις που επιτρέπουν σε όλους όσοι επηρεάζονται από τα συστήματα ΤΝ να αντλούν τη βεβαιότητα ότι ο σχεδιασμός, η ανάπτυξη και η χρήση τους έχουν σύννομο, δεοντολογικό και στιβαρό χαρακτήρα. Οι παρούσες κατευθυντήριες γραμμές έχουν σκοπό την προώθηση της υπεύθυνης και βιώσιμης καινοτομίας στον τομέα της ΤΝ στην Ευρώπη. Επιδιώκουν να καταστήσουν τη δεοντολογία βασικό πυλώνα της ανάπτυξης μιας ενιαίας προσέγγισης της ΤΝ, η οποία θα επιχειρεί να ωφελεί, να ενισχύει και να προστατεύει τόσο την ατομική ευημερία του ανθρώπου όσο και το κοινό καλό της κοινωνίας. Πιστεύουμε ότι με τον τρόπο αυτό η Ευρώπη θα μπορέσει να αναδειχθεί σε πρωτοπόρο της προηγμένης ΤΝ παγκοσμίως, που θα αξίζει την εμπιστοσύνη μας σε ατομικό και συλλογικό επίπεδο. Η διασφάλιση της αξιοπιστίας αποτελεί απαραίτητη προϋπόθεση ώστε οι Ευρωπαίοι πολίτες να είναι σε θέση να αποκομίζουν πλήρως τα οφέλη των συστημάτων ΤΝ, με τη σιγουριά ότι εφαρμόζονται μέτρα για την πρόληψη των δυνητικών κινδύνων τους.
- (18) Η χρήση των συστημάτων ΤΝ δεν σταματά στο εθνικά σύνορα, και το ίδιο ισχύει για τις επιπτώσεις τους. Ως εκ τούτου, για τις παγκόσμιες ευκαιρίες και προκλήσεις που δημιουργεί η ΤΝ απαιτούνται παγκόσμιες λύσεις. Γι' αυτό, λοιπόν, προτρέπουμε όλα τα ενδιαφερόμενα μέρη να εργαστούν προς την κατεύθυνση της δημιουργίας ενός παγκόσμιου πλαισίου για αξιόπιστη ΤΝ, καλλιεργώντας τη συναίνεση σε διεθνές επίπεδο και, παράλληλα, προωθώντας και τηρώντας τη βασισμένη στα θεμελιώδη δικαιώματα προσέγγισή μας.

#### *Κοινό και πεδίο εφαρμογής*

- (19) Οι παρούσες κατευθυντήριες γραμμές απευθύνονται σε όλα τα ενδιαφερόμενα μέρη στον τομέα της ΤΝ που συμμετέχουν στον σχεδιασμό, στην ανάπτυξη, την εγκατάσταση, την υλοποίηση και τη χρήση ΤΝ, ή που επηρεάζονται από την ΤΝ, δηλαδή, μεταξύ άλλων, σε επιχειρήσεις, οργανισμούς, ερευνητές, δημόσιες υπηρεσίες και οργανισμούς, ιδρύματα, οργανώσεις της κοινωνίας των πολιτών, μεμονωμένους πολίτες, εργαζόμενους και καταναλωτές. Τα ενδιαφερόμενα μέρη που είναι προσηλωμένα στην επίτευξη αξιόπιστης ΤΝ μπορούν σε εθελοντική βάση να επιλέξουν να χρησιμοποιούν τις παρούσες κατευθυντήριες γραμμές ως μέθοδο για την επιχειρησιακή υλοποίηση της προσήλωσής τους, ιδίως κάνοντας χρήση του πρακτικού καταλόγου αξιολόγησης του κεφαλαίου III στις διεργασίες τους για την ανάπτυξη και εγκατάσταση των συστημάτων ΤΝ. Ο εν λόγω κατάλογος αξιολόγησης μπορεί επίσης να λειτουργεί συμπληρωματικά —και επομένως να ενσωματώνεται— σε υφιστάμενες διαδικασίες αξιολόγησης.
- (20) Στόχος των κατευθυντήριων γραμμών είναι η παροχή κατευθύνσεων για τις εφαρμογές ΤΝ εν γένει, ώστε να δημιουργηθεί μια οριζόντια βάση για την επίτευξη αξιόπιστης ΤΝ. Ωστόσο, **διαφορετικές καταστάσεις παρουσιάζουν διαφορετικές προκλήσεις**. Για τα συστήματα ΤΝ που προτείνουν μουσική δεν εγείρονται οι ίδιοι δεοντολογικοί προβληματισμοί όπως για τα συστήματα ΤΝ που προτείνουν κρίσιμες θεραπευτικές αγωγές. Ομοίως, διαφορετικές είναι οι ευκαιρίες και οι προκλήσεις που δημιουργούνται από τη χρήση συστημάτων ΤΝ στο πλαίσιο των σχέσεων επιχείρησης με καταναλωτή, επιχείρησης με επιχείρηση, εργοδότη με εργαζόμενο και δημόσιας υπηρεσίας με πολίτη ή, πιο γενικά, σε διαφορετικούς τομείς ή περιπτώσεις χρήσης. Επομένως, δεδομένων των ιδιαιτεροτήτων των συστημάτων ΤΝ ανάλογα με το εκάστοτε πλαίσιο, αναγνωρίζεται ότι η εφαρμογή αυτών των κατευθυντήριων γραμμών θα πρέπει να προσαρμόζεται στη συγκεκριμένη εφαρμογή ΤΝ. Επίσης, θα πρέπει να διερευνηθεί η αναγκαιότητα μιας πρόσθετης τομεακής προσέγγισης, η οποία θα συμπληρώνει το πιο γενικό οριζόντιο πλαίσιο που προτείνεται στο παρόν έγγραφο.

Για να κατανοήσουμε καλύτερα το πώς οι παρούσες κατευθύνσεις είναι δυνατό να εφαρμοστούν σε

<sup>11</sup> Αυτό σημαίνει επίσης ότι οι νομοθέτες ή οι φορείς χάραξης πολιτικής χρειάζεται να επανεξετάζουν την επάρκεια της κείμενης νομοθεσίας όταν αυτή δεν ευθυγραμμίζεται με τις δεοντολογικές αρχές.

οριζόντιο επίπεδο, καθώς και τα ζητήματα για τα οποία απαιτείται τομεακή προσέγγιση, καλούμε όλα τα ενδιαφερόμενα μέρη να εφαρμόσουν πιλοτικά τον κατάλογο αξιολόγησης της αξιοπιστίας της TN (κεφάλαιο III) με τον οποίο υλοποιείται επιχειρησιακά το πλαίσιο αυτό και να μας παράσχουν σχετική αναπληροφόρηση. Με βάση την αναπληροφόρηση που θα συλλεχθεί στο πλαίσιο αυτής της πιλοτικής φάσης θα αναθεωρήσουμε τον κατάλογο αξιολόγησης αυτών των κατευθυντήριων γραμμών έως τις αρχές του 2020. Η πιλοτική φάση θα έχει τεθεί σε εφαρμογή έως το καλοκαίρι του 2019 και θα διαρκέσει έως το τέλος του έτους. Θα μπορούν να συμμετέχουν σε αυτήν όλα τα ενδιαφερόμενα μέρη, δηλώνοντας το ενδιαφέρον τους μέσω της Ευρωπαϊκής Συμμαχίας για την TN.

## **B. ΈΝΑ ΠΛΑΙΣΙΟ ΓΙΑ ΑΞΙΟΠΙΣΤΗ ΤΝ**

- (21) Στις παρούσες κατευθυντήριες γραμμές διατυπώνεται ένα πλαίσιο για την επίτευξη αξιόπιστης ΤΝ με βάση τα θεμελιώδη δικαιώματα, όπως αυτά κατοχυρώνονται στον Χάρτη των Θεμελιωδών Δικαιωμάτων της ΕΕ (Χάρτης της ΕΕ) και στο συναφές διεθνές δίκαιο των ανθρωπίνων δικαιωμάτων. Ακολούθως εξετάζονται συνοπτικά οι τρεις συνιστώσες της αξιόπιστης ΤΝ.

### *Σύννομη ΤΝ*

- (22) Τα συστήματα ΤΝ δεν λειτουργούν σε έναν άνομο κόσμο. Μια σειρά από νομικά δεσμευτικούς κανόνες σε ευρωπαϊκό, εθνικό και διεθνές επίπεδο ήδη εφαρμόζονται ή είναι συναφή ως προς την ανάπτυξη, εγκατάσταση και χρήση συστημάτων ΤΝ σήμερα. Στις συναφείς νομικές πηγές περιλαμβάνονται, ενδεικτικά, το πρωτογενές δίκαιο της ΕΕ (οι συνθήκες της Ευρωπαϊκής Ένωσης και ο Χάρτης Θεμελιωδών Δικαιωμάτων της), το παράγωγο δίκαιο της ΕΕ (όπως ο γενικός κανονισμός για την προστασία δεδομένων, οι οδηγίες κατά των διακρίσεων, η οδηγία για τα μηχανήματα, η οδηγία για την ευθύνη λόγω ελαττωματικών προϊόντων, ο κανονισμός για την ελεύθερη ροή των δεδομένων μη προσωπικού χαρακτήρα, το δίκαιο σε θέματα προστασίας των καταναλωτών και οι οδηγίες για την ασφάλεια και την υγεία κατά την εργασία), αλλά και οι συνθήκες του ΟΗΕ για τα ανθρώπινα δικαιώματα και οι συμβάσεις του Συμβουλίου της Ευρώπης (όπως η Ευρωπαϊκή Σύμβαση ανθρωπίνων δικαιωμάτων) και πολυάριθμοι νόμοι των κρατών μελών της ΕΕ. Εκτός από τους οριζόντιους εφαρμοστέους κανόνες υπάρχουν διάφοροι κανόνες που αφορούν συγκεκριμένα πεδία, οι οποίοι εφαρμόζονται σε συγκεκριμένες εφαρμογές ΤΝ (όπως για παράδειγμα ο κανονισμός για τα ιατροτεχνολογικά προϊόντα στον τομέα της υγειονομικής περίθαλψης).
- (23) Η νομοθεσία προβλέπει τόσο θετικές όσο και αρνητικές υποχρεώσεις, που σημαίνει ότι ο νόμος δεν θα πρέπει να ερμηνεύεται μόνο σε συνάρτηση με το τι δεν επιτρέπεται να γίνεται, αλλά και σε συνάρτηση με το τι *θα πρέπει* να γίνεται. Ο νόμος αφενός απαγορεύει συγκεκριμένες ενέργειες αλλά, αφετέρου, επιτρέπει άλλες. Εν προκειμένω, επισημαίνεται ότι ο Χάρτης της ΕΕ περιέχει άρθρα σχετικά με την «επιχειρηματική ελευθερία» και την «ελευθερία των τεχνών και των επιστημών», παράλληλα με άρθρα που πραγματεύονται πεδία με τα οποία είμαστε πιο εξοικειωμένοι όταν το ζητούμενο είναι η διασφάλιση της αξιοπιστίας της ΤΝ, όπως για παράδειγμα η προστασία των δεδομένων και η απαγόρευση των διακρίσεων.
- (24) Οι παρούσες κατευθυντήριες γραμμές δεν πραγματεύονται ρητά την πρώτη συνιστώσα της αξιόπιστης ΤΝ (σύννομη ΤΝ). Στόχος τους είναι, αντ' αυτού, να παρέχουν κατευθύνσεις όσον αφορά την προαγωγή και τη διασφάλιση της δεύτερης και της τρίτης συνιστώσας (δεοντολογική και στιβαρή ΤΝ). Παρότι οι δύο τελευταίες ως ένα βαθμό αποτυπώνονται ήδη συχνά στην κείμενη νομοθεσία, η πλήρης πραγμάτωσή τους μπορεί να υπερβαίνει τις υφιστάμενες νομικές υποχρεώσεις.
- (25) Κανένα σημείο του παρόντος εγγράφου δεν πρέπει να εκληφθεί ή να ερμηνευθεί ως παροχή νομικών συμβουλών ή κατευθύνσεων σχετικά με το πώς μπορεί να επιτυγχάνεται η συμμόρφωση με εφαρμοστέους κείμενους νομικούς κανόνες και απαιτήσεις. Κανένα σημείο του παρόντος εγγράφου δεν γεννά νομικά δικαιώματα ούτε επιβάλλει νομικές υποχρεώσεις σε τρίτους. Ωστόσο, υπενθυμίζουμε ότι κάθε φυσικό ή νομικό πρόσωπο οφείλει να συμμορφώνεται με την νομοθεσία —τόσο αυτήν που εφαρμόζεται σήμερα όσο και εκείνη που θα θεσπιστεί στο μέλλον σε συνάρτηση με την ανάπτυξη της ΤΝ. Οι παρούσες κατευθυντήριες

γραμμές βασίζονται στην παραδοχή ότι **όλα τα νομικά δικαιώματα και οι υποχρεώσεις που ισχύουν για τις διεργασίες και τις δραστηριότητες οι οποίες σχετίζονται με την ανάπτυξη, εγκατάσταση και χρήση της ΤΝ διατηρούν τον υποχρεωτικό τους χαρακτήρα και πρέπει να τηρούνται δεόντως.**

#### *Δεοντολογική ΤΝ*

- (26) Για την επίτευξη αξιόπιστης ΤΝ δεν αρκεί η τήρηση της νομοθεσίας, καθώς αποτελεί τη μία μόνο από τις τρεις συνιστώσες της. Η νομοθεσία δεν συμβαδίζει πάντοτε με τις τεχνολογικές εξελίξεις, ενίοτε δεν είναι ευθυγραμμισμένη με τους κανόνες δεοντολογίας ή, απλώς, δεν είναι κατάλληλη για την αντιμετώπιση ορισμένων ζητημάτων. Για να είναι αξιόπιστα, τα συστήματα ΤΝ θα πρέπει επομένως να είναι δεοντολογικά, ώστε να διασφαλίζεται η ευθυγράμμισή τους με τους κανόνες δεοντολογίας.

#### *Στιβαρή ΤΝ*

- (27) Ακόμη κι αν διασφαλίζεται η ύπαρξη δεοντολογικού σκοπού, τα άτομα και η κοινωνία θα πρέπει να έχουν και τη βεβαιότητα ότι τα συστήματα ΤΝ δεν θα προκαλέσουν ακούσια βλάβη. Τα συστήματα αυτά θα πρέπει να λειτουργούν με ασφαλή και αξιόπιστο τρόπο, και θα πρέπει να προβλέπονται ασφαλιστικές δικλείδες που θα προλαμβάνουν τις όποιες ακούσιες δυσμενείς επιπτώσεις. Είναι επομένως σημαντικό να διασφαλίζεται ότι τα συστήματα ΤΝ είναι στιβαρά. Αυτό είναι αναγκαίο τόσο από τεχνικής άποψης (ώστε να διασφαλίζεται η τεχνική στιβαρότητα του συστήματος με τον ενδεδειγμένο τρόπο σε δεδομένο πλαίσιο, όπως ο τομέας εφαρμογής ή η φάση του κύκλου ζωής), όσο και από κοινωνικής άποψης (λαμβανομένου δεόντως υπόψη του πλαισίου και του περιβάλλοντος στο οποίο λειτουργεί το σύστημα). Η δεοντολογική και η στιβαρή ΤΝ είναι, επομένως, στενά συνυφασμένες και αλληλοσυμπληρούμενες. Οι αρχές που παρουσιάζονται στο κεφάλαιο I και οι απαιτήσεις του κεφαλαίου II που προκύπτουν από τις εν λόγω αρχές αφορούν και τις δύο αυτές συνιστώσες.

#### *Το πλαίσιο*

- (28) Στο παρόν έγγραφο οι κατευθύνσεις παρέχονται σε τρία επίπεδα, από τα πιο αφηρημένα στο κεφάλαιο I έως τα πιο συγκεκριμένα στο κεφάλαιο III:

**I) Βάσεις της αξιόπιστης ΤΝ.** Στο κεφάλαιο I παρουσιάζονται οι βάσεις της αξιόπιστης ΤΝ και αναλύεται η βασισμένη στα θεμελιώδη δικαιώματα<sup>12</sup> προσέγγισή της. Προσδιορίζονται και περιγράφονται οι δεοντολογικές αρχές που πρέπει να τηρούνται προκειμένου να διασφαλιστεί η δεοντολογική και στιβαρή ΤΝ.

**II) Πραγμάτωση της αξιόπιστης ΤΝ.** Στο κεφάλαιο II, αυτές οι δεοντολογικές αρχές μεταφράζονται σε επτά απαιτήσεις τις οποίες τα συστήματα ΤΝ θα πρέπει να εφαρμόζουν και να τηρούν καθόλη τη διάρκεια του κύκλου ζωής τους. Επίσης, παρουσιάζονται τόσο τεχνικές όσο και μη τεχνικές μέθοδοι που μπορούν να χρησιμοποιούνται για την εφαρμογή τους.

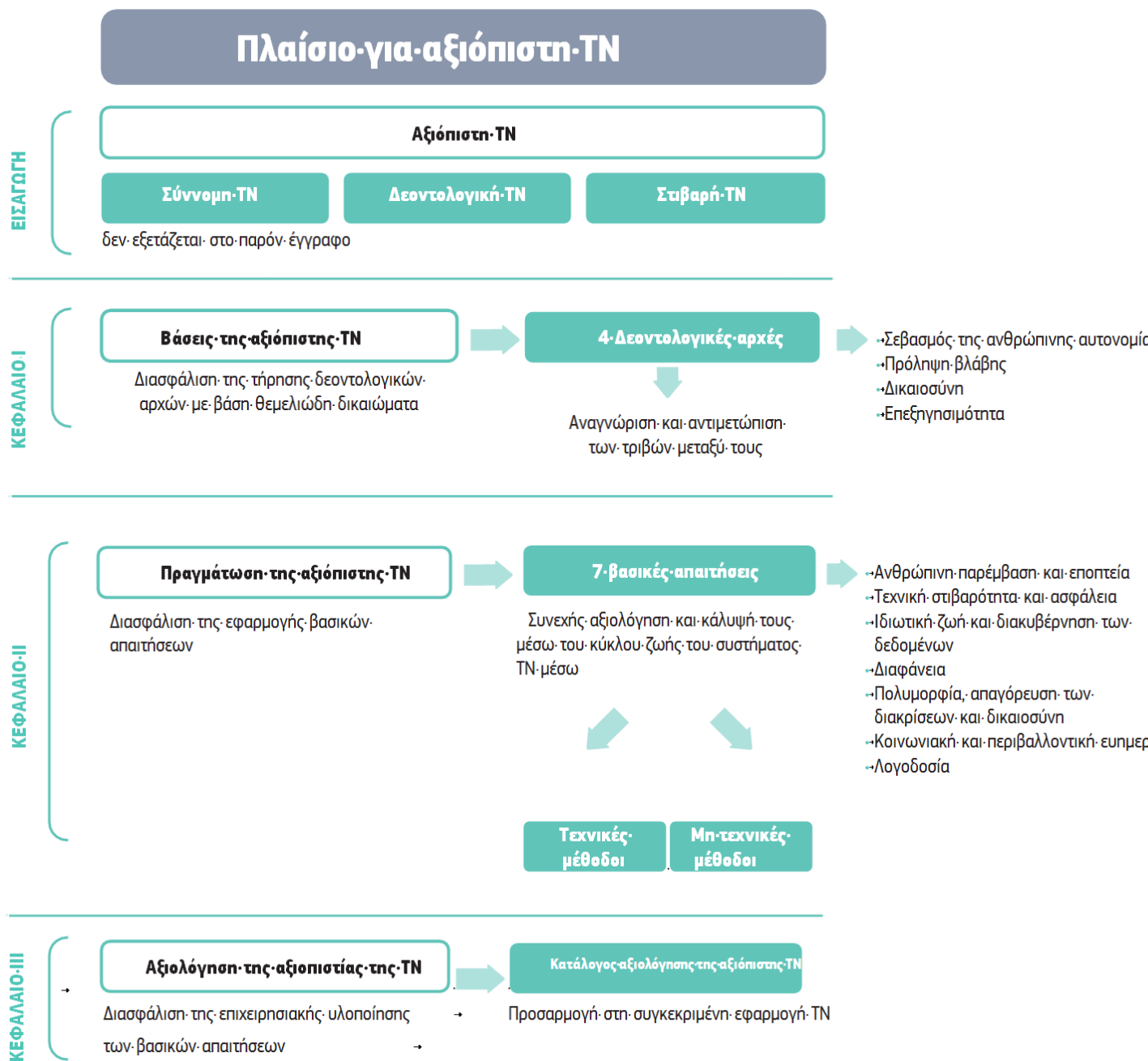
**III) Αξιολόγηση της αξιοπιστίας της ΤΝ.** Οι επαγγελματίες της ΤΝ χρειάζονται συγκεκριμένες κατευθύνσεις. Ως εκ τούτου, στο κεφάλαιο III παρατίθεται ένας προκαταρκτικός και μη εξαντλητικός κατάλογος αξιολόγησης της αξιοπιστίας της ΤΝ για την επιχειρησιακή υλοποίηση των απαιτήσεων του κεφαλαίου II. Η εν λόγω αξιολόγηση θα πρέπει να προσαρμόζεται στη συγκεκριμένη εφαρμογή του συστήματος.

- (29) Στην τελευταία ενότητα του εγγράφου παρουσιάζονται χρήσιμες ευκαιρίες και καίριοι προβληματισμοί που εγείρουν τα συστήματα ΤΝ και θα έπρεπε να εξεταστούν, επί των οποίων θα θέλαμε να δώσουμε το έναυσμα για περαιτέρω συζήτηση.

---

<sup>12</sup> Τα θεμελιώδη δικαιώματα αποτελούν τη βάση τόσο του διεθνούς όσο και του ενωσιακού δικαίου των ανθρωπίνων δικαιωμάτων σε αυτά εδράζονται τα νομικώς ισχυρά δικαιώματα που κατοχυρώνονται από τις Συνθήκες της ΕΕ και τον Χάρτη των Θεμελιωδών Δικαιωμάτων της ΕΕ. Επομένως, η τήρηση των θεμελιωδών δικαιωμάτων, δεδομένου ότι είναι νομικώς δεσμευτικά, εμπίπτει στην πρώτη συνιστώσα της αξιόπιστης ΤΝ, τη «σύννομη ΤΝ». Τα θεμελιώδη δικαιώματα μπορεί ωστόσο να θεωρηθεί ότι αντικατοπτρίζουν επίσης ειδικά ηθικά δικαιώματα όλων των ατόμων που προκύπτουν από την ανθρώπινη υπόστασή τους, ανεξάρτητα από τον νομικά δεσμευτικό χαρακτήρα τους. Υπό την έννοια αυτή, αποτελούν επίσης τμήμα και της δεύτερης συνιστώσας της αξιόπιστης ΤΝ, της «δεοντολογικής ΤΝ».

(30) Η δομή των κατευθυντήριων γραμμών απεικονίζεται στο *σχήμα 1* κατωτέρω.



Σχήμα 1: Οι κατευθυντήριες γραμμές ως ένα πλαίσιο για αξιόπιστη TN

## I. Κεφάλαιο I: Βάσεις της αξιόπιστης TN

- (31) Στο παρόν κεφάλαιο παρουσιάζονται οι βάσεις της αξιόπιστης TN, που εδράζονται στα θεμελιώδη δικαιώματα και αποτυπώνονται με τέσσερις δεοντολογικές αρχές που θα πρέπει να τηρούνται προκειμένου να διασφαλιστεί η δεοντολογική και στιβαρή TN. Το παρόν κεφάλαιο βασίζεται σε μεγάλο βαθμό στο πεδίο της δεοντολογίας.
- (32) Η δεοντολογία της TN αποτελεί υποπεδίο της εφαρμοσμένης δεοντολογίας, το οποίο εστιάζει στα δεοντολογικά ζητήματα που θέτει η ανάπτυξη, η εγκατάσταση και η χρήση της TN. Το βασικό της μέλημα είναι να προσδιορίσει τον τρόπο με τον οποίο η TN μπορεί να προωθήσει ή να εγείρει προβληματισμούς για την ευημερία των ατόμων, είτε από την άποψη της ποιότητας ζωής είτε της ανθρώπινης αυτονομίας και ελευθερίας που είναι απαραίτητες για μια δημοκρατική κοινωνία.
- (33) Ο δεοντολογικός αναστοχασμός σχετικά με την τεχνολογία TN μπορεί να εξυπηρετήσει πολλούς σκοπούς. Πρώτον, μπορεί να τροφοδοτήσει τον αναστοχασμό σχετικά με την ανάγκη προστασίας των ατόμων και των ομάδων στο πιο στοιχειώδες επίπεδο. Δεύτερον, μπορεί να ενθαρρύνει νέες μορφές καινοτομίας που επιδιώκουν να προωθήσουν δεοντολογικές αξίες, όπως εκείνες που συμβάλλουν στην επίτευξη των στόχων βιώσιμης ανάπτυξης των Ηνωμένων Εθνών<sup>13</sup>, οι οποίοι είναι στενά συνδεδεμένοι με το επικείμενο θεματολόγιο της ΕΕ για το 2030<sup>14</sup>. Παρότι το παρόν έγγραφο αφορά κυρίως τον πρώτο σκοπό που αναφέρεται ανωτέρω, δεν θα πρέπει να υποτιμάται η σημασία που θα μπορούσε να έχει η δεοντολογία για τον δεύτερο σκοπό. Η αξιόπιστη TN μπορεί να βελτιώσει την ατομική και τη συλλογική ευημερία, παράγοντας ευμάρεια, δημιουργώντας αξία και μεγιστοποιώντας τον πλούτο. Μπορεί να συνεισφέρει στην επίτευξη μιας δίκαιης κοινωνίας, συμβάλλοντας στη βελτίωση της υγείας και της ευημερίας των πολιτών με τρόπους που προωθούν την ισότητα στην κατανομή των οικονομικών, κοινωνικών και πολιτικών ευκαιριών.
- (34) Είναι λοιπόν επιτακτική ανάγκη να κατανοήσουμε τον τρόπο με τον οποίο μπορούμε να στηρίξουμε καλύτερα την ανάπτυξη, την εγκατάσταση και τη χρήση της TN, έτσι ώστε να διασφαλιστεί η δυνατότητα ευημερίας για όλους σε έναν κόσμο που βασίζεται στην TN, καθώς και η οικοδόμηση ενός καλύτερου μέλλοντος, διατηρώντας ταυτόχρονα την ανταγωνιστικότητά μας σε παγκόσμιο επίπεδο. Όπως συμβαίνει με κάθε ισχυρή τεχνολογία, η χρήση συστημάτων TN στην κοινωνία μας θέτει αρκετές δεοντολογικές προκλήσεις, παραδείγματος χάρη σχετικά με τις επιπτώσεις τους στον άνθρωπο και την κοινωνία, τις δυνατότητες λήψης αποφάσεων και την ασφάλεια. Εάν πρόκειται να χρησιμοποιούμε ολοένα και περισσότερο τη βοήθεια των συστημάτων TN ή να μεταβιβάζουμε τη λήψη αποφάσεων σε αυτά, θα πρέπει να διασφαλίζουμε ότι τα συγκεκριμένα συστήματα έχουν δίκαιες επιπτώσεις στη ζωή των ανθρώπων, ότι συμμορφώνονται με μη διαπραγματεύσιμες αξίες και ότι είναι σε θέση να ενεργήσουν ανάλογα, καθώς και ότι κατάλληλες διαδικασίες λογοδοσίας μπορούν να το διασφαλίζουν αυτό.
- (35) Η Ευρώπη θα πρέπει να ορίσει ποιο είναι το κανονιστικό όραμα που επιθυμεί να υλοποιήσει σε ένα μέλλον όπου η TN κατέχει κυρίαρχη θέση και κατά συνέπεια να αντιληφθεί ποια έννοια της TN θα πρέπει να μελετηθεί, να αναπτυχθεί, να εγκατασταθεί και να χρησιμοποιηθεί στην Ευρώπη για την επίτευξη του συγκεκριμένου οράματος. Με το παρόν έγγραφο, σκοπεύουμε να συνεισφέρουμε στη συγκεκριμένη προσπάθεια εισάγοντας την έννοια της αξιόπιστης TN, η οποία πιστεύουμε ότι είναι ο σωστός τρόπος για να οικοδομήσουμε ένα μέλλον με την TN. Ένα μέλλον όπου τα συστήματα TN βασίζονται στην έννοια της δημοκρατίας, του κράτους δικαίου και των θεμελιωδών δικαιωμάτων και όπου τα συστήματα αυτού του είδους βελτιώνουν συνεχώς και προασπίζουν το δημοκρατικό πνεύμα θα διαμορφώσει επίσης ένα περιβάλλον όπου μπορούν να ευδοκιμήσουν η καινοτομία και η υπεύθυνη ανταγωνιστικότητα.
- (36) Ένας κώδικας δεοντολογίας ειδικά για κάθε τομέα —όσο συνεπείς, ανεπτυγμένες και λεπτομερείς κι αν είναι οι μελλοντικές εκδόσεις του— δεν θα μπορούσε ποτέ να λειτουργήσει ως υποκατάστατο της ίδιας της

<sup>13</sup> [https://ec.europa.eu/commission/publications/reflection-paper-towards-sustainable-europe-2030\\_en](https://ec.europa.eu/commission/publications/reflection-paper-towards-sustainable-europe-2030_en)

<sup>14</sup> <https://sustainabledevelopment.un.org/?menu=1300>

δεοντολογικής συλλογιστικής, για την οποία θα πρέπει πάντα να λαμβάνονται υπόψη οι λεπτομέρειες του εκάστοτε πλαισίου που δεν είναι δυνατό να αποτυπωθούν σε γενικές κατευθυντήριες γραμμές. Πέρα από την ανάπτυξη ενός συνόλου κανόνων, η διασφάλιση της αξιόπιστης ΤΝ απαιτεί από εμάς να οικοδομήσουμε και να διατηρήσουμε ένα πνεύμα και μια νοοτροπία δεοντολογίας μέσα από τον δημόσιο διάλογο, την εκπαίδευση και την πρακτική μάθηση.

### 1. Τα θεμελιώδη δικαιώματα ως ηθικά και νομικά δικαιώματα

- (37) Πιστεύουμε σε μια προσέγγιση της δεοντολογίας της ΤΝ που βασίζεται στα θεμελιώδη δικαιώματα τα οποία κατοχυρώνονται στις Συνθήκες της ΕΕ,<sup>15</sup> τον Χάρτη των Θεμελιωδών Δικαιωμάτων της ΕΕ (Χάρτης της ΕΕ) και το διεθνές δίκαιο των ανθρωπίνων δικαιωμάτων<sup>16</sup>. Ο σεβασμός των θεμελιωδών δικαιωμάτων, μέσα σε ένα πλαίσιο δημοκρατίας και κράτους δικαίου, προσφέρει τις πιο ελπιδοφόρες βάσεις για τον προσδιορισμό αφηρημένων δεοντολογικών αρχών και αξιών που είναι δυνατό να υλοποιηθούν επιχειρησιακά στο πλαίσιο της ΤΝ.
- (38) Οι Συνθήκες της ΕΕ και ο Χάρτης της ΕΕ ορίζουν μια σειρά θεμελιωδών δικαιωμάτων τα οποία τα κράτη μέλη της ΕΕ και τα θεσμικά όργανα της ΕΕ υποχρεούνται νομικά να σέβονται κατά την εφαρμογή της νομοθεσίας της ΕΕ. Τα συγκεκριμένα δικαιώματα περιγράφονται στον Χάρτη της ΕΕ και αφορούν συγκεκριμένα την αξιοπρέπεια, τις ελευθερίες, την ισότητα και την αλληλεγγύη, τα δικαιώματα των πολιτών και τη δικαιοσύνη. Η κοινή βάση που συνενώνει τα εν λόγω δικαιώματα μπορεί να θεωρηθεί ότι εδράζεται στον σεβασμό για την ανθρώπινη αξιοπρέπεια —αποτυπώνοντας αυτό που περιγράφεται ως «ανθρωποκεντρική προσέγγιση», σύμφωνα με την οποία ο άνθρωπος απολαμβάνει τη μοναδική και αναφαίρετη ηθική υπόσταση υπεροχής στους τομείς των ατομικών, πολιτικών, οικονομικών και κοινωνικών δικαιωμάτων<sup>17</sup>.
- (39) Ενώ τα δικαιώματα που ορίζονται στον Χάρτη της ΕΕ είναι νομικώς δεσμευτικά,<sup>18</sup> είναι σημαντικό να αναγνωριστεί ότι τα θεμελιώδη δικαιώματα δεν παρέχουν σε κάθε περίπτωση ολοκληρωμένη νομική προστασία. Για παράδειγμα, για τον Χάρτη της ΕΕ, είναι σημαντικό να υπογραμμιστεί ότι το πεδίο εφαρμογής του περιορίζεται σε τομείς του ενωσιακού δικαίου. Το διεθνές δίκαιο των ανθρωπίνων δικαιωμάτων και ιδίως η Ευρωπαϊκή Σύμβαση για τα Δικαιώματα του Ανθρώπου είναι νομικώς δεσμευτικά για τα κράτη μέλη της ΕΕ, ακόμη και σε τομείς που δεν εμπίπτουν στο πεδίο εφαρμογής του ενωσιακού δικαίου. Ταυτόχρονα, θα πρέπει να υπογραμμιστεί ότι τα θεμελιώδη δικαιώματα απονέμονται επίσης σε άτομα και (σε ορισμένο βαθμό) σε ομάδες λόγω της ηθικής τους υπόστασης ως ανθρώπων, ανεξάρτητα από τη νομική τους ισχύ. Επομένως, ως νομικώς ισχυρά δικαιώματα, τα θεμελιώδη δικαιώματα εμπίπτουν στην πρώτη συνιστώσα της αξιόπιστης ΤΝ (σύννομη ΤΝ), η οποία διασφαλίζει τη συμμόρφωση με τον νόμο. Ως δικαιώματα με καθολικό χαρακτήρα, που έχει τις ρίζες του στην εγγενή ηθική υπόσταση των ανθρώπων, τα εν λόγω δικαιώματα στηρίζουν επίσης τη δεύτερη συνιστώσα της αξιόπιστης ΤΝ (δεοντολογική ΤΝ) και αφορούν τους δεοντολογικούς κανόνες που δεν είναι απαραίτητα νομικώς δεσμευτικοί, αλλά έχουν καίρια σημασία για την εξασφάλιση της αξιοπιστίας. Δεδομένου ότι το παρόν έγγραφο δεν αποσκοπεί στην παροχή καθοδήγησης σχετικά με την προηγούμενη συνιστώσα, για τους σκοπούς αυτών των μη δεσμευτικών κατευθυντήριων γραμμών, οι αναφορές στα θεμελιώδη δικαιώματα αντικατοπτρίζουν τη δεύτερη συνιστώσα.

<sup>15</sup> Η ΕΕ βασίζεται σε μια συνταγματική δέσμευση για την προστασία των θεμελιωδών και αδιαίρετων δικαιωμάτων των ανθρώπων, την εξασφάλιση του σεβασμού του κράτους δικαίου, την ενίσχυση της δημοκρατικής ελευθερίας και την προώθηση του κοινού καλού. Τα συγκεκριμένα δικαιώματα αποτυπώνονται στα άρθρα 2 και 3 της συνθήκης για την Ευρωπαϊκή Ένωση και στον Χάρτη των Θεμελιωδών Δικαιωμάτων της ΕΕ.

<sup>16</sup> Άλλες νομικές πράξεις αποτυπώνουν και παρέχουν περαιτέρω διευκρινίσεις σχετικά με τις συγκεκριμένες δεσμεύσεις, όπως για παράδειγμα ο Ευρωπαϊκός Κοινωνικός Χάρτης του Συμβουλίου της Ευρώπης ή συγκεκριμένη νομοθεσία όπως ο γενικός κανονισμός της ΕΕ για την προστασία των δεδομένων.

<sup>17</sup> Θα πρέπει να σημειωθεί ότι η προσήλωση στην ανθρωποκεντρική ΤΝ και την έδρασή της στα θεμελιώδη δικαιώματα απαιτεί συλλογικές κοινωνιακές και συνταγματικές βάσεις οι οποίες καθιστούν πρακτικά εφικτή και ουσιαστική την ατομική ελευθερία και τον σεβασμό της ανθρώπινης αξιοπρέπειας, αντί να υποδηλώνουν μια υπέρμετρα ατομικιστική θεώρηση της ανθρώπινης υπόστασης.

<sup>18</sup> Σύμφωνα με το άρθρο 51 του Χάρτη, εφαρμόζεται στα θεσμικά όργανα της ΕΕ και στα κράτη μέλη της ΕΕ κατά την εφαρμογή της νομοθεσίας της ΕΕ.

## 2. Από τα θεμελιώδη δικαιώματα στις δεοντολογικές αρχές

### 2.1 Θεμελιώδη δικαιώματα ως βάση για αξιόπιστη ΤΝ

- (40) Από τη συνολική δέσμη αδιαίρετων δικαιωμάτων που ορίζονται στο διεθνές δίκαιο των ανθρωπίνων δικαιωμάτων, στις Συνθήκες της ΕΕ και στον Χάρτη της ΕΕ, οι παρακάτω οικογένειες θεμελιωδών δικαιωμάτων είναι ιδιαίτερες πρόσφορες για να καλύψουν τα συστήματα ΤΝ. Πολλά από τα εν λόγω δικαιώματα είναι, υπό συγκεκριμένες συνθήκες, νομικώς ισχυρά στην ΕΕ, έτσι ώστε η τήρηση των όρων τους να είναι νομικώς υποχρεωτική. Αλλά ακόμη και μετά τη συμμόρφωση με τα νομικώς ισχυρά θεμελιώδη δικαιώματα, ο δεοντολογικός προβληματισμός μπορεί να μας βοηθήσει να κατανοήσουμε τον τρόπο με τον οποίο η ανάπτυξη, η εγκατάσταση και η χρήση της ΤΝ μπορεί να επηρεάζει τα θεμελιώδη δικαιώματα και τις υποκείμενες αξίες τους, ενώ μπορεί να συμβάλλει στην παροχή πιο λεπτομερούς καθοδήγησης κατά την εξέταση του τι *θα πρέπει* να κάνουμε αντί του τι *μπορούμε* να κάνουμε (επί του παρόντος) με την τεχνολογία.
- (41) **Σεβασμός της ανθρώπινης αξιοπρέπειας.** Η ανθρώπινη αξιοπρέπεια περιλαμβάνει την ιδέα ότι κάθε άνθρωπος έχει μια «εγγενή αξία», η οποία δεν πρέπει ποτέ να υποβαθμίζεται, να παραβιάζεται ή να καταστέλλεται από άλλους —ούτε από τις νέες τεχνολογίες όπως τα συστήματα ΤΝ.<sup>19</sup> Στο πλαίσιο της ΤΝ, ο σεβασμός της ανθρώπινης αξιοπρέπειας συνεπάγεται ότι όλοι οι άνθρωποι αντιμετωπίζονται με σεβασμό λόγω του ότι αποτελούν ηθικά *υποκείμενα*, και όχι απλώς *αντικείμενα* που υφίστανται διαλογή, ταξινόμηση, χαρακτηρισμό, αγελοποίηση, κοινωνικό προγραμματισμό ή χειραγώγηση. Επομένως, τα συστήματα ΤΝ θα πρέπει να αναπτυχθούν κατά τέτοιο τρόπο ώστε να σέβονται, να εξυπηρετούν και να προστατεύουν τη σωματική και πνευματική ακεραιότητα του ανθρώπου, την προσωπική και πολιτισμική αίσθηση της ταυτότητας και την ικανοποίηση των βασικών αναγκών του<sup>20</sup>.
- (42) **Ελευθερία του ατόμου.** Οι άνθρωποι θα πρέπει να παραμένουν ελεύθεροι να λαμβάνουν οι ίδιοι αποφάσεις για τη ζωή τους. Τούτο συνεπάγεται απαλλαγή από κυριαρχική εισβολή, αλλά απαιτεί επίσης παρέμβαση από κυβερνητικές και μη κυβερνητικές οργανώσεις για να διασφαλιστεί ότι τα άτομα ή οι άνθρωποι που αντιμετωπίζουν κίνδυνο αποκλεισμού έχουν ίση πρόσβαση στα οφέλη και τις ευκαιρίες της ΤΝ. Σε ένα περιβάλλον ΤΝ, η ελευθερία του ατόμου απαιτεί μετριασμό του άμεσου και έμμεσου παράνομου εξαναγκασμού, των απειλών στην ψυχική αυτονομία και ψυχική υγεία, της αδικαιολόγητης επιτήρησης, της εξαπάτησης και της αθέμιτης χειραγώγησης. Στην πραγματικότητα, η ελευθερία του ατόμου συνεπάγεται τη δέσμευση να δοθεί στα άτομα η δυνατότητα να ασκούν ακόμη μεγαλύτερο έλεγχο στη ζωή τους, συμπεριλαμβανομένης της προστασίας (μεταξύ άλλων δικαιωμάτων) της επιχειρηματικής ελευθερίας, της ελευθερίας της τέχνης και της επιστήμης, της ελευθερίας έκφρασης, της ιδιωτικής ζωής και της προστασίας προσωπικών δεδομένων, καθώς και της ελευθερίας του συνέρχεσθαι και του συνεταιρίζεσθαι.
- (43) **Σεβασμός της δημοκρατίας, της δικαιοσύνης και του κράτους δικαίου.** Κάθε κυβερνητική εξουσία στις συνταγματικές δημοκρατίες πρέπει να είναι νομίμως εξουσιοδοτημένη και περιορισμένη από τον νόμο. Τα συστήματα ΤΝ θα πρέπει να χρησιμοποιούνται για τη διατήρηση και την ενθάρρυνση των δημοκρατικών διαδικασιών και τον σεβασμό προς την πολυφωνία των αξιών και των επιλογών ζωής των ατόμων. Τα συστήματα ΤΝ δεν θα πρέπει να υπονομεύουν τις δημοκρατικές διαδικασίες, τις ανθρώπινες διαβουλεύσεις ή τα συστήματα δημοκρατικής ψηφοφορίας. Τα συστήματα ΤΝ θα πρέπει επίσης να ενσωματώνουν τη δέσμευση να διασφαλίσουν ότι δεν λειτουργούν με τρόπους που υπονομεύουν τις θεμελιώδεις δεσμεύσεις στις οποίες στηρίζεται το κράτος δικαίου, τους υποχρεωτικούς νόμους και κανονισμούς, καθώς επίσης και να διασφαλίσουν τις δέουσες διαδικασίες και την ισότητα ενώπιον του νόμου.
- (44) **Ισότητα, απαγόρευση των διακρίσεων και αλληλεγγύη** —*συμπεριλαμβανομένων των δικαιωμάτων των*

<sup>19</sup> C. McCrudden, Human Dignity and Judicial Interpretation of Human Rights, *EJIL*, 19(4), 2008.

<sup>20</sup> Για την κατανόηση της έννοιας της «ανθρώπινης αξιοπρέπειας» σε αυτό το πλαίσιο, βλ. E. Hilgendorf, Problem Areas in the Dignity Debate and the Ensemble Theory of Human Dignity, στο: D. Grimm, A. Kemmerer, C. Möllers (eds.), *Human Dignity in Context. Explorations of a Contested Concept*, 2018, σ. 325 και επ.



ατόμων που αντιμετωπίζουν κίνδυνο αποκλεισμού. Θα πρέπει να διασφαλίζεται ο ισότιμος σεβασμός της ηθικής αξίας και της αξιοπρέπειας όλων των ανθρώπων. Τούτο υπερβαίνει την απαγόρευση των διακρίσεων, σύμφωνα με την οποία είναι ανεκτές οι διακρίσεις μεταξύ ανόμιων καταστάσεων βάσει αντικειμενικής αιτιολόγησης. Σε ένα περιβάλλον ΤΝ, η ισότητα συνεπάγεται ότι οι λειτουργίες του συστήματος δεν είναι δυνατό να οδηγούν σε αποτελέσματα που χαρακτηρίζονται από αθέμιτη μεροληψία (π.χ. τα δεδομένα που χρησιμοποιούνται για την εκπαίδευση των συστημάτων ΤΝ θα πρέπει να είναι όσο το δυνατόν πιο συμπεριληπτικά, αντιπροσωπεύοντας διαφορετικές ομάδες πληθυσμού). Το γεγονός αυτό απαιτεί επίσης τον κατάλληλο σεβασμό των δυνητικά ευάλωτων ατόμων και ομάδων<sup>21</sup>, όπως οι εργαζόμενοι, οι γυναίκες, τα άτομα με αναπηρίες, οι εθνοτικές μειονότητες, τα παιδιά, οι καταναλωτές ή άλλα άτομα που αντιμετωπίζουν κίνδυνο αποκλεισμού.

- (45) **Δικαιώματα πολιτών.** Οι πολίτες επωφελούνται από ένα ευρύ φάσμα δικαιωμάτων, μεταξύ των οποίων το δικαίωμα ψήφου, το δικαίωμα στη χρηστή διοίκηση ή την πρόσβαση σε δημόσια έγγραφα και το δικαίωμα υποβολής αναφοράς στη διοίκηση. Τα συστήματα ΤΝ προσφέρουν σημαντικές δυνατότητες βελτίωσης της κλίμακας και της αποδοτικότητας της διακυβέρνησης όσον αφορά την παροχή δημόσιων αγαθών και υπηρεσιών προς την κοινωνία. Ταυτόχρονα, τα δικαιώματα των πολιτών θα ήταν επίσης δυνατό να επηρεαστούν αρνητικά από τις εφαρμογές της ΤΝ και θα πρέπει να διασφαλιστούν. Όταν χρησιμοποιείται στο παρόν κείμενο ο όρος «δικαιώματα των πολιτών», δεν αποκλείονται ούτε παραβλέπονται τα δικαιώματα των υπηκόων τρίτων χωρών και των αντικανονικών (ή παράνομων) ατόμων στην ΕΕ που έχουν επίσης δικαιώματα βάσει του διεθνούς δικαίου και —ως εκ τούτου— στον τομέα της ΤΝ.

## 2.2 Δεοντολογικές αρχές στο πλαίσιο των συστημάτων ΤΝ<sup>22</sup>

- (46) Πολλές δημόσιες και ιδιωτικές οργανώσεις, καθώς και οργανώσεις πολιτών, άντλησαν έμπνευση από τα θεμελιώδη δικαιώματα για την εκπόνηση δεοντολογικών πλαισίων για την ΤΝ<sup>23</sup>. Στην ΕΕ, η «Ευρωπαϊκή ομάδα για τη δεοντολογία της επιστήμης και των νέων τεχνολογιών» («ΕΟΔ»), πρότεινε ένα σύνολο 9 βασικών αρχών, βασισμένων στις θεμελιώδεις αξίες που ορίζονται στις Συνθήκες της ΕΕ και στον Χάρτη των Θεμελιωδών Δικαιωμάτων της ΕΕ.<sup>24</sup> Εξελίσσουμε περαιτέρω το συγκεκριμένο έργο, αναγνωρίζοντας τις περισσότερες από τις αρχές που προωθούσαν διάφορες ομάδες μέχρι στιγμής, αποσαφηνίζοντας παράλληλα τους στόχους που όλες οι αρχές επιδιώκουν να καλλιεργήσουν και να υποστηρίξουν. Οι εν λόγω δεοντολογικές αρχές είναι δυνατό να εμπνεύσουν νέα και ειδικά ρυθμιστικά μέσα, να διευκολύνουν την ερμηνεία των θεμελιωδών δικαιωμάτων κατά την εξέλιξη του κοινωνικοτεχνικού μας περιβάλλοντος με την πάροδο του χρόνου, ενώ μπορούν να καθοδηγήσουν το σκεπτικό για την ανάπτυξη, τη χρήση και την εφαρμογή των συστημάτων ΤΝ, προσαρμοζόμενες δυναμικά με την εξέλιξη της ίδιας της κοινωνίας.
- (47) Τα συστήματα ΤΝ θα πρέπει να βελτιώνουν την ατομική και συλλογική ευημερία. Στην παρούσα ενότητα παρατίθενται **τέσσερις δεοντολογικές αρχές**, που εδράζονται στα θεμελιώδη δικαιώματα, οι οποίες θα πρέπει να τηρούνται προκειμένου να διασφαλιστεί ότι τα συστήματα ΤΝ αναπτύσσονται, εγκαθίστανται και χρησιμοποιούνται με αξιόπιστο τρόπο. Ορίζονται ως **δεοντολογικές επιταγές**, έτσι ώστε οι επαγγελματίες της ΤΝ να προσπαθούν πάντα να συμμορφώνονται με αυτές. Χωρίς να επιβάλλουμε μια ιεράρχηση, παραθέτουμε τις παρακάτω αρχές με τρόπο που αντικατοπτρίζει τη σειρά εμφάνισης των θεμελιωδών

<sup>21</sup> Για μια περιγραφή του όρου όπως χρησιμοποιείται στο σύνολο του εγγράφου, ανατρέξτε στο γλωσσάριο.

<sup>22</sup> Οι συγκεκριμένες αρχές ισχύουν επίσης για την ανάπτυξη, την εγκατάσταση και τη χρήση άλλων τεχνολογιών και ως εκ τούτου δεν αφορούν συγκεκριμένα τα συστήματα ΤΝ. Εν συνεχεία, επιδίωξαμε να προσδιορίσουμε τη συνάφειά τους ειδικά σε ένα πλαίσιο που αφορά την ΤΝ.

<sup>23</sup> Η στήριξη στα θεμελιώδη δικαιώματα συμβάλλει επίσης στον περιορισμό της κανονιστικής αβεβαιότητας, καθώς μπορεί να βασιστεί σε δεκαετίες πρακτικής προστασίας των θεμελιωδών δικαιωμάτων στην ΕΕ, παρέχοντας συνεπώς σαφήνεια, καλύτερη κατανόηση και προβλεψιμότητα.

<sup>24</sup> Πιο πρόσφατα, η ομάδα εργασίας του AI4People εξέτασε τις προαναφερθείσες αρχές της ΕΟΔ, καθώς και άλλες 36 δεοντολογικές αρχές που έχουν υποβληθεί μέχρι σήμερα, και τις συνόψισε σε τέσσερις γενικές αρχές: L. Floridi, J. Cowls, M. Beltrametti, R. Chatila, P. Chazerand, V. Dignum, C. Luetge, R. Madelin, U. Pagallo, F. Rossi, B. Schafer, P. Valcke, E. J. M. Vayena (2018), «AI4People — An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations», Minds and Machines 28(4): 689-707.

δικαιωμάτων στα οποία βασίζονται στον Χάρτη της ΕΕ<sup>25</sup>.

(48) Οι εν λόγω αρχές είναι:

- i) Σεβασμός της ανθρώπινης αυτονομίας
- ii) Πρόληψη βλάβης
- iii) Δικαιοσύνη
- iv) Επεξηγησιμότητα

(49) Πολλές από αυτές αποτυπώνονται ήδη σε μεγάλο βαθμό στις υφιστάμενες νομικές απαιτήσεις για τις οποίες απαιτείται υποχρεωτική συμμόρφωση και κατά συνέπεια εμπίπτουν επίσης στο πεδίο εφαρμογής της «σύννομης ΤΝ», η οποία αποτελεί την πρώτη συνιστώσα της αξιόπιστης ΤΝ<sup>26</sup>. Ωστόσο, όπως προαναφέρθηκε, ενώ πολλές νομικές υποχρεώσεις αντικατοπτρίζουν δεοντολογικές αρχές, η τήρηση των δεοντολογικών αρχών υπερβαίνει την επίσημη συμμόρφωση με την κείμενη νομοθεσία<sup>27</sup>.

- Η αρχή του σεβασμού της ανθρώπινης αυτονομίας

(50) Τα θεμελιώδη δικαιώματα στα οποία βασίζεται η ΕΕ έχουν γνώμονα τον σεβασμό της ελευθερίας και της αυτονομίας των ανθρώπων. Οι άνθρωποι που αλληλεπιδρούν με τα συστήματα ΤΝ πρέπει να είναι σε θέση να διατηρούν πλήρη και αποτελεσματική αυτοδιάθεση και να είναι σε θέση να συμμετέχουν στη δημοκρατική διαδικασία. Τα συστήματα ΤΝ δεν θα πρέπει αδικαιολόγητα να προβαίνουν σε καθυπόταξη, εξαναγκασμό, εξαπάτηση, χειραγώγηση, κοινωνικό προγραμματισμό ή αγελοποίηση των ανθρώπων. Αντ' αυτού, τα συστήματα ΤΝ θα πρέπει να σχεδιάζονται έτσι ώστε να αυξάνουν, να συμπληρώνουν και να ενισχύουν τις ανθρώπινες γνωστικές, κοινωνικές και πολιτισμικές δεξιότητες. Η κατανομή των λειτουργιών μεταξύ των ανθρώπων και των συστημάτων ΤΝ θα πρέπει να ακολουθεί τις αρχές του ανθρωποκεντρικού σχεδιασμού και να προσφέρει ουσιαστικές ευκαιρίες για ανθρώπινη επιλογή. Αυτό σημαίνει εξασφάλιση της ανθρώπινης εποπτείας<sup>28</sup> και ελέγχου των διαδικασιών εργασίας στα συστήματα ΤΝ. Τα συστήματα ΤΝ ενδέχεται επίσης να μετασχηματίσουν ουσιαστικά τον τομέα της εργασίας. Θα πρέπει να στηρίζουν τους ανθρώπους στο εργασιακό περιβάλλον και να αποσκοπούν στη δημιουργία ουσιαστικού έργου.

- Η αρχή της πρόληψης βλάβης

(51) Τα συστήματα ΤΝ δεν θα πρέπει ούτε να προκαλούν ούτε να επιδεινώνουν τυχόν βλάβες<sup>29</sup> ή να επηρεάζουν με άλλο τρόπο τον άνθρωπο αρνητικά<sup>30</sup>. Αυτό συνεπάγεται την προστασία της ανθρώπινης αξιοπρέπειας, καθώς και της πνευματικής και σωματικής ακεραιότητας. Τα συστήματα ΤΝ και τα περιβάλλοντα στα οποία λειτουργούν θα πρέπει να είναι ασφαλή και προστατευμένα. Θα πρέπει να είναι στιβαρά από τεχνικής άποψης, ενώ θα πρέπει να διασφαλίζεται ότι δεν είναι ανοικτά σε κακόβουλη χρήση. Θα πρέπει να δοθεί περισσότερη προσοχή στα ευάλωτα άτομα και να συμπεριληφθούν στην ανάπτυξη και την εγκατάσταση συστημάτων ΤΝ. Ιδιαίτερη προσοχή θα πρέπει να δοθεί επίσης σε καταστάσεις όπου τα συστήματα ΤΝ είναι δυνατό να προκαλέσουν ή να επιδεινώσουν αρνητικές επιπτώσεις λόγω ασυμμετρίας εξουσίας ή πληροφόρησης, όπως μεταξύ εργοδοτών και εργαζομένων, επιχειρήσεων και καταναλωτών ή κυβερνήσεων

<sup>25</sup> Ο σεβασμός της ανθρώπινης αυτονομίας συνδέεται στενά με το δικαίωμα στην ανθρώπινη αξιοπρέπεια και την ελευθερία (που αποτυπώνονται στα άρθρα 1 και 6 του Χάρτη). Η πρόληψη βλάβης συνδέεται στενά με την προστασία της σωματικής ή πνευματικής ακεραιότητας (που αποτυπώνεται στο άρθρο 3). Η δικαιοσύνη συνδέεται στενά με τα δικαιώματα της απαγόρευσης των διακρίσεων, της αλληλεγγύης και της δικαιοσύνης (που αποτυπώνονται στα άρθρα 21 και εξής). Η επεξηγησιμότητα και η ευθύνη συνδέονται στενά με τα δικαιώματα που σχετίζονται με τη δικαιοσύνη (που αποτυπώνεται στο άρθρο 47).

<sup>26</sup> Επί παραδείγματι, ο ΓΚΠΔ ή ο κανονισμός για την προστασία των καταναλωτών.

<sup>27</sup> Για περισσότερες πληροφορίες σχετικά με το συγκεκριμένο ζήτημα, βλ. για παράδειγμα L. Floridi, *Soft Ethics and the Governance of the Digital*, *Philosophy & Technology*, Μάρτιος 2018, τόμος 31, τεύχος 1, σ. 1–8.

<sup>28</sup> Η έννοια της ανθρώπινης εποπτείας αναπτύσσεται περαιτέρω στην παράγραφο 65 παρακάτω.

<sup>29</sup> Οι βλάβες είναι δυνατό να είναι ατομικές ή συλλογικές και μπορεί να περιλαμβάνουν άυλη βλάβη σε κοινωνικά, πολιτιστικά και πολιτικά περιβάλλοντα.

<sup>30</sup> Αυτό περιλαμβάνει επίσης τον τρόπο διαβίωσης των ατόμων και των κοινωνικών ομάδων, για παράδειγμα με την αποφυγή πολιτιστικής βλάβης.

και πολιτών. Η πρόληψη βλάβης συνεπάγεται επίσης την εξέταση του φυσικού περιβάλλοντος και όλων των έμβιων όντων.

- Η αρχή της δικαιοσύνης

(52) Η ανάπτυξη, η εγκατάσταση και η χρήση συστημάτων ΤΝ θα πρέπει να γίνεται με τρόπο δίκαιο. Ενώ αναγνωρίζουμε ότι υπάρχουν πολλές διαφορετικές ερμηνείες της δικαιοσύνης, πιστεύουμε ότι η δικαιοσύνη έχει τόσο ουσιαστική όσο και διαδικαστική διάσταση. Η ουσιαστική διάσταση συνεπάγεται δέσμευση για: εξασφάλιση της ισότιμης και δίκαιης κατανομής τόσο των ωφελειών όσο και του κόστους, καθώς και εξασφάλιση ότι τα άτομα και οι ομάδες δεν υφίστανται αθέμιτη μεροληψία, διακρίσεις και στιγματισμό. Εάν καταστεί δυνατό να αποφευχθεί η αθέμιτη μεροληψία, τα συστήματα ΤΝ μπορεί ακόμη και να αυξήσουν την κοινωνιακή δικαιοσύνη. Θα πρέπει επίσης να προωθηθούν οι ίσες ευκαιρίες όσον αφορά την πρόσβαση στην εκπαίδευση, τα αγαθά, τις υπηρεσίες και την τεχνολογία. Επιπλέον, η χρήση συστημάτων ΤΝ δεν θα πρέπει ποτέ να οδηγεί στην εξαπάτηση των (τελικών) χρηστών ή τον περιορισμό της ελευθερίας επιλογής τους. Επιπλέον, η δικαιοσύνη συνεπάγεται ότι οι επαγγελματίες του τομέα της ΤΝ θα πρέπει να σέβονται την αρχή της αναλογικότητας μεταξύ μέσων και σκοπών και να εξετάζουν προσεκτικά τον τρόπο εξισορρόπησης ανταγωνιστικών συμφερόντων και στόχων.<sup>31</sup> Η διαδικαστική διάσταση της δικαιοσύνης συνεπάγεται τη δυνατότητα αμφισβήτησης και αποτελεσματικής έννομης προστασίας έναντι αποφάσεων που λαμβάνονται από συστήματα ΤΝ και από τους ανθρώπους που τα χειρίζονται.<sup>32</sup> Προκειμένου να γίνει αυτό, η οντότητα που είναι υπεύθυνη για την απόφαση θα πρέπει να είναι αναγνωρίσιμη, ενώ θα πρέπει να επεξηγούνται οι διαδικασίες λήψης αποφάσεων.

- Η αρχή της επεξηγησιμότητας

(53) Η επεξηγησιμότητα είναι καθοριστική για την οικοδόμηση και διατήρηση της εμπιστοσύνης των χρηστών απέναντι στα συστήματα ΤΝ. Αυτό σημαίνει ότι οι διαδικασίες πρέπει να είναι διαφανείς, οι δυνατότητες και ο σκοπός των συστημάτων ΤΝ να κοινοποιούνται ανοικτά και οι αποφάσεις, στο μέτρο του δυνατού, να επεξηγούνται στους άμεσα και έμμεσα επηρεαζόμενους. Χωρίς αυτές τις πληροφορίες, μια απόφαση δεν είναι δυνατό να αμφισβητηθεί δεόντως. Δεν είναι πάντα εφικτή η επεξήγηση των λόγων για τους οποίους ένα μοντέλο έχει οδηγήσει σε ένα συγκεκριμένο αποτέλεσμα ή απόφαση (και ποιος συνδυασμός εισροών συνέβαλε σε αυτό). Αυτές οι περιπτώσεις αναφέρονται ως αλγόριθμοι «μαύρου κουτιού» και απαιτούν ιδιαίτερη προσοχή. Υπό αυτές τις συνθήκες, ενδέχεται να απαιτούνται άλλα μέτρα επεξηγησιμότητας (π.χ. ιχνηλασιμότητα, ελεγκσιμότητα και διαφανής επικοινωνία σχετικά με τις δυνατότητες του συστήματος), υπό την προϋπόθεση ότι το σύστημα στο σύνολό του σέβεται τα θεμελιώδη δικαιώματα. Ο βαθμός στον οποίο απαιτείται επεξηγησιμότητα εξαρτάται σημαντικά από το πλαίσιο και τη σοβαρότητα των συνεπειών, στην περίπτωση που το συγκεκριμένο αποτέλεσμα είναι εσφαλμένο ή κατ' άλλον τρόπο ανακριβές.<sup>33</sup>

### 2.3 Τριβές μεταξύ των αρχών

(54) Ενδέχεται να προκύψουν τριβές μεταξύ των ως άνω αρχών και δεν υπάρχει παγιωμένη λύση για την

<sup>31</sup> Τούτο αφορά την αρχή της αναλογικότητας (που αποτυπώνεται στο ρητό ότι «δεν θα πρέπει κανείς να κυνηγά κουνούπια με κανονιές»). Τα μέτρα που λαμβάνονται για την επίτευξη ενός στόχου (π.χ. τα μέτρα εξόρυξης δεδομένων που εφαρμόζονται για την πραγματοποίηση της λειτουργίας βελτιστοποίησης της ΤΝ) θα πρέπει να περιορίζονται σε αυτά που είναι απολύτως απαραίτητα. Επίσης, στην περίπτωση που πολλά μέτρα ανταγωνίζονται για την ικανοποίηση ενός στόχου, θα πρέπει να προτιμάται εκείνο που είναι λιγότερο επιζήμιο για τα θεμελιώδη δικαιώματα και τους δεοντολογικούς κανόνες (π.χ. οι προγραμματιστές της ΤΝ θα πρέπει πάντα να προτιμούν τα δεδομένα του δημόσιου τομέα έναντι των προσωπικών δεδομένων). Μπορεί επίσης να γίνει αναφορά στην αναλογικότητα μεταξύ χρήστη και εγκαταστάτη, λαμβανομένων υπόψη των δικαιωμάτων των εταιρειών (συμπεριλαμβανομένης της πνευματικής ιδιοκτησίας και της εμπιστευτικότητας), αφενός, και των δικαιωμάτων του χρήστη, αφετέρου.

<sup>32</sup> Συμπεριλαμβανομένου του δικαιώματος του συνεταιρίζεσθαι και της προσχώρησης σε συνδικαλιστικές ενώσεις σε ένα εργασιακό περιβάλλον, όπως προβλέπεται στο άρθρο 12 του Χάρτη των Θεμελιωδών Δικαιωμάτων της ΕΕ.

<sup>33</sup> Για παράδειγμα, οι ανακριβείς προτάσεις για αγορές που δημιουργούνται από ένα σύστημα ΤΝ εγείρουν ελάχιστους δεοντολογικούς προβληματισμούς, σε αντιδιαστολή με τα συστήματα ΤΝ που αξιολογούν αν ένα πρόσωπο που έχει καταδικαστεί για ποινικό αδίκημα θα πρέπει να αφεθεί ελεύθερο υπό όρους.

αντιμετώπισή τους. Σύμφωνα με τη θεμελιώδη δέσμευση της ΕΕ για συμμετοχή στις δημοκρατικές διαδικασίες, δίκαιη δίκη και ανοικτή πολιτική συμμετοχή, θα πρέπει να προσδιοριστούν μέθοδοι διαβούλευσης για την αντιμετώπιση των τριβών αυτού του είδους με λογοδοσία. Για παράδειγμα, σε διάφορους τομείς εφαρμογής, η αρχή της πρόληψης βλάβης και η αρχή της ανθρώπινης αυτονομίας ενδέχεται να συγκρούονται. Παραδείγματος χάρη, η χρήση των συστημάτων ΤΝ για «προληπτική αστυνόμευση», που ενδεχομένως να συμβάλλει στη μείωση της εγκληματικότητας, αλλά με τρόπους που συνεπάγονται δραστηριότητες επιτήρησης οι οποίες θίγουν την ατομική ελευθερία και την ιδιωτική ζωή. Επιπλέον, τα συνολικά οφέλη των συστημάτων της ΤΝ θα πρέπει να υπερβαίνουν σημαντικά τους προβλεπόμενους ατομικούς κινδύνους. Παρότι οι εν λόγω αρχές αναμφίβολα προσφέρουν καθοδήγηση για λύσεις, παραμένουν αφηρημένες δεοντολογικές επιταγές. Συνεπώς, οι επαγγελματίες της ΤΝ δεν είναι δυνατό να αναμένεται ότι θα είναι σε θέση να βρίσκουν τη σωστή λύση με βάση τις ανωτέρω αρχές, θα πρέπει όμως να προσεγγίζουν τα δεοντολογικά διλήμματα και τις αντισταθμιστικές ρυθμίσεις μέσω αιτιολογημένης ανάλυσης που βασίζεται στην τεκμηρίωση και όχι στη διαίσθηση ή την τυχαία διακριτική ευχέρεια. Ενδέχεται, ωστόσο, να ανακύψουν καταστάσεις στις οποίες να μην είναι δυνατό να υπάρξουν δεοντολογικά αποδεκτές αντισταθμιστικές ρυθμίσεις. Ορισμένα θεμελιώδη δικαιώματα και συσχετιζόμενες αρχές έχουν απόλυτο χαρακτήρα και δεν είναι δυνατό να αποτελούν αντικείμενο στάθμισης (π.χ. ανθρώπινη αξιοπρέπεια).

#### **Βασικές κατευθύνσεις που προκύπτουν από το κεφάλαιο I:**

- ✓ Ανάπτυξη, εγκατάσταση και χρήση των συστημάτων ΤΝ με τρόπους που τηρούν τις ακόλουθες δεοντολογικές αρχές: *σεβασμός της ανθρώπινης αυτονομίας, πρόληψη βλάβης, δικαιοσύνη και επεξηγησιμότητα*. Αναγνώριση και αντιμετώπιση των πιθανών τριβών μεταξύ αυτών των αρχών.
- ✓ Απόδοση ιδιαίτερης προσοχής σε καταστάσεις που αφορούν πιο ευάλωτες ομάδες όπως τα παιδιά, τα άτομα με αναπηρίες και άλλες ομάδες που βρίσκονται ανέκαθεν σε μειονεκτική θέση, ή αντιμετωπίζουν κίνδυνο αποκλεισμού, και/ή σε καταστάσεις που χαρακτηρίζονται από ασυμμετρικές εξουσίας ή πληροφόρησης, όπως μεταξύ εργοδοτών και εργαζομένων ή επιχειρήσεων και καταναλωτών<sup>34</sup>.
- ✓ Αναγνώριση και συνεκτίμηση του γεγονότος ότι ορισμένες εφαρμογές ΤΝ, αν και έχουν τη δυνατότητα να αποφέρουν πολλά ουσιαστικά οφέλη για τα άτομα και την κοινωνία, μπορεί επίσης να έχουν αρνητικές επιπτώσεις, συμπεριλαμβανομένων επιπτώσεων που ενδεχομένως να είναι δύσκολο να προβλεφθούν, να προσδιοριστούν ή να μετρηθούν (π.χ. στη δημοκρατία, στο κράτος δικαίου και στη διανεμητική δικαιοσύνη ή και στον ίδιο τον ανθρώπινο νου.) Λήψη επαρκών μέτρων για τον μετριασμό των εν λόγω κινδύνων, όταν κρίνεται σκόπιμο, και ανάλογα με το μέγεθος του κινδύνου.

## **II. Κεφάλαιο II: Πραγμάτωση της αξιόπιστης ΤΝ**

(55) Το παρόν κεφάλαιο παρέχει κατευθύνσεις σχετικά με την εφαρμογή και την πραγμάτωση της αξιόπιστης ΤΝ, μέσω ενός καταλόγου επτά απαιτήσεων που πρέπει να τηρούνται βάσει των αρχών που περιγράφονται στο κεφάλαιο I. Επιπλέον, παρουσιάζονται τεχνικές και μη τεχνικές μέθοδοι που είναι σήμερα διαθέσιμες για την εφαρμογή των συγκεκριμένων απαιτήσεων σε ολόκληρο τον κύκλο ζωής του συστήματος ΤΝ.

### **1. Απαιτήσεις της αξιόπιστης ΤΝ**

(56) Οι αρχές που περιγράφονται στο κεφάλαιο I θα πρέπει να μεταφραστούν σε συγκεκριμένες απαιτήσεις για την επίτευξη αξιόπιστης ΤΝ. Οι εν λόγω απαιτήσεις ισχύουν για τα διάφορα ενδιαφερόμενα μέρη που συμμετέχουν στον κύκλο ζωής των συστημάτων ΤΝ: προγραμματιστές, εγκαταστάτες και τελικούς χρήστες, καθώς και την ευρύτερη κοινωνία. Με τον όρο προγραμματιστές, εννοούμε τα άτομα που ερευνούν,

<sup>34</sup> Βλ. άρθρα 24 έως 27 του Χάρτη της ΕΕ σχετικά με τα δικαιώματα του παιδιού και των ηλικιωμένων, την ένταξη των ατόμων με αναπηρίες και τα δικαιώματα των εργαζομένων. Βλ. επίσης άρθρο 38 σχετικά με την προστασία του καταναλωτή.

σχεδιάζουν και/ή αναπτύσσουν συστήματα ΤΝ. Με τον όρο εγκαταστάτες, εννοούμε τους δημόσιους ή ιδιωτικούς οργανισμούς που χρησιμοποιούν συστήματα ΤΝ στις επιχειρηματικές τους διαδικασίες και προσφέρουν προϊόντα και υπηρεσίες σε άλλους. Οι τελικοί χρήστες είναι τα άτομα που ασχολούνται άμεσα ή έμμεσα με το σύστημα ΤΝ. Τέλος, η ευρύτερη κοινωνία περιλαμβάνει όλους τους άλλους που επηρεάζονται άμεσα ή έμμεσα από τα συστήματα ΤΝ.

(57) Οι διαφορετικές κατηγορίες ενδιαφερόμενων μερών έχουν διαφορετικούς ρόλους στη διασφάλιση της τήρησης των απαιτήσεων:

- a. οι προγραμματιστές θα πρέπει να εφαρμόζουν και να τηρούν τις απαιτήσεις κατά τις διαδικασίες σχεδιασμού και ανάπτυξης·
- b. οι εγκαταστάτες θα πρέπει να διασφαλίζουν ότι τα συστήματα που χρησιμοποιούν και τα προϊόντα και οι υπηρεσίες που προσφέρουν πληρούν τις απαιτήσεις·
- c. οι τελικοί χρήστες και η ευρύτερη κοινωνία θα πρέπει να ενημερώνονται σχετικά με τις συγκεκριμένες απαιτήσεις, ώστε να είναι σε θέση να ζητούν την τήρησή τους.

(58) Ο κατωτέρω κατάλογος απαιτήσεων δεν είναι εξαντλητικός<sup>35</sup>. Περιλαμβάνει συστημικές, ατομικές και κοινωνιακές πτυχές:

**1 Ανθρώπινη παρέμβαση και εποπτεία**

*Συμπεριλαμβανομένων των θεμελιωδών δικαιωμάτων, της ανθρώπινης παρέμβασης και της ανθρώπινης εποπτείας*

**2 Τεχνική στιβαρότητα και ασφάλεια**

*Συμπεριλαμβανομένων της ανθεκτικότητας σε επιθέσεις και της προστασίας, του εφεδρικού σχεδίου και της γενικής ασφάλειας, της ακρίβειας, της αξιοπιστίας και της αναπαραγωγιμότητας*

**3 Ιδιωτική ζωή και διακυβέρνηση των δεδομένων**

*Συμπεριλαμβανομένων του σεβασμού της ιδιωτικής ζωής, της ποιότητας και της ακεραιότητας των δεδομένων και της πρόσβασης στα δεδομένα*

**4 Διαφάνεια**

*Συμπεριλαμβανομένης της ανιχνευσιμότητας, της επεξηγησιμότητας και της επικοινωνίας*

**5 Πολυμορφία, απαγόρευση των διακρίσεων και δικαιοσύνη**

*Συμπεριλαμβανομένων της αποφυγής της αθέμιτης μεροληψίας, της προσβασιμότητας και του καθολικού σχεδιασμού και της συμμετοχής των ενδιαφερόμενων μερών*

**6 Κοινωνική και περιβαλλοντική ευημερία**

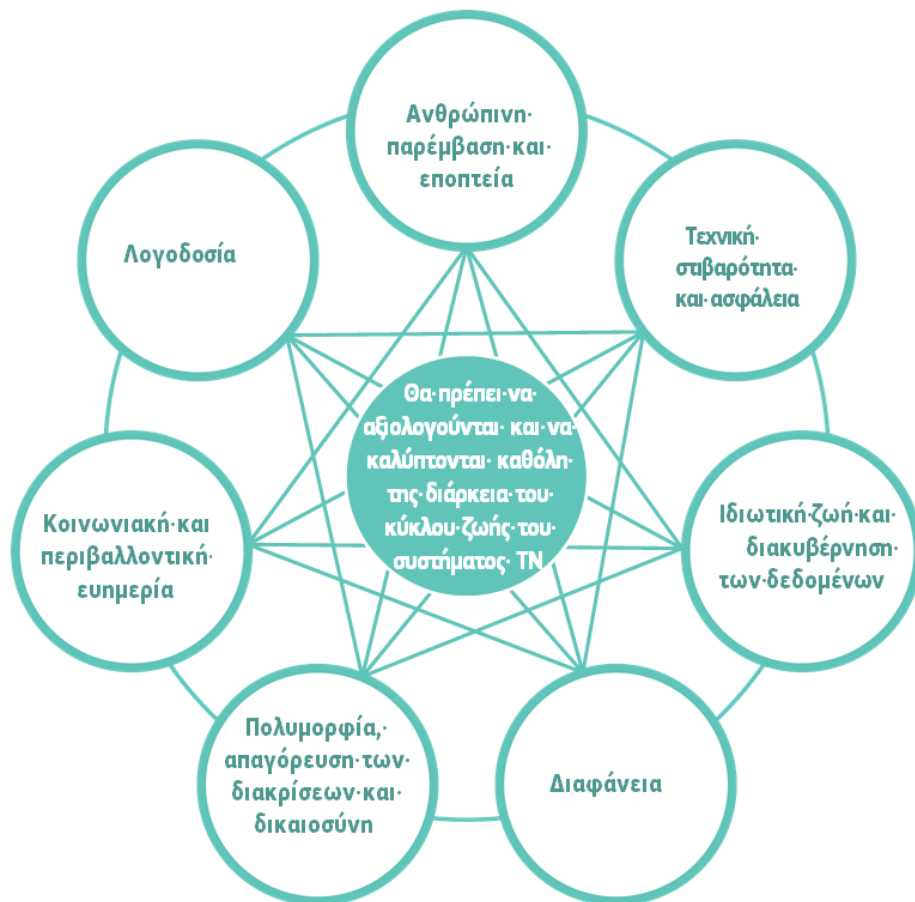
*Συμπεριλαμβανομένων της βιωσιμότητας και της φιλικότητας προς το περιβάλλον, των κοινωνικών επιπτώσεων, της κοινωνίας και της δημοκρατίας*

**7 Λογοδοσία**

*Συμπεριλαμβανομένων της ελεγχιμότητας, της ελαχιστοποίησης και της γνωστοποίησης των αρνητικών επιπτώσεων, των αντισταθμιστικών ρυθμίσεων και της έννομης προστασίας.*

---

<sup>35</sup> Χωρίς να επιβάλλουμε μια ιεράρχηση, παραθέτουμε τις παρακάτω αρχές με τρόπο που αντικατοπτρίζει τη σειρά εμφάνισης των αρχών και των δικαιωμάτων με τα οποία σχετίζονται στον Χάρτη της ΕΕ.



Σχήμα 2: Η αλληλεξάρτηση των επτά απαιτήσεων: όλες είναι εξίσου σημαντικές, αλληλοϋποστηρίζονται και θα πρέπει να εφαρμόζονται και να αξιολογούνται καθόλη τη διάρκεια του κύκλου ζωής του συστήματος TN

- (59) Παρότι όλες οι απαιτήσεις είναι εξίσου σημαντικές, το πλαίσιο και οι ενδεχόμενες τριβές μεταξύ τους θα πρέπει να λαμβάνονται υπόψη κατά την εφαρμογή τους σε διάφορους τομείς και κλάδους. Οι εν λόγω απαιτήσεις θα πρέπει να εφαρμόζονται καθόλη της διάρκεια του κύκλου ζωής του συστήματος TN, ανάλογα με την εκάστοτε εφαρμογή. Αν και οι περισσότερες απαιτήσεις ισχύουν για όλα τα συστήματα TN, ιδιαίτερη προσοχή δίνεται σε εκείνες που επηρεάζουν άμεσα ή έμμεσα τα άτομα. Επομένως, για ορισμένες εφαρμογές (για παράδειγμα σε βιομηχανικά περιβάλλοντα), ενδέχεται να έχουν μικρότερη σημασία.
- (60) Οι ανωτέρω απαιτήσεις περιλαμβάνουν στοιχεία που σε ορισμένες περιπτώσεις αποτυπώνονται ήδη στην κείμενη νομοθεσία. Επαναλαμβάνουμε ότι —σύμφωνα με την πρώτη συνιστώσα της αξιόπιστης TN— αποτελεί ευθύνη των προγραμματιστών και των εγκαταστατών συστημάτων TN να διασφαλίζουν ότι συμμορφώνονται με τις νομικές τους υποχρεώσεις, τόσο όσον αφορά τους οριζόντιους εφαρμοστέους κανόνες όσο και τους κανόνες που αφορούν τον εκάστοτε τομέα.
- (61) Στις παραγράφους που ακολουθούν, η κάθε απαίτηση παρουσιάζεται αναλυτικότερα.

### 1. Ανθρώπινη παρέμβαση και εποπτεία

- (62) Τα συστήματα TN θα πρέπει να στηρίζουν την ανθρώπινη αυτονομία και τη λήψη αποφάσεων, όπως προβλέπεται στην αρχή του σεβασμού της ανθρώπινης αυτονομίας. Η απαίτηση αυτή προϋποθέτει ότι τα συστήματα TN θα πρέπει και να λειτουργούν ως εργαλεία ανάπτυξης μιας δημοκρατικής, ευημερούσας και ισότιμης κοινωνίας, υποστηρίζοντας την παρέμβαση του χρήστη, αλλά και να προάγουν τα θεμελιώδη δικαιώματα, καθώς και να αφήνει περιθώριο για ανθρώπινη εποπτεία.

- (63) **Θεμελιώδη δικαιώματα.** Όπως πολλές τεχνολογίες, τα συστήματα TN μπορούν εξίσου να ενισχύουν ή να καταστρατηγούν τα θεμελιώδη δικαιώματα. Μπορούν να ωφελήσουν τους ανθρώπους, για παράδειγμα, διευκολύνοντάς τους να παρακολουθούν τα προσωπικά τους δεδομένα ή αυξάνοντας την προσβασιμότητα στην εκπαίδευση, υποστηρίζοντας συνεπώς το δικαίωμά τους στην εκπαίδευση. Ωστόσο, δεδομένης της εμβέλειας και των ικανοτήτων των συστημάτων TN, μπορούν επίσης να επηρεάσουν αρνητικά τα θεμελιώδη δικαιώματα. Σε καταστάσεις που ενέχουν κίνδυνους αυτού του είδους, θα πρέπει να διενεργείται εκτίμηση επιπτώσεων για τα θεμελιώδη δικαιώματα. Αυτή θα πρέπει να γίνεται πριν από την ανάπτυξή τους και να συμπεριλαμβάνει αξιολόγηση του κατά πόσον οι συγκεκριμένοι κίνδυνοι είναι δυνατό να μειωθούν ή να δικαιολογηθούν ως απαραίτητοι, σε μια δημοκρατική κοινωνία, προκειμένου να γίνονται σεβαστά τα δικαιώματα και οι ελευθερίες των άλλων. Επιπλέον, θα πρέπει να δημιουργηθούν μηχανισμοί για τη λήψη εξωτερικής αναπληροφόρησης σχετικά με συστήματα TN που ενδεχομένως παραβιάζουν τα θεμελιώδη δικαιώματα.
- (64) **Ανθρώπινη παρέμβαση.** Οι χρήστες θα πρέπει να είναι σε θέση να λαμβάνουν τεκμηριωμένες αυτόνομες αποφάσεις σχετικά με τα συστήματα TN. Θα πρέπει να έχουν στη διάθεσή τους τις γνώσεις και τα εργαλεία για να κατανοούν και να αλληλεπιδρούν με τα συστήματα TN σε ικανοποιητικό βαθμό και, όπου είναι δυνατό, να έχουν τη δυνατότητα αυτοαξιολόγησης ή αμφισβήτησης του συστήματος μέσα σε λογικά πλαίσια. Τα συστήματα TN θα πρέπει να βοηθούν τα άτομα να κάνουν καλύτερες και πιο τεκμηριωμένες επιλογές σύμφωνα με τους στόχους τους. Τα συστήματα TN είναι δυνατό ορισμένες φορές να εγκαθίστανται έτσι ώστε να διαμορφώνουν και να επηρεάζουν την ανθρώπινη συμπεριφορά μέσω μηχανισμών που ενδεχομένως να είναι δύσκολο να εντοπιστούν,, καθώς μπορεί να εκμεταλλευτούν υποσυνείδητες διαδικασίες, συμπεριλαμβανομένων διαφόρων μορφών αθέμιτης χειραγώγησης, εξαπάτησης, αγελοποίησης και κοινωνικού προγραμματισμού, δηλαδή πρακτικών που ενδέχεται να συνιστούν απειλή για την ατομική αυτονομία. Στο επίκεντρο της λειτουργίας του συστήματος θα πρέπει να βρίσκεται η συνολική αρχή της αυτονομίας του χρήστη. Ιδιαίτερη σημασία στο πλαίσιο αυτό έχει το δικαίωμα των χρηστών να μην υπόκεινται σε απόφαση που λαμβάνεται αποκλειστικά βάσει αυτοματοποιημένης επεξεργασίας, εάν η εν λόγω απόφαση παράγει έννομα αποτελέσματα για τους χρήστες ή τους επηρεάζει σημαντικά με παρόμοιο τρόπο<sup>36</sup>.
- (65) **Ανθρώπινη εποπτεία.** Η ανθρώπινη εποπτεία συμβάλλει στη διασφάλιση ότι ένα σύστημα TN δεν υπονομεύει την ανθρώπινη αυτονομία ούτε προκαλεί άλλες δυσμενείς επιπτώσεις. Η εποπτεία μπορεί να επιτευχθεί μέσω μηχανισμών διακυβέρνησης, όπως η προσέγγιση στην οποία ο άνθρωπος είναι εκείνος που παρεμβαίνει (human-in-the-loop - HITL), εκείνος που εποπτεύει (human-on-the-loop - HOTL) ή εκείνος που ελέγχει (human-in-command - HIC). Η προσέγγιση της ανθρώπινης παρέμβασης (HITL) αναφέρεται στην ανθρώπινη παρέμβαση σε κάθε κύκλο λήψης αποφάσεων του συστήματος, κάτι που σε πολλές περιπτώσεις δεν είναι ούτε εφικτό ούτε και επιθυμητό. Η προσέγγιση της ανθρώπινης εποπτείας (HOTL) αναφέρεται στην ικανότητα ανθρώπινης παρέμβασης κατά τη διάρκεια του κύκλου σχεδιασμού του συστήματος και παρακολούθησης της λειτουργίας του. Η προσέγγιση του ανθρώπινου ελέγχου (HIC) αναφέρεται στην ικανότητα εποπτείας της συνολικής δραστηριότητας του συστήματος TN (συμπεριλαμβανομένου του ευρύτερου οικονομικού, κοινωνικού, νομικού και δεοντολογικού αντίκτυπού του), καθώς και στην ικανότητα λήψης αποφάσεων σχετικά με το πότε και πώς χρησιμοποιείται το σύστημα σε κάθε συγκεκριμένη περίπτωση. Αυτή η προσέγγιση μπορεί να περιλαμβάνει την απόφαση να μη χρησιμοποιηθεί ένα σύστημα TN σε μια συγκεκριμένη περίπτωση, να οριστούν επίπεδα ανθρώπινης διακριτικής ευχέρειας κατά τη χρήση του συστήματος ή να εξασφαλιστεί η δυνατότητα παράκαμψης μιας απόφασης που ελήφθη από το σύστημα. Επιπλέον, πρέπει να διασφαλιστεί ότι οι δημόσιες αρχές έχουν την ικανότητα να ασκούν εποπτεία σύμφωνα με την εντολή που έχουν λάβει. Είναι δυνατό να απαιτηθούν μηχανισμοί εποπτείας σε διάφορους βαθμούς για την υποστήριξη άλλων μέτρων ασφάλειας και ελέγχου,

<sup>36</sup>

Μπορεί να γίνει αναφορά στο άρθρο 22 του ΓΚΠΔ, όπου έχει ήδη κατοχυρωθεί το συγκεκριμένο δικαίωμα.

ανάλογα με τον τομέα εφαρμογής του συστήματος TN και τον πιθανό κίνδυνο. Με όλες τις άλλες παραμέτρους σταθερές, όσο μικρότερη είναι η εποπτεία που μπορεί να ασκήσει ένας άνθρωπος σε ένα σύστημα TN, τόσο εκτενέστερες οι δοκιμές και αυστηρότερη η διακυβέρνηση που απαιτείται.

## 2. Τεχνική στιβαρότητα και ασφάλεια

- (66) Σημαντική συνιστώσα για την επίτευξη αξιόπιστης TN είναι η τεχνική στιβαρότητα, η οποία συνδέεται στενά με την *αρχή της πρόληψης βλάβης*. Η τεχνική στιβαρότητα προϋποθέτει την ανάπτυξη συστημάτων TN με προληπτική προσέγγιση των κινδύνων και με τέτοιο τρόπο ώστε να συμπεριφέρονται αξιόπιστα με τον αναμενόμενο τρόπο, ενώ παράλληλα ελαχιστοποιείται η ακούσια και μη αναμενόμενη βλάβη και αποτρέπεται η απaráδεκτη βλάβη. Το ίδιο θα πρέπει να ισχύει επίσης για πιθανές αλλαγές στο λειτουργικό τους περιβάλλον ή για την παρουσία άλλων πρακτόρων (ανθρώπινων και τεχνητών) που ενδέχεται να αλληλεπιδρούν κατ' αντιπαράθεση με το σύστημα. Επιπλέον, θα πρέπει να διασφαλίζεται η σωματική και διανοητική ακεραιότητα των ανθρώπων.
- (67) **Ανθεκτικότητα σε επιθέσεις και προστασία.** Τα συστήματα TN, όπως όλα τα συστήματα λογισμικού, θα πρέπει να προστατεύονται από ευπάθειες που ενδέχεται να καθιστούν δυνατή την εκμετάλλευσή τους από αντίπαλες πλευρές, π.χ. δικτυοπαραβίαση (χακάρισμα). Οι επιθέσεις ενδέχεται να αφορούν τα δεδομένα (δηλητηρίαση δεδομένων), το μοντέλο (διαρροή μοντέλου) ή την υποκείμενη υποδομή, τόσο το λογισμικό όσο και το υλισμικό. Εάν ένα σύστημα TN δεχθεί επίθεση, π.χ. επιθέσεις από αντίπαλες πλευρές, τα δεδομένα, καθώς και η συμπεριφορά του συστήματος είναι δυνατό να τροποποιηθούν, με αποτέλεσμα το σύστημα να λαμβάνει διαφορετικές αποφάσεις ή να παύσει εντελώς να λειτουργεί. Τα συστήματα και τα δεδομένα είναι δυνατό επίσης να αλλοιωθούν με κακόβουλη πρόθεση ή με έκθεση σε απροσδόκητες καταστάσεις. Οι ανεπαρκείς διαδικασίες προστασίας είναι επίσης δυνατό να οδηγήσουν σε εσφαλμένες αποφάσεις ή ακόμα και σε σωματική βλάβη. Προκειμένου τα συστήματα TN να θεωρούνται ασφαλή,<sup>37</sup> θα πρέπει να λαμβάνονται υπόψη ενδεχόμενες ακούσιες εφαρμογές της TN (π.χ. εφαρμογές διπλής χρήσης), καθώς και ενδεχόμενη κατάχρηση ενός συστήματος TN από κακόβουλους παράγοντες, ενώ επίσης θα πρέπει να λαμβάνονται μέτρα για την πρόληψη και τον μετριασμό τους.<sup>38</sup>
- (68) **Εφεδρικό σχέδιο και γενική ασφάλεια.** Τα συστήματα TN θα πρέπει να διαθέτουν εγγυήσεις για την εφαρμογή εφεδρικού σχεδίου σε περίπτωση που προκύψουν προβλήματα. Αυτό μπορεί να σημαίνει ότι τα συστήματα TN μεταβαίνουν από μια στατιστική διαδικασία σε μια διαδικασία που βασίζεται σε κανόνες ή ότι απαιτείται η συμμετοχή ανθρώπινου χειριστή πριν συνεχίσουν τη δράση τους.<sup>39</sup> Πρέπει να διασφαλίζεται ότι το σύστημα θα κάνει αυτό που οφείλει χωρίς να βλάπτει τα έμβια όντα ή το περιβάλλον. Τούτο περιλαμβάνει την ελαχιστοποίηση των ακούσιων συνεπειών και σφαλμάτων. Επιπλέον, θα πρέπει να θεσπιστούν διαδικασίες για την αποσαφήνιση και την αξιολόγηση των δυνητικών κινδύνων που σχετίζονται με τη χρήση συστημάτων TN σε διάφορους τομείς εφαρμογής. Το επίπεδο των απαιτούμενων μέτρων ασφαλείας εξαρτάται από το μέγεθος του κινδύνου που ενέχει ένα σύστημα TN, ο οποίος με τη σειρά του εξαρτάται από τις δυνατότητες του συστήματος. Στις περιπτώσεις στις οποίες είναι δυνατό να προβλεφθεί ότι η διαδικασία ανάπτυξης ή το σύστημα καθεαυτό θα ενέχουν ιδιαίτερα υψηλό κίνδυνο, είναι καίριας σημασίας να αναπτύσσονται και να δοκιμάζονται προληπτικά τα μέτρα ασφαλείας.
- (69) **Ακρίβεια.** Η ακρίβεια σχετίζεται με την ικανότητα ενός συστήματος TN να διατυπώνει σωστές κρίσεις, για παράδειγμα για να ταξινομεί σωστά τις πληροφορίες στις κατάλληλες κατηγορίες ή την ικανότητά του να

<sup>37</sup> Βλ. π.χ. τους προβληματισμούς στο σημείο 2.7 του Συντονισμένου Σχεδίου της Ευρωπαϊκής Ένωσης για την Τεχνητή Νοημοσύνη.

<sup>38</sup> Για την προστασία των συστημάτων TN, ενδέχεται να είναι επιτακτική η ανάγκη για τη δυνατότητα ανάπτυξης ενός «ενάρετου κύκλου» στην έρευνα και την ανάπτυξη μεταξύ της κατανόησης των επιθέσεων, της ανάπτυξης κατάλληλων προστατευτικών μέτρων και της βελτίωσης των μεθοδολογιών αξιολόγησης. Προκειμένου να επιτευχθεί αυτό, θα πρέπει να προωθηθεί η σύγκλιση μεταξύ της κοινότητας TN και της κοινότητας προστασίας. Επιπλέον, αποτελεί ευθύνη όλων των εμπλεκόμενων φορέων να δημιουργήσουν κοινά πρότυπα διασυνοριακής προστασίας και ασφαλείας και να διαμορφώσουν ένα περιβάλλον αμοιβαίας εμπιστοσύνης, προωθώντας τη διεθνή συνεργασία. Για πιθανά μέτρα, βλ. Malicious Use of AI (Avin S., Brundage M., κ.ά., 2018).

<sup>39</sup> Θα πρέπει να λαμβάνονται υπόψη επίσης και σενάρια στα οποία η ανθρώπινη παρέμβαση δεν είναι άμεσα δυνατή.



κάνει σωστές προβλέψεις, συστάσεις ή να λαμβάνει αποφάσεις που βασίζονται σε δεδομένα ή μοντέλα. Μια σαφής και καλά διαμορφωμένη διαδικασία ανάπτυξης και αξιολόγησης μπορεί να υποστηρίξει, να μετριάσει και να διορθώσει ακούσιους κινδύνους που προκύπτουν από ανακριβείς προβλέψεις. Όταν δεν είναι δυνατή η αποφυγή περιστασιακών ανακριβών προβλέψεων, είναι σημαντικό το σύστημα να είναι σε θέση να υποδεικνύει πόσο πιθανό είναι να προκύψουν τα εν λόγω σφάλματα. Ένα υψηλό επίπεδο ακρίβειας είναι ιδιαίτερα σημαντικό σε καταστάσεις όπου το σύστημα TN επηρεάζει άμεσα τις ανθρώπινες ζωές.

- (70) **Αξιοπιστία και αναπαραγωγιμότητα.** Είναι ζωτικής σημασίας τα αποτελέσματα των συστημάτων TN να μπορούν να αναπαραχθούν, αλλά να είναι και αξιόπιστα. Ένα αξιόπιστο σύστημα TN είναι αυτό που λειτουργεί σωστά με μια σειρά εισροών και σε μια σειρά καταστάσεων. Αυτό είναι απαραίτητο για τον ενδελεχή έλεγχο ενός συστήματος TN και την αποφυγή τυχόν ακούσιων βλαβών. Η αναπαραγωγιμότητα περιγράφει κατά πόσον ένα πείραμα TN παρουσιάζει την ίδια συμπεριφορά όταν επαναλαμβάνεται υπό τις ίδιες συνθήκες. Τούτο επιτρέπει στους επιστήμονες και στους υπεύθυνους χάραξης πολιτικής να περιγράψουν με ακρίβεια τις ενέργειες των συστημάτων TN. Τα αρχεία αναπαραγωγής<sup>40</sup> μπορούν να διευκολύνουν τη διαδικασία ελέγχου και αναπαραγωγής συμπεριφορών.

### 3. Ιδιωτική ζωή και διακυβέρνηση των δεδομένων

- (71) Στενά συνδεδεμένη με την αρχή της πρόληψης βλάβης, η ιδιωτική ζωή αποτελεί θεμελιώδες δικαίωμα το οποίο επηρεάζεται σε μεγάλο βαθμό από τα συστήματα TN. Η πρόληψη βλάβης στην ιδιωτική ζωή απαιτεί επίσης επαρκή διακυβέρνηση δεδομένων που καλύπτει την ποιότητα και την ακεραιότητα των χρησιμοποιούμενων δεδομένων, τη συνάφειά τους σε σχέση με τον τομέα στον οποίο τα συστήματα TN θα εγκαθίστανται, τα πρωτόκολλα πρόσβασής τους και την ικανότητα επεξεργασίας δεδομένων κατά τρόπο που προστατεύει την ιδιωτική ζωή.
- (72) **Προστασία ιδιωτικής ζωής και δεδομένων.** Τα συστήματα TN θα πρέπει να εγγυώνται την προστασία της ιδιωτικής ζωής και των δεδομένων καθόλη τη διάρκεια του κύκλου ζωής ενός συστήματος<sup>41</sup>. Αυτό περιλαμβάνει τις πληροφορίες που παρέχονται αρχικά από τον χρήστη, καθώς και τις πληροφορίες που δημιουργούνται σχετικά με τον χρήστη κατά τη διάρκεια της αλληλεπίδρασής του με το σύστημα (π.χ. τα αποτελέσματα που παράγει το σύστημα TN για συγκεκριμένους χρήστες ή τον τρόπο με τον οποίο οι χρήστες ανταποκρίθηκαν σε συγκεκριμένες συστάσεις). Η ψηφιακή καταγραφή της ανθρώπινης συμπεριφοράς μπορεί να επιτρέψει στα συστήματα TN να συνάγουν όχι μόνο τις προτιμήσεις του ατόμου, αλλά και τον σεξουαλικό προσανατολισμό, την ηλικία, το φύλο, τις θρησκευτικές ή πολιτικές απόψεις του. Προκειμένου η διαδικασία συλλογής δεδομένων να εμπνέει εμπιστοσύνη στους χρήστες, θα πρέπει να διασφαλίζεται ότι τα δεδομένα που συλλέγονται σχετικά με αυτούς δεν θα χρησιμοποιηθούν για αθέμιτες ή παράνομες διακρίσεις απέναντί τους.
- (73) **Ποιότητα και ακεραιότητα των δεδομένων.** Η ποιότητα των χρησιμοποιούμενων συνόλων δεδομένων είναι υψίστης σημασίας για τις επιδόσεις των συστημάτων TN. Τα δεδομένα που συλλέγονται μπορεί να περιέχουν κοινωνικά κατασκευασμένες προκαταλήψεις ή να περιέχουν ανακριβείς, σφάλματα και λάθη. Ο κίνδυνος αυτός πρέπει να αντιμετωπίζεται πριν από την εκπαίδευση με οποιοδήποτε σύνολο δεδομένων. Επιπλέον, πρέπει να διασφαλίζεται η ακεραιότητα των δεδομένων. Η τροφοδοσία ενός συστήματος TN με κακόβουλα δεδομένα μπορεί να αλλοιώσει τη συμπεριφορά του, ιδίως με αυτοεκπαιδευόμενα συστήματα. Οι διεργασίες και τα σύνολα δεδομένων που χρησιμοποιούνται πρέπει να υποβάλλονται σε δοκιμές και να τεκμηριώνονται σε κάθε στάδιο, όπως για παράδειγμα στα στάδια του σχεδιασμού, της εκπαίδευσης, των δοκιμών και της εγκατάστασης. Το ίδιο θα πρέπει να ισχύει και για τα συστήματα TN που δεν αναπτύσσονται στο εσωτερικό

<sup>40</sup> Αυτό αφορά αρχεία που θα αναπαράγουν κάθε βήμα της διαδικασίας ανάπτυξης του συστήματος TN, από την έρευνα και την αρχική συλλογή δεδομένων έως τα αποτελέσματα.

<sup>41</sup> Μπορεί να γίνει αναφορά στους υφιστάμενους νόμους για την προστασία της ιδιωτικής ζωής, όπως ο ΓΚΠΔ ή ο επικείμενος κανονισμός για την προστασία της ιδιωτικής ζωής στις ηλεκτρονικές επικοινωνίες (κανονισμός ePrivacy).

ενός οργανισμού, αλλά αγοράζονται από τρίτους.

- (74) **Πρόσβαση σε δεδομένα.** Σε κάθε συγκεκριμένο οργανισμό που χειρίζεται δεδομένα ατόμων (είτε πρόκειται για χρήστες του συστήματος είτε όχι), θα πρέπει να εφαρμόζονται πρωτόκολλα δεδομένων που διέπουν την πρόσβαση σε δεδομένα. Τα εν λόγω πρωτόκολλα θα πρέπει να προσδιορίζουν τους χρήστες που μπορούν να έχουν πρόσβαση στα δεδομένα και τις συνθήκες υπό τις οποίες αποκτούν πρόσβαση. Θα πρέπει να επιτρέπεται η πρόσβαση στα δεδομένα ενός ατόμου μόνο σε δεόντως ειδικευμένο προσωπικό με τη σχετική αρμοδιότητα και ανάγκη.

#### **4. Διαφάνεια**

- (75) Η συγκεκριμένη απαίτηση συνδέεται στενά με την αρχή της επεξηγησιμότητας και περιλαμβάνει τη διαφάνεια των στοιχείων που σχετίζονται με ένα σύστημα TN: τα δεδομένα, το σύστημα και τα επιχειρηματικά μοντέλα.
- (76) **Ιχνηλασιμότητα.** Τα σύνολα δεδομένων και οι διαδικασίες που οδηγούν στην απόφαση του συστήματος TN, συμπεριλαμβανομένων εκείνων που αφορούν τη συλλογή και την επισήμανση των δεδομένων, καθώς και οι αλγόριθμοι που χρησιμοποιούνται, θα πρέπει να τεκμηριώνονται στο βέλτιστο δυνατό επίπεδο ώστε να καθίσταται δυνατή η ιχνηλασιμότητα και η αύξηση της διαφάνειας. Τούτο ισχύει και για τις αποφάσεις που λαμβάνονται από το σύστημα TN. Έτσι καθίσταται δυνατός ο εντοπισμός των λόγων για τους οποίους μια απόφαση ενός συστήματος TN ήταν εσφαλμένη, κάτι που με τη σειρά του θα μπορούσε να συμβάλει στην αποφυγή μελλοντικών λαθών. Επομένως, η ιχνηλασιμότητα διευκολύνει την ελεγχσιμότητα και την επεξηγησιμότητα.
- (77) **Επεξηγησιμότητα.** Η επεξηγησιμότητα αφορά την δυνατότητα επεξήγησης τόσο των τεχνικών διεργασιών ενός συστήματος TN όσο και των σχετικών ανθρώπινων αποφάσεων (π.χ. τομείς εφαρμογής ενός συστήματος TN). Η τεχνική επεξηγησιμότητα προϋποθέτει τη δυνατότητα κατανόησης και ανίχνευσης, από ανθρώπους, των αποφάσεων που λαμβάνονται από ένα σύστημα TN. Επιπλέον, ενδέχεται να χρειαστούν αντισταθμιστικές ρυθμίσεις όσον αφορά τη βελτίωση της επεξηγησιμότητας του συστήματος (κάτι που μπορεί να περιορίσει την ακρίβειά του) ή την αύξηση της ακρίβειάς του (εις βάρος της επεξηγησιμότητας). Σε κάθε περίπτωση που ένα σύστημα TN έχει σημαντικό αντίκτυπο στη ζωή των ανθρώπων, θα πρέπει να είναι δυνατό να αξιώνεται η παροχή κατάλληλων επεξηγήσεων σχετικά με τη διαδικασία λήψης αποφάσεων του συστήματος TN. Οι επεξηγήσεις αυτές θα πρέπει να παρέχονται εγκαίρως και να προσαρμόζονται στην εμπειρογνωμοσύνη του ενδιαφερόμενου μέρους (π.χ. μη ειδήμων, ρυθμιστικός φορέας ή ερευνητής). Επιπλέον, θα πρέπει να παρέχονται επεξηγήσεις σχετικά με τον βαθμό στον οποίο ένα σύστημα TN επηρεάζει και διαμορφώνει τη διαδικασία λήψης αποφάσεων ενός οργανισμού, τις σχεδιαστικές επιλογές του συστήματος, καθώς και το σκεπτικό της εγκατάστασής του (ώστε να διασφαλίζεται η διαφάνεια του επιχειρηματικού μοντέλου).
- (78) **Επικοινωνία.** Τα συστήματα TN δεν πρέπει να παρουσιάζονται ως άνθρωποι στον χρήστη. Οι άνθρωποι έχουν το δικαίωμα να ενημερώνονται ότι αλληλεπιδρούν με ένα σύστημα TN. Τούτο συνεπάγεται ότι τα συστήματα TN θα πρέπει να είναι αναγνωρίσιμα ως τέτοια. Επιπλέον, θα πρέπει να παρέχεται η δυνατότητα να παρακαμφθεί η συγκεκριμένη αλληλεπίδραση προς όφελος της ανθρώπινης αλληλεπίδρασης, όπου αυτό είναι αναγκαίο, προκειμένου να διασφαλίζεται η τήρηση των θεμελιωδών δικαιωμάτων. Πέραν αυτού, οι δυνατότητες και οι περιορισμοί του συστήματος TN θα πρέπει να κοινοποιούνται στους επαγγελματίες στον τομέα της TN ή στους τελικούς χρήστες με τρόπο κατάλληλο ανάλογα με την εκάστοτε περίπτωση χρήσης. Η κοινοποίηση αυτή θα μπορούσε να περιλαμβάνει την ανακοίνωση του επιπέδου ακρίβειας του συστήματος TN, καθώς και τους περιορισμούς του.

#### **5. Πολυμορφία, απαγόρευση των διακρίσεων και δικαιοσύνη**

- (79) Προκειμένου να επιτύχουμε αξιόπιστη ΤΝ, ολόκληρος ο κύκλος ζωής του συστήματος ΤΝ θα πρέπει να χαρακτηρίζεται από συμπεριληπτικότητα και πολυμορφία. Εκτός από την εξέταση και τη συμμετοχή όλων των επηρεαζόμενων μερών καθόλη τη διάρκεια της διαδικασίας, τούτο συνεπάγεται επίσης την εξασφάλιση ισότιμης πρόσβασης μέσω συμπεριληπτικών διαδικασιών σχεδιασμού, καθώς και μέσω της ίσης μεταχείρισης. Η συγκεκριμένη απαίτηση συνδέεται στενά με την αρχή της δικαιοσύνης.
- (80) **Αποφυγή αθέμιτης μεροληψίας.** Τα σύνολα δεδομένων που χρησιμοποιούνται από συστήματα ΤΝ (τόσο για την εκπαίδευση όσο και για τη λειτουργία τους) ενδέχεται να περιλαμβάνουν εκ παραδρομής προϋπάρχοντα σφάλματα μεροληψίας, να είναι ελλιπή και να βασίζονται σε ακατάλληλα μοντέλα διακυβέρνησης. Η συνέχιση των μεροληψιών αυτού του είδους θα μπορούσε να οδηγήσει σε ακούσιες άμεσες (και έμμεσες) προκαταλήψεις και διακρίσεις<sup>42</sup> κατά ορισμένων ομάδων ή ανθρώπων, επιδεινώνοντας ενδεχομένως τις υφιστάμενες προκαταλήψεις και την περιθωριοποίηση. Βλάβη μπορεί επίσης να προκύψει από την εκούσια εκμετάλλευση της μεροληψίας (των καταναλωτών) ή από την άσκηση αθέμιτου ανταγωνισμού, όπως η ομογενοποίηση των τιμών μέσω συμπράξεων ή μη διαφάνειας της αγοράς.<sup>43</sup> Η αναγνωρίσιμη και διακριτική μεροληψία θα πρέπει να παρακάμπτεται κατά τη φάση συλλογής, όπου είναι δυνατό. Αθέμιτη μεροληψία ενδέχεται να παρατηρείται επίσης στον τρόπο ανάπτυξης των συστημάτων ΤΝ (π.χ. προγραμματισμός αλγορίθμων). Η μεροληψία αυτή θα μπορούσε να αντιμετωπιστεί με την εφαρμογή διαδικασιών εποπτείας για την ανάλυση και την αντιμετώπιση του σκοπού, των περιορισμών, των απαιτήσεων και των αποφάσεων του συστήματος κατά τρόπο σαφή και διαφανή. Επιπλέον, η πρόσληψη ατόμων με διαφορετικές καταβολές, διαφορετικούς πολιτισμούς και από διαφορετικά επιστημονικά πεδία μπορεί να εξασφαλίσει πολυφωνία απόψεων και θα πρέπει να ενθαρρύνεται.
- (81) **Προσβασιμότητα και καθολικός σχεδιασμός.** Ιδίως στις σχέσεις επιχείρησης με καταναλωτή, τα συστήματα θα πρέπει να είναι προσανατολισμένα στον χρήστη και να σχεδιάζονται με τρόπο που να επιτρέπει σε όλους τους ανθρώπους να χρησιμοποιούν προϊόντα ή υπηρεσίες ΤΝ, ανεξάρτητα από την ηλικία, το φύλο, τις ικανότητες ή τα χαρακτηριστικά τους. Η προσβασιμότητα στη συγκεκριμένη τεχνολογία για τα άτομα με αναπηρίες, όλων των κοινωνικών ομάδων, έχει ιδιαίτερη σημασία. Τα συστήματα ΤΝ δεν θα πρέπει να έχουν ενιαία προσέγγιση και θα πρέπει να λαμβάνουν υπόψη τις αρχές του καθολικού σχεδιασμού<sup>44</sup> που απευθύνεται σε όσο το δυνατόν ευρύτερο φάσμα χρηστών, σύμφωνα με τα σχετικά πρότυπα προσβασιμότητας<sup>45</sup>. Κάτι τέτοιο θα καταστήσει δυνατή την ισότιμη πρόσβαση και την ενεργή συμμετοχή όλων των ανθρώπων στις υφιστάμενες και αναδυόμενες ανθρώπινες δραστηριότητες μέσω υπολογιστών και σε σχέση με τις τεχνολογίες υποβοήθησης.<sup>46</sup>
- (82) **Συμμετοχή των ενδιαφερόμενων μερών.** Προκειμένου να αναπτυχθούν αξιόπιστα συστήματα ΤΝ, συνιστάται να πραγματοποιούνται διαβουλεύσεις με τα ενδιαφερόμενα μέρη που ενδέχεται να επηρεάζονται άμεσα ή έμμεσα από το σύστημα καθόλη τη διάρκεια του κύκλου ζωής του. Παράλληλα, κρίνεται επωφελές να ζητείται τακτική αναπληροφόρηση ακόμη και μετά την εγκατάσταση, καθώς και να δημιουργούνται μακροπρόθεσμοι μηχανισμοί για τη συμμετοχή των ενδιαφερόμενων μερών, για παράδειγμα με τη διασφάλιση της ενημέρωσης, διαβούλευσης και συμμετοχής των εργαζομένων καθόλη τη διάρκεια της εφαρμογής των συστημάτων ΤΝ σε οργανισμούς.

## 6. Κοινωνική και περιβαλλοντική ευημερία

<sup>42</sup> Για τον ορισμό των άμεσων και έμμεσων διακρίσεων βλ., για παράδειγμα, το άρθρο 2 της οδηγίας 2000/78/ΕΚ του Συμβουλίου, της 27ης Νοεμβρίου 2000, για τη διαμόρφωση γενικού πλαισίου για την ίση μεταχείριση στην απασχόληση και την εργασία. Βλ. επίσης άρθρο 21 του Χάρτη των Θεμελιωδών Δικαιωμάτων της ΕΕ.

<sup>43</sup> Βλ. έγγραφο του Οργανισμού Θεμελιωδών Δικαιωμάτων της Ευρωπαϊκής Ένωσης: «BigData: Discrimination in data-supported decision making (2018)» <http://fra.europa.eu/en/publication/2018/big-data-discrimination>.

<sup>44</sup> Το άρθρο 42 της οδηγίας για τις δημόσιες συμβάσεις απαιτεί οι τεχνικές προδιαγραφές να λαμβάνουν υπόψη την προσβασιμότητα και τον σχεδιασμό για όλους.

<sup>45</sup> Για παράδειγμα, το πρότυπο EN 301 549.

<sup>46</sup> Η συγκεκριμένη απαίτηση συνδέεται με τη σύμβαση των Ηνωμένων Εθνών για τα δικαιώματα των ατόμων με αναπηρίες.

- (83) Σύμφωνα με τις αρχές της δικαιοσύνης και της πρόληψης βλάβης, η ευρύτερη κοινωνία, τα άλλα ευαίσθητα όντα και το περιβάλλον θα πρέπει επίσης να θεωρούνται ενδιαφερόμενα μέρη σε ολόκληρο τον κύκλο ζωής των συστημάτων TN. Θα πρέπει να ενθαρρύνεται η βιωσιμότητα και η οικολογική υπευθυνότητα των συστημάτων TN και να ενισχύεται η έρευνα σε λύσεις TN που να καλύπτουν τομείς παγκόσμιου ενδιαφέροντος, όπως για παράδειγμα οι στόχοι βιώσιμης ανάπτυξης. Ιδανικά, η TN θα πρέπει να χρησιμοποιείται προς όφελος όλων των ανθρώπων, συμπεριλαμβανομένων των μελλοντικών γενεών.
- (84) **Βιώσιμη και φιλική προς το περιβάλλον TN.** Τα συστήματα TN υπόσχονται να συμβάλλουν στην αντιμετώπιση ορισμένων από τα πιο πιεστικά κοινωνικά ζητήματα, ωστόσο θα πρέπει να διασφαλίζεται ότι αυτό γίνεται με τον πλέον φιλικό τρόπο προς το περιβάλλον. Η διαδικασία ανάπτυξης, εγκατάστασης και χρήσης του συστήματος, καθώς και ολόκληρη η αλυσίδα εφοδιασμού του, θα πρέπει να αξιολογείται σε αυτό το πλαίσιο, π.χ. μέσω κριτικής εξέτασης της χρήσης πόρων και της κατανάλωσης ενέργειας κατά τη διάρκεια της εκπαίδευσης, και να επιλέγονται λιγότερο επιβλαβείς επιλογές. Θα πρέπει να ενθαρρύνονται μέτρα που εξασφαλίζουν την φιλικότητα προς το περιβάλλον ολόκληρης της αλυσίδας εφοδιασμού του συστήματος TN.
- (85) **Κοινωνικές επιπτώσεις.** Η γενικευμένη έκθεση σε κοινωνικά συστήματα TN σε όλους τους τομείς της ζωής μας (στην εκπαίδευση, την εργασία, τη φροντίδα ή την ψυχαγωγία) μπορεί να αλλοιώσει την αντίληψή μας για την κοινωνική παρέμβαση ή να επηρεάσει τις κοινωνικές σχέσεις και τις συνδέσεις μας.<sup>47</sup> Μολονότι τα συστήματα TN είναι δυνατό να χρησιμοποιηθούν για τη βελτίωση των κοινωνικών δεξιοτήτων<sup>48</sup>, είναι επίσης δυνατό να συμβάλουν στην υποβάθμισή τους. Κάτι τέτοιο θα μπορούσε να επηρεάσει επίσης τη σωματική και πνευματική ευημερία των ανθρώπων. Συνεπώς, οι επιπτώσεις των συγκεκριμένων συστημάτων θα πρέπει να παρακολουθούνται και να εξετάζονται προσεκτικά.
- (86) **Κοινωνία και Δημοκρατία.** Πέρα από την αξιολόγηση των επιπτώσεων της ανάπτυξης, εγκατάστασης και χρήσης των συστημάτων TN σε άτομα, οι συγκεκριμένες επιπτώσεις θα πρέπει επίσης να αξιολογούνται από κοινωνική άποψη, λαμβανομένης υπόψη της επίδρασής τους στα ιδρύματα, τη δημοκρατία και ευρύτερα την κοινωνία. Η χρήση συστημάτων TN θα πρέπει να εξεταστεί προσεκτικά, ιδίως σε περιπτώσεις που αφορούν τη δημοκρατική διαδικασία, όχι μόνο σχετικά με τη λήψη πολιτικών αποφάσεων, αλλά και με το εκλογικό περιβάλλον.

## 7. Λογοδοσία

- (87) Η απαίτηση της λογοδοσίας συμπληρώνει τις ανωτέρω απαιτήσεις, που συνδέονται στενά με την αρχή της δικαιοσύνης. Προϋποθέτει τη δημιουργία μηχανισμών μέσω των οποίων θα διασφαλίζεται η υπευθυνότητα και η λογοδοσία για τα συστήματα TN και τα αποτελέσματά τους, τόσο πριν όσο και μετά την υλοποίησή τους.
- (88) **Ελεγχιμότητα.** Η ελεγχιμότητα συνεπάγεται τη δυνατότητα αξιολόγησης των αλγορίθμων, των δεδομένων και των διαδικασιών σχεδιασμού. Κάτι τέτοιο δεν σημαίνει απαραίτητα ότι οι πληροφορίες σχετικά με τα επιχειρηματικά μοντέλα και την πνευματική ιδιοκτησία που σχετίζονται με το σύστημα TN θα πρέπει πάντοτε να είναι ευρέως διαθέσιμες. Η αξιολόγηση από εσωτερικούς και εξωτερικούς ελεγκτές και η διαθεσιμότητα εκθέσεων αξιολόγησης αυτού του είδους μπορούν να συνεισφέρουν στην αξιοπιστία της τεχνολογίας. Στις εφαρμογές που επηρεάζουν τα θεμελιώδη δικαιώματα, συμπεριλαμβανομένων των εφαρμογών που είναι

<sup>47</sup> Εδώ υποδηλώνονται τα συστήματα TN που επικοινωνούν και αλληλεπιδρούν με τον άνθρωπο με προσομοίωση της κοινωνικότητας στην αλληλεπίδραση με τη μορφή ανθρώπινων ρομπότ (ενσωματωμένη TN) ή ως άβαταρ σε εικονική πραγματικότητα. Με αυτόν τον τρόπο, τα συγκεκριμένα συστήματα έχουν τη δυνατότητα να μεταβάλλουν τις κοινωνικοπολιτισμικές πρακτικές μας και τον ιστό της κοινωνικής μας ζωής.

<sup>48</sup> Βλ. για παράδειγμα το χρηματοδοτούμενο από την ΕΕ πρόγραμμα που αναπτύσσει λογισμικό βασισμένο σε TN, το οποίο επιτρέπει στα ρομπότ να αλληλεπιδρούν αποτελεσματικότερα με αυτιστικά παιδιά σε θεραπευτικές συνεδρίες υπό ανθρώπινη καθοδήγηση, συμβάλλοντας στη βελτίωση των κοινωνικών και επικοινωνιακών τους δεξιοτήτων:  
[http://ec.europa.eu/research/infocentre/article\\_en.cfm?id=research/headlines/news/article\\_19\\_03\\_12\\_en.html?infocentre&item=Infocentre&artid=49968](http://ec.europa.eu/research/infocentre/article_en.cfm?id=research/headlines/news/article_19_03_12_en.html?infocentre&item=Infocentre&artid=49968)

καίριας σημασίας για την ασφάλεια, τα συστήματα TN θα πρέπει να είναι δυνατό να ελέγχονται ανεξάρτητα.

- (89) **Ελαχιστοποίηση και γνωστοποίηση των αρνητικών επιπτώσεων.** Θα πρέπει να διασφαλίζεται η ικανότητα τόσο γνωστοποίησης για δράσεις ή αποφάσεις που συμβάλλουν σε ένα συγκεκριμένο αποτέλεσμα του συστήματος, όσο και αντιμετώπισης των συνεπειών ενός τέτοιου αποτελέσματος. Ο εντοπισμός, η αξιολόγηση, η γνωστοποίηση και η ελαχιστοποίηση των δυνητικών αρνητικών επιπτώσεων των συστημάτων TN είναι ιδιαίτερως σημαντικές για εκείνους που επηρεάζονται άμεσα (ή έμμεσα). Θα πρέπει να υπάρχει δέουσα προστασία για τους καταγγέλτες, τις ΜΚΟ, τα συνδικάτα ή άλλες οντότητες, κατά τη γνωστοποίηση των θεμιτών προβληματισμών τους σχετικά με ένα σύστημα TN. Η χρήση εκτιμήσεων επιπτώσεων (π.χ. από ανεξάρτητες ομάδες εικονικών δοκιμών ή μορφές αλγοριθμικών εκτιμήσεων επιπτώσεων) τόσο πριν όσο και κατά τη διάρκεια της ανάπτυξης, εγκατάστασης και χρήσης συστημάτων TN μπορεί να αποδειχθεί χρήσιμη για την ελαχιστοποίηση των αρνητικών επιπτώσεων. Οι εκτιμήσεις αυτές πρέπει να είναι ανάλογες με τους κινδύνους που ενέχουν τα συστήματα TN.
- (90) **Αντισταθμιστικές ρυθμίσεις.** Κατά την εφαρμογή των ανωτέρω απαιτήσεων, ενδέχεται να προκύψουν τριβές μεταξύ τους, οι οποίες ενδέχεται να οδηγήσουν σε αναπόφευκτες αντισταθμιστικές ρυθμίσεις. Οι εν λόγω αντισταθμιστικές ρυθμίσεις θα πρέπει να αντιμετωπίζονται με ορθολογικό και μεθοδολογικό τρόπο στο πλαίσιο της πλέον σύγχρονης τεχνολογίας. Αυτό συνεπάγεται ότι θα πρέπει να προσδιορίζονται τα σχετικά συμφέροντα και οι αξίες που επηρεάζονται από το σύστημα TN και ότι, σε περίπτωση που προκύπτει σύγκρουση, οι αντισταθμιστικές ρυθμίσεις που γίνονται μεταξύ τους θα πρέπει να αναγνωρίζονται ρητά και να αξιολογούνται όσον αφορά τον κίνδυνο που ενέχουν για τις δεοντολογικές αρχές, συμπεριλαμβανομένων των θεμελιωδών δικαιωμάτων. Σε περιπτώσεις που δεν είναι δυνατό να προσδιοριστούν δεοντολογικά αποδεκτές αντισταθμιστικές ρυθμίσεις, η ανάπτυξη, η εγκατάσταση και η χρήση του συστήματος TN δεν θα πρέπει να προχωρά με τη συγκεκριμένη μορφή. Οποιαδήποτε απόφαση σχετικά με την αντισταθμιστική ρύθμιση που πρέπει να γίνει θα πρέπει να αιτιολογείται και να τεκμηριώνεται δεόντως. Ο υπεύθυνος για τη λήψη αποφάσεων θα πρέπει να λογοδοτεί για τον τρόπο με τον οποίο γίνεται η κατάλληλη αντισταθμιστική ρύθμιση και θα πρέπει να εξετάζει συνεχώς την καταλληλότητα της απόφασης που προκύπτει, προκειμένου να διασφαλίζει ότι μπορούν να γίνουν οι απαραίτητες αλλαγές στο σύστημα όπου αυτό είναι απαραίτητο.<sup>49</sup>
- (91) **Έννομη προστασία.** Θα πρέπει να προβλέπονται προσβάσιμοι μηχανισμοί που να εξασφαλίζουν κατάλληλη έννομη προστασία σε περιπτώσεις άδικων, δυσμενών επιπτώσεων<sup>50</sup>. Η επίγνωση ότι προβλέπεται έννομη προστασία σε περίπτωση που προκύψει κάποιο πρόβλημα είναι πολύ σημαντική για την εξασφάλιση της εμπιστοσύνης. Ιδιαίτερη προσοχή θα πρέπει να δίνεται σε ευάλωτα άτομα ή ομάδες.

## 2. Τεχνικές και μη τεχνικές μέθοδοι για την πραγμάτωση της αξιόπιστης TN

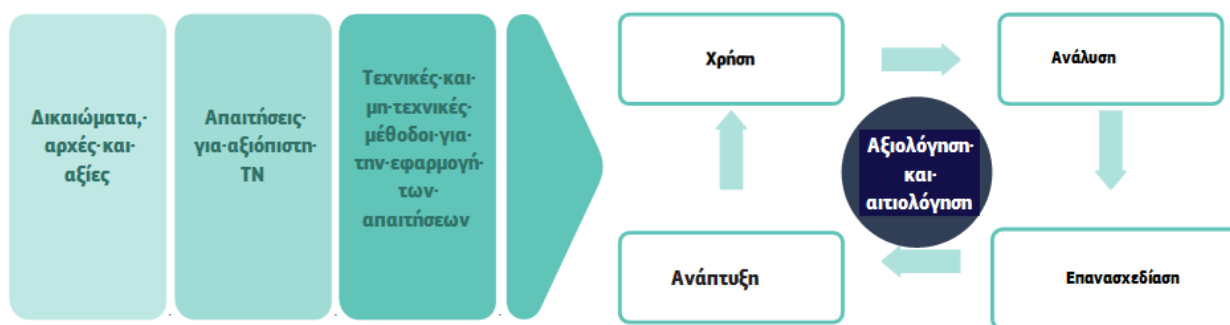
- (92) Για την εφαρμογή των ανωτέρω απαιτήσεων είναι δυνατό να χρησιμοποιούνται τεχνικές και μη τεχνικές μέθοδοι. Αυτές περιλαμβάνουν όλα τα στάδια του κύκλου ζωής ενός συστήματος TN. Η αξιολόγηση των μεθόδων που χρησιμοποιούνται για την εφαρμογή των απαιτήσεων, καθώς και η υποβολή εκθέσεων και η αιτιολόγηση των αλλαγών στις διαδικασίες εφαρμογής, θα πρέπει να πραγματοποιούνται σε διαρκή βάση.<sup>51</sup> Καθώς τα συστήματα TN εξελίσσονται συνεχώς και λειτουργούν σε δυναμικό περιβάλλον, η πραγμάτωση της

<sup>49</sup> Διάφορα μοντέλα διακυβέρνησης μπορούν να συμβάλλουν στην επίτευξη αυτού του στόχου. Π.χ. η παρουσία εσωτερικού και/ή εξωτερικού εμπειρογνώμονα ή συμβουλίου σε θέματα δεοντολογίας (και ειδικού ανά τομέα) ενδεχομένως να είναι χρήσιμη για να επισημαίνονται τομείς πιθανών συγκρούσεων και να προτείνονται τρόποι για την καλύτερη δυνατή επίλυση των εν λόγω συγκρούσεων. Χρήσιμη κρίνεται επίσης η ουσιαστική διαβούλευση και η συζήτηση με τα ενδιαφερόμενα μέρη, συμπεριλαμβανομένων εκείνων που κινδυνεύουν να επηρεαστούν αρνητικά από ένα σύστημα TN. Τα ευρωπαϊκά πανεπιστήμια θα πρέπει να αναλάβουν ηγετικό ρόλο στην κατάρτιση των απαραίτητων εμπειρογνομών σε θέματα δεοντολογίας.

<sup>50</sup> Βλ. επίσης τη γνωμοδότηση του Οργανισμού Θεμελιωδών Δικαιωμάτων της Ευρωπαϊκής Ένωσης «Improving access to remedy in the area of business and human rights at the EU level» (2017), <https://fra.europa.eu/en/opinion/2017/business-human-rights>.

<sup>51</sup> Τούτο συνεπάγεται, για παράδειγμα, την αιτιολόγηση των επιλογών που έγιναν κατά τον σχεδιασμό, την ανάπτυξη και την εγκατάσταση του συστήματος προκειμένου να ενσωματωθούν οι προαναφερθείσες απαιτήσεις.

αξιόπιστης TN αποτελεί συνεχή διαδικασία, που απεικονίζεται στο σχήμα 3 παρακάτω.



Σχήμα 3: Πραγμάτωση αξιόπιστης TN σε ολόκληρο τον κύκλο ζωής του συστήματος

(93) Οι ακόλουθες μέθοδοι είναι δυνατό να θεωρηθούν συμπληρωματικές ή εναλλακτικές μεταξύ τους, δεδομένου ότι διαφορετικές απαιτήσεις —και διαφορετικές ευαισθησίες— ενδεχομένως να ενισχύσουν την ανάγκη για διαφορετικές μεθόδους υλοποίησης. Η συγκεκριμένη επισκόπηση δεν είναι ούτε ολοκληρωμένη, ούτε εξαντλητική, ούτε υποχρεωτική. Αντιθέτως, στόχος της είναι να προσφέρει έναν κατάλογο προτεινόμενων μεθόδων που ενδεχομένως να συμβάλλουν στην υλοποίηση της αξιόπιστης TN.

#### 1. Τεχνικές μέθοδοι

(94) Η παρούσα ενότητα περιγράφει τεχνικές μεθόδους για την εξασφάλιση αξιόπιστης TN που μπορούν να ενσωματωθούν στις φάσεις σχεδιασμού, ανάπτυξης και χρήσης ενός συστήματος TN. Οι παρακάτω μέθοδοι ποικίλλουν σε επίπεδο ωριμότητας<sup>52</sup>.

##### ▪ Αρχιτεκτονικές για αξιόπιστη TN

(95) Οι απαιτήσεις για αξιόπιστη TN θα πρέπει να «μεταφραστούν» σε διαδικασίες και/ή περιορισμούς στις διαδικασίες, που θα πρέπει να εδράζονται στην αρχιτεκτονική του συστήματος TN. Κάτι τέτοιο θα ήταν δυνατό να επιτευχθεί μέσω ενός συνόλου κανόνων «καταλόγου επιτρεπόμενων» (για συμπεριφορές ή καταστάσεις) που θα πρέπει να εφαρμόζει πάντα το σύστημα, περιορισμούς «καταλόγου αποκλεισμού» στις συμπεριφορές ή καταστάσεις που δεν θα πρέπει ποτέ να παραβιάζει το σύστημα, καθώς και συνδυασμούς των εν λόγω καταλόγων ή πιο πολύπλοκων τεκμηριωμένων εγγυήσεων σχετικά με τη συμπεριφορά του συστήματος. Η παρακολούθηση της συμμόρφωσης του συστήματος με τους εν λόγω περιορισμούς κατά τη διάρκεια των εργασιών είναι δυνατό να επιτυγχάνεται με ξεχωριστή διαδικασία.

(96) Τα συστήματα TN με δυνατότητες μάθησης που έχουν τη δυνατότητα να προσαρμόζουν δυναμικά τη συμπεριφορά τους είναι δυνατό να εκληφθούν ως ένα μη προσδιοριστικό σύστημα που ενδεχομένως παρουσιάζει απροσδόκητη συμπεριφορά. Αυτά συχνά εξετάζονται μέσω του θεωρητικού πρίσματος ενός κύκλου «αίσθησης - σχεδιασμού δράσης - δράσης» (sense-plan-act). Η προσαρμογή της συγκεκριμένης αρχιτεκτονικής για την εξασφάλιση της αξιόπιστης TN απαιτεί την ενσωμάτωση των απαιτήσεων και στα τρία στάδια του κύκλου: i) στο βήμα «αίσθησης» (sense), το σύστημα θα πρέπει να αναπτύσσεται έτσι ώστε να αναγνωρίζει όλα τα στοιχεία του περιβάλλοντος που είναι απαραίτητα για να εξασφαλίζεται η συμμόρφωση με τις απαιτήσεις· ii) στο βήμα «σχεδιασμός δράσης» (plan), το σύστημα θα πρέπει να εξετάζει μόνο τα σχέδια που συμμορφώνονται με τις απαιτήσεις· iii) στο βήμα «δράση» (act), οι ενέργειες του συστήματος θα πρέπει να περιορίζονται σε συμπεριφορές που τηρούν τις απαιτήσεις.

<sup>52</sup> Ενώ ορισμένες από τις συγκεκριμένες μεθόδους είναι ήδη διαθέσιμες επί του παρόντος, για άλλες απαιτείται ακόμα περισσότερη έρευνα. Οι συγκεκριμένοι τομείς στους οποίους απαιτείται περαιτέρω έρευνα θα εξεταστούν επίσης στο δεύτερο παραδοτέο της ΟΕΥΕ για την TN, ήτοι τις Συστάσεις Πολιτικής και Επενδύσεων.

(97) Η αρχιτεκτονική όπως σκιαγραφείται ανωτέρω είναι γενικής μορφής και παρέχει μόνο μια ατελή περιγραφή για τα περισσότερα συστήματα TN. Ωστόσο, προσφέρει σημεία έδρασης για τους περιορισμούς και τις πολιτικές που πρέπει να αντικατοπτρίζονται σε συγκεκριμένες λειτουργικές μονάδες, ώστε να προκύψει ένα συνολικό σύστημα που να είναι και να θεωρείται αξιόπιστο.

▪ *Δεοντολογία και κράτος δικαίου ήδη από τον σχεδιασμό (X ήδη από τον σχεδιασμό)*

(98) Οι μέθοδοι για τη διασφάλιση των αξιών ήδη από τον σχεδιασμό παρέχουν ακριβείς και σαφείς συνδέσεις μεταξύ των αφηρημένων αρχών με τις οποίες θα πρέπει να συμμορφώνεται το σύστημα και των συγκεκριμένων αποφάσεων υλοποίησης. Η ιδέα ότι η συμμόρφωση με τους κανόνες μπορεί να εφαρμοστεί στον σχεδιασμό του συστήματος TN αποτελεί κεντρικό στοιχείο για τη συγκεκριμένη μέθοδο. Οι εταιρείες είναι υπεύθυνες για τον προσδιορισμό των επιπτώσεων των συστημάτων TN που διαθέτουν ήδη από την αρχή, καθώς και για τους κανόνες που θα πρέπει να τηρεί το οικείο σύστημα TN για την αποφυγή των αρνητικών επιπτώσεων. Διαφορετικές έννοιες του «ήδη από τον σχεδιασμό» χρησιμοποιούνται ήδη ευρέως, π.χ. *προστασία της ιδιωτικής ζωής ήδη από τον σχεδιασμό και προστασία ήδη από τον σχεδιασμό*. Όπως αναφέρθηκε ανωτέρω, προκειμένου να εμπνεεί εμπιστοσύνη η TN, οι διαδικασίες, τα δεδομένα και τα αποτελέσματά της θα πρέπει να είναι ασφαλή, ενώ θα πρέπει να σχεδιάζεται έτσι ώστε να είναι στιβαρή απέναντι σε δεδομένα και επιθέσεις από αντίπαλες πλευρές. Θα πρέπει να εφαρμόζει μηχανισμό για ασφαλή τερματισμό λειτουργίας και να παρέχει τη δυνατότητα συνέχισης της λειτουργίας μετά από αναγκαστική διακοπή (π.χ. επίθεση).

▪ *Μέθοδοι επεξήγησης*

(99) Προκειμένου ένα σύστημα να είναι αξιόπιστο, θα πρέπει να είμαστε σε θέση να κατανοήσουμε τους λόγους για τους οποίους συμπεριφέρθηκε με συγκεκριμένο τρόπο και τους λόγους για τους οποίους παρείχε μια δεδομένη ερμηνεία. Σε ένα ολόκληρο ερευνητικό πεδίο, την επεξηγήσιμη TN (Explainable AI – XAI), καταβάλλεται προσπάθεια να καλυφθεί το συγκεκριμένο ζήτημα για την καλύτερη κατανόηση των υποκείμενων μηχανισμών του συστήματος, καθώς και την εξεύρεση λύσεων. Επί του παρόντος, αυτό εξακολουθεί να αποτελεί μια ανοικτή πρόκληση για συστήματα TN που βασίζονται σε νευρωνικά δίκτυα. Οι διαδικασίες εκπαίδευσης με νευρωνικά δίκτυα μπορούν να οδηγήσουν σε παραμέτρους του δικτύου που ορίζονται σε αριθμητικές τιμές οι οποίες είναι δύσκολο να συσχετιστούν με τα αποτελέσματα. Επιπλέον, μερικές φορές μικρές αλλαγές στις τιμές των δεδομένων μπορεί να οδηγήσουν σε σημαντικές αλλαγές στην ερμηνεία, προκαλώντας σύγχυση στο σύστημα π.χ. μεταξύ ενός σχολικού λεωφορείου με μια στρουθοκάμηλο. Η εν λόγω ευπάθεια μπορεί επίσης να αξιοποιηθεί κατά τις επιθέσεις στο σύστημα. Οι μέθοδοι που περιλαμβάνουν την έρευνα XAI είναι ζωτικής σημασίας όχι μόνο προκειμένου να εξηγείται η συμπεριφορά του συστήματος στους χρήστες, αλλά και για να εγκαθίσταται αξιόπιστη τεχνολογία.

▪ *Δοκιμή και επικύρωση*

(100) Λόγω της μη προσδιοριστικής φύσης των συστημάτων TN, η οποία εξαρτάται από το συγκεκριμένο πλαίσιο, οι παραδοσιακές δοκιμές δεν αρκούν. Οι αποτυχίες των εννοιών και των αναπαραστάσεων που χρησιμοποιούνται από το σύστημα μπορούν να εκδηλωθούν μόνο όταν ένα πρόγραμμα εφαρμόζεται σε επαρκώς ρεαλιστικά δεδομένα. Συνεπώς, προκειμένου να επαληθευτεί και να επικυρωθεί η επεξεργασία των δεδομένων, το υποκείμενο μοντέλο θα πρέπει να παρακολουθείται προσεκτικά κατά τη διάρκεια τόσο της εκπαίδευσης όσο και της εγκατάστασης για τη σταθερότητα, τη στιβαρότητα και τη λειτουργία του εντός σαφών και προβλέψιμων ορίων. Θα πρέπει να διασφαλίζεται ότι το αποτέλεσμα της διαδικασίας σχεδιασμού είναι συνεπές με τις εισροές και ότι οι αποφάσεις λαμβάνονται κατά τρόπο που επιτρέπει την επικύρωση της υποκείμενης διαδικασίας.

(101) Η δοκιμή και η επικύρωση του συστήματος θα πρέπει να πραγματοποιούνται όσο το δυνατό νωρίτερα και να εξασφαλίζεται ότι το σύστημα συμπεριφέρεται με τον αναμενόμενο τρόπο καθόλη τη διάρκεια του κύκλου ζωής του και ιδίως μετά την εγκατάσταση. Θα πρέπει να περιλαμβάνει όλα τα στοιχεία ενός συστήματος TN, συμπεριλαμβανομένων των δεδομένων, των προ-εκπαιδευμένων μοντέλων, των περιβαλλόντων και της

συμπεριφοράς του συστήματος στο σύνολό του. Θα πρέπει να σχεδιάζεται και να εκτελείται από μια όσο το δυνατόν πιο ποικιλόμορφη ομάδα ανθρώπων. Θα πρέπει να αναπτύσσονται πολλαπλές μέθοδοι μετρήσεων για να καλύπτουν τις κατηγορίες που δοκιμάζονται για διαφορετικές προοπτικές. Είναι δυνατό να εξετάζεται η χρήση ελέγχων κατ' αντιπαράθεση από αξιόπιστες και πολύμορφες ανεξάρτητες ομάδες εικονικών δοκιμών (red teams) που προσπαθούν σκοπίμως να «σπάσουν» το σύστημα προκειμένου να εντοπίσουν ευπάθειες, καθώς και η χρήση προγραμμάτων επιβράβευσης για εντοπισμό σφαλμάτων (bug bounties) που προσφέρουν κίνητρο σε τρίτους να ανιχνεύουν και να αναφέρουν υπεύθυνα τα σφάλματα και τις αδυναμίες του συστήματος. Τέλος, θα πρέπει να διασφαλίζεται ότι τα αποτελέσματα ή οι ενέργειες είναι συνεπείς με τα αποτελέσματα των προηγούμενων διαδικασιών, σε σύγκριση με τις πολιτικές που έχουν καθοριστεί στο παρελθόν, προκειμένου να διασφαλίζεται ότι δεν παραβιάζονται.

- *Δείκτες ποιότητας υπηρεσίας*

(102) Είναι δυνατό να ορίζονται κατάλληλοι δείκτες ποιότητας υπηρεσιών για τα συστήματα TN προκειμένου να διασφαλίζεται μια βασική κατανόηση σε σχέση με το κατά πόσο έχουν δοκιμαστεί και αναπτυχθεί λαμβανομένων υπόψη των θεμάτων ασφάλειας και προστασίας. Οι εν λόγω δείκτες θα μπορούσαν να περιλαμβάνουν μέτρα για την αξιολόγηση της δοκιμής και της εκπαίδευσης των αλγορίθμων, καθώς και παραδοσιακές μεθόδους μέτρησης της λειτουργικότητας του λογισμικού, επιδόσεις, χρηστικότητα, αξιοπιστία, προστασία και συντηρησιμότητα.

## 2. Μη τεχνικές μέθοδοι

(103) Στην παρούσα ενότητα περιγράφονται ποικίλες μη τεχνικές μέθοδοι που μπορούν να διαδραματίσουν πολύτιμο ρόλο στη διασφάλιση και διατήρηση αξιόπιστης TN. Και αυτές επίσης θα πρέπει να αξιολογούνται σε **διαρκή βάση**.

- *Κανονιστική ρύθμιση*

(104) Όπως προαναφέρθηκε, υπάρχει ήδη κανονιστικό πλαίσιο για την υποστήριξη της αξιοπιστίας της TN επί του παρόντος —μεταξύ άλλων νομοθεσία για την ασφάλεια των προϊόντων και πλαίσια λογοδοσίας. Στον βαθμό που θεωρούμε ότι το κανονιστικό πλαίσιο ενδέχεται να χρειάζεται αναθεώρηση, προσαρμογή ή θέσπιση, τόσο ως διασφάλιση όσο και ως παράγοντας διευκόλυνσης, το ζήτημα αυτό θα εξεταστεί και στο δεύτερο παραδοτέο μας, το οποίο θα αποτελείται από τις Συστάσεις Πολιτικής και Επενδύσεων TN.

- *Κώδικες δεοντολογίας*

(105) Οι οργανισμοί και τα ενδιαφερόμενα μέρη μπορούν να προσυπογράψουν τις κατευθυντήριες γραμμές και να προσαρμόσουν τον χάρτη εταιρικής ευθύνης, τους κύριους δείκτες επιδόσεων, τους κώδικες δεοντολογίας τους ή τα έγγραφα εσωτερικής πολιτικής, ώστε να συμβάλουν στις προσπάθειες για αξιόπιστη TN. Ένας οργανισμός που εργάζεται σε ένα σύστημα TN μπορεί, γενικότερα, να τεκμηριώνει τις προθέσεις του, καθώς και να τις προστατεύει με πρότυπα ορισμένων επιθυμητών αξιών, όπως τα θεμελιώδη δικαιώματα, η διαφάνεια και η αποφυγή βλάβης.

- *Τυποποίηση*

(106) Τα πρότυπα, π.χ. για τον σχεδιασμό, την κατασκευή και τις επιχειρηματικές πρακτικές, μπορεί να λειτουργούν ως σύστημα διαχείρισης ποιότητας για τους χρήστες της TN, τους καταναλωτές, τους οργανισμούς, τα ερευνητικά ιδρύματα και τις κυβερνήσεις, προσφέροντας την ικανότητα να αναγνωρίζουν και να ενθαρρύνουν τη δεοντολογική συμπεριφορά μέσω των αποφάσεων αγοράς τους. Πέρα από τα συμβατικά πρότυπα, υφίστανται προσεγγίσεις συρρύθμισης: συστήματα διαπίστευσης, επαγγελματικοί κώδικες δεοντολογίας ή πρότυπα για τον σχεδιασμό που συμμορφώνεται με τα θεμελιώδη δικαιώματα. Τρέχοντα παραδείγματα είναι π.χ. τα πρότυπα ISO ή η σειρά προτύπων IEEE P7000, αλλά στο μέλλον ενδέχεται να κριθεί κατάλληλη ή χρήση κάποιας σήμανσης όπως, ενδεχομένως, «Αξιόπιστη TN», με την οποία θα επιβεβαιώνεται διά της αναφοράς σε συγκεκριμένα τεχνικά πρότυπα ότι, για παράδειγμα, το σύστημα



συμμορφώνεται με τις αρχές της ασφάλειας, της τεχνικής στιβαρότητας και της επεξηγησιμότητας.

- *Πιστοποίηση*

(107) Καθώς δεν είναι δυνατό να αναμένεται ότι όλοι είναι σε θέση να κατανοήσουν πλήρως τις λειτουργίες και τα αποτελέσματα των συστημάτων TN, θα πρέπει να εξετάζεται το ενδεχόμενο αξιοποίησης οργανισμών που είναι σε θέση να επιβεβαιώσουν στο ευρύτερο κοινό ότι ένα σύστημα TN χαρακτηρίζεται από διαφάνεια, υπευθυνότητα και δικαιοσύνη<sup>53</sup>. Οι εν λόγω πιστοποιήσεις θα εφαρμόζουν πρότυπα που έχουν αναπτυχθεί για διαφορετικούς τομείς εφαρμογών και τεχνικές TN, κατάλληλα ευθυγραμμισμένα με τα βιομηχανικά και κοινωνικά πρότυπα του εκάστοτε πλαισίου. Ωστόσο, η πιστοποίηση δεν μπορεί ποτέ να αντικαταστήσει την ευθύνη. Επομένως, θα πρέπει να συμπληρωθεί από πλαίσια λογοδοσίας, συμπεριλαμβανομένων αποποιήσεων ευθύνης, καθώς και μηχανισμών ελέγχου και αποκατάστασης<sup>54</sup>.

- *Λογοδοσία μέσω πλαισίων διακυβέρνησης*

(108) Οι οργανισμοί θα πρέπει να δημιουργούν πλαίσια διακυβέρνησης, τόσο εσωτερικά όσο και εξωτερικά, ώστε να εξασφαλίζεται η λογοδοσία για τις δεοντολογικές διαστάσεις των αποφάσεων που σχετίζονται με την ανάπτυξη, την εγκατάσταση και τη χρήση της TN. Κάτι τέτοιο μπορεί, για παράδειγμα, να περιλαμβάνει τον διορισμό ενός ατόμου υπεύθυνου για ζητήματα δεοντολογίας που σχετίζονται με την TN ή μια εσωτερική/εξωτερική επιτροπή ή συμβούλιο δεοντολογίας. Μεταξύ των πιθανών ρόλων ενός τέτοιου ατόμου, μιας επιτροπής ή ενός συμβουλίου, είναι η παροχή εποπτείας και συμβουλευτικής. Όπως προαναφέρθηκε, οι προδιαγραφές και/ή οι οργανισμοί πιστοποίησης μπορούν επίσης να διαδραματίσουν κάποιον ρόλο προς τον σκοπό αυτό. Θα πρέπει να διασφαλίζονται διάυλοι επικοινωνίας με ομάδες του κλάδου και/ή δημόσιας εποπτείας, για ανταλλαγή βέλτιστων πρακτικών, συζητήσεις για τα διλήμματα ή γνωστοποίηση των αναδυόμενων ζητημάτων που εγείρουν δεοντολογικούς προβληματισμούς. Οι εν λόγω μηχανισμοί μπορούν να συμπληρώσουν, αλλά δεν μπορούν να αντικαταστήσουν τη νομική εποπτεία (π.χ. με τη μορφή διορισμού υπεύθυνου προστασίας δεδομένων ή ισοδύναμων μέτρων, που απαιτούνται διά νόμου από τη νομοθεσία περί προστασίας δεδομένων).

- *Εκπαίδευση και ευαισθητοποίηση για την προώθηση πνεύματος δεοντολογίας*

(109) Η αξιόπιστη TN ενθαρρύνει την ενημερωμένη συμμετοχή όλων των ενδιαφερόμενων μερών. Η επικοινωνία, η εκπαίδευση και η κατάρτιση διαδραματίζουν καίριο ρόλο, τόσο προκειμένου να εξασφαλίζεται η ευρεία διάδοση της γνώσης των δυνητικών επιπτώσεων των συστημάτων TN όσο και προκειμένου οι πολίτες να γνωρίζουν ότι μπορούν να συμμετάσχουν στη διαμόρφωση της κοινωνιακής ανάπτυξης. Τούτο περιλαμβάνει όλα τα ενδιαφερόμενα μέρη, π.χ. όσους συμμετέχουν στην κατασκευή των προϊόντων (σχεδιαστές και προγραμματιστές), τους χρήστες (εταιρείες ή ιδιώτες) και άλλες επηρεαζόμενες ομάδες (αυτούς που δεν μπορούν να αγοράσουν ή να χρησιμοποιήσουν ένα σύστημα TN, αλλά για τους οποίους λαμβάνονται αποφάσεις από ένα σύστημα TN, καθώς και την κοινωνία γενικότερα). Η βασική παιδεία σε θέματα TN θα πρέπει να προωθείται σε ολόκληρη την κοινωνία. Προϋπόθεση για την εκπαίδευση του κοινού είναι η εξασφάλιση των κατάλληλων δεξιοτήτων και της κατάρτισης των ειδικών σε θέματα δεοντολογίας στον συγκεκριμένο χώρο.

- *Συμμετοχή των ενδιαφερόμενων μερών και κοινωνικός διάλογος*

(110) Τα οφέλη της TN είναι πολλαπλά και η Ευρώπη θα πρέπει να εξασφαλίσει ότι είναι διαθέσιμα σε όλους. Για να το πετύχει απαιτείται ανοικτή συζήτηση και η συμμετοχή των κοινωνικών εταίρων, των ενδιαφερόμενων μερών, καθώς και του ευρύτερου κοινού. Πολλοί οργανισμοί βασίζονται ήδη σε επιτροπές ενδιαφερόμενων

<sup>53</sup> Όπως υποστηρίζεται π.χ. από την πρωτοβουλία IEEE Ethically Aligned Design: <https://standards.ieee.org/industry-connections/ec/autonomous-systems.html>.

<sup>54</sup> Για περισσότερα σχετικά με τους περιορισμούς της πιστοποίησης, βλ.: [https://ainowinstitute.org/AI\\_Now\\_2018\\_Report.pdf](https://ainowinstitute.org/AI_Now_2018_Report.pdf).

μερών για τον διάλογο σχετικά με τη χρήση συστημάτων TN και αναλύσεων δεδομένων. Στις συγκεκριμένες επιτροπές περιλαμβάνονται διάφορα μέλη, όπως νομικοί εμπειρογνώμονες, τεχνικοί εμπειρογνώμονες, ειδικοί σε θέματα δεοντολογίας, εκπρόσωποι καταναλωτών και εργαζόμενοι. Με την ενεργή επιδίωξη συμμετοχής, η οποία μπορεί να είναι ιδιαίτερως χρήσιμη σε πολύπλοκες υποθέσεις, και διαλόγου σχετικά με τη χρήση και τις επιπτώσεις των συστημάτων TN στηρίζεται η αξιολόγηση των αποτελεσμάτων και των προσεγγίσεων..

▪ *Πολυμορφία και συμπεριληπτικές ομάδες σχεδιασμού*

(111) Η πολυμορφία και η συμπεριληπτικότητα διαδραματίζουν ουσιαστικό ρόλο κατά την ανάπτυξη συστημάτων TN που θα χρησιμοποιηθούν στον πραγματικό κόσμο. Καθώς τα συστήματα TN εκτελούν περισσότερες εργασίες μόνα τους, είναι ιδιαίτερως σημαντικό οι ομάδες που σχεδιάζουν, αναπτύσσουν, δοκιμάζουν και συντηρούν, εγκαθιστούν και/ή προμηθεύονται αυτά τα συστήματα να αντικατοπτρίζουν την πολυμορφία των χρηστών και της κοινωνίας γενικότερα, διότι αυτό συμβάλλει στην αντικειμενικότητα και την εξέταση διαφορετικών προοπτικών, αναγκών και στόχων. Ιδανικά, οι ομάδες παρουσιάζουν πολυμορφία όχι μόνο από άποψη φύλου, πολιτισμού ή ηλικίας, αλλά και από άποψη επαγγελματικού υποβάθρου και δεξιοτήτων.

**Βασικές κατευθύνσεις που προκύπτουν από το κεφάλαιο II:**

- ✓ Διασφάλιση της συμμόρφωσης του κύκλου ζωής του συστήματος TN με τις απαιτήσεις για την αξιόπιστη TN: 1) ανθρώπινη παρέμβαση και εποπτεία, 2) τεχνική στιβαρότητα και ασφάλεια, 3) ιδιωτική ζωή και διακυβέρνηση των δεδομένων, 4) διαφάνεια, 5) πολυμορφία, απαγόρευση των διακρίσεων και δικαιοσύνη, 6) περιβαλλοντική και κοινωνιακή ευημερία και 7) λογοδοσία.
- ✓ Εξέταση τεχνικών και μη τεχνικών μεθόδων για την εξασφάλιση της εφαρμογής των εν λόγω απαιτήσεων.
- ✓ Προώθηση της έρευνας και της καινοτομίας ώστε να συμβάλλουν στην αξιολόγηση των συστημάτων TN και να προάγουν την επίτευξη των απαιτήσεων· διάδοση των αποτελεσμάτων και άνοιγμα των ερωτημάτων στο ευρύτερο κοινό, καθώς και συστηματική εκπαίδευση μιας νέας γενιάς εμπειρογνομόνων στη δεοντολογία της TN.
- ✓ Ενημέρωση, με σαφή και ενεργό τρόπο, των ενδιαφερόμενων μερών σχετικά με τις δυνατότητες και τους περιορισμούς του συστήματος TN, ώστε να δημιουργούνται ρεαλιστικές προσδοκίες, καθώς και σχετικά με τον τρόπο με τον οποίο εφαρμόζονται οι απαιτήσεις. Διαφάνεια έναντι των εν λόγω ενδιαφερόμενων μερών ως προς το γεγονός ότι έχουν να κάνουν με ένα σύστημα TN.
- ✓ Διευκόλυνση της ιχνηλασιμότητας και της ελεγχιμότητας των συστημάτων TN, ιδίως σε κρίσιμα πλαίσια και καταστάσεις.
- ✓ Συμμετοχή των ενδιαφερόμενων μερών καθόλη τη διάρκεια του κύκλου ζωής του συστήματος TN. Ενίσχυση της κατάρτισης και της εκπαίδευσης, έτσι ώστε όλα τα ενδιαφερόμενα μέρη να είναι κατατοπισμένα και εκπαιδευμένα σε σχέση με την αξιόπιστη TN.
- ✓ Επίγνωση του γεγονότος ότι μπορεί να υπάρχουν θεμελιώδεις τριβές μεταξύ διαφόρων αρχών και απαιτήσεων. Διαρκής προσδιορισμός, αξιολόγηση, τεκμηρίωση και γνωστοποίηση των εν λόγω τριβών και των λύσεών τους με τη μορφή αντισταθμιστικών ρυθμίσεων.

**III. Κεφάλαιο III: Αξιολόγηση της αξιοπιστίας της τεχνητής νοημοσύνης**

(112) Με βάση τις βασικές απαιτήσεις του κεφαλαίου II, στο παρόν κεφάλαιο παρατίθεται μη εξαντλητικός **κατάλογος αξιολόγησης της αξιοπιστίας της TN** (πιλοτική έκδοση) **για την επιχειρησιακή υλοποίηση της αξιόπιστης TN**. Ο κατάλογος αξιολόγησης αφορά ιδίως τα συστήματα TN που αλληλεπιδρούν απευθείας

με χρήστες και απευθύνεται κυρίως σε προγραμματιστές και εγκαταστάτες συστημάτων TN (είτε τα αναπτύσσουν οι ίδιοι είτε τα προμηθεύονται από τρίτους). Αυτός ο κατάλογος αξιολόγησης δεν καλύπτει την επιχειρησιακή υλοποίηση της πρώτης συνιστώσας της αξιόπιστης TN (σύννομη TN). Η συμμόρφωση με αυτόν τον κατάλογο αξιολόγησης δεν αποτελεί τεκμήριο νομικής συμμόρφωσης ούτε προορίζεται για την παροχή κατευθύνσεων για τη διασφάλιση της συμμόρφωσης με την εφαρμοστέα νομοθεσία. Δεδομένων των ιδιοτήτων των συστημάτων TN αναλόγως της εκάστοτε εφαρμογής, ο κατάλογος αξιολόγησης θα πρέπει να προσαρμόζεται στις συγκεκριμένες περιπτώσεις χρήσης και τα πλαίσια εντός των οποίων λειτουργούν τα συστήματα. Επιπλέον, στο παρόν κεφάλαιο παρέχεται μια γενική σύσταση που αφορά το πώς μπορεί να εφαρμόζεται ο κατάλογος αξιολόγησης της αξιοπιστίας της TN μέσω μιας δομής διακυβέρνησης που συνδυάζει το επιχειρησιακό με το διοικητικό επίπεδο.

(113) Ο κατάλογος αξιολόγησης και η δομή διακυβέρνησης θα διαμορφωθούν σε στενή συνεργασία με ενδιαφερόμενα μέρη από ολόκληρο τον δημόσιο και τον ιδιωτικό τομέα. Η διαδικασία θα έχει πιλοτικό χαρακτήρα και θα καθιστά δυνατή τη συλλογή εκτενούς αναπληροφόρησης μέσω δύο παράλληλων διεργασιών:

- a. μιας ποιοτικής διεργασίας με εχέγγυα αντιπροσωπευτικότητας, στην οποία θα μετάσχει μικρός αριθμός επιλεγμένων εταιρειών, οργανισμών και ιδρυμάτων (από διάφορους τομείς και διαφόρων μεγεθών) προκειμένου να εφαρμόσουν πιλοτικά τον κατάλογο αξιολόγησης και τη δομή διακυβέρνησης στην πράξη και να παράσχουν εμπειριστατωμένη αναπληροφόρηση.
- b. μια ποσοτική διεργασία, στην οποία μπορούν να μετάσχουν όλα τα ενδιαφερόμενα μέρη, προκειμένου να εφαρμόσουν πιλοτικά τον κατάλογο αξιολόγησης και να παράσχουν αναπληροφόρηση στο πλαίσιο ανοιχτής διαβούλευσης.

(114) Μετά την πιλοτική φάση θα ενσωματώσουμε τα αποτελέσματα της διαδικασίας αναπληροφόρησης στον κατάλογο αξιολόγησης και θα εκπονήσουμε αναθεωρημένη έκδοσή του στις αρχές του 2020. Στόχος είναι η επίτευξη ενός πλαισίου το οποίο θα μπορεί να χρησιμοποιείται οριζόντια σε όλες τις εφαρμογές και, ως εκ τούτου, θα παρέχει τη βάση για τη διασφάλιση αξιόπιστης TN σε όλα τα πεδία. Αφού παγιωθεί η βάση αυτή, θα μπορεί να αναπτυχθεί ένα πλαίσιο ειδικά για τον εκάστοτε τομέα ή εφαρμογή.

#### Διακυβέρνηση

(115) Σε περίπτωση που εταιρείες, οργανισμοί και ιδρύματα επιθυμούν να εξετάσουν τρόπους με τους οποίους ο κατάλογος αξιολόγησης της αξιοπιστίας της TN μπορεί να εφαρμοστεί στην οργανωτική τους δομή, αυτό μπορεί να γίνει είτε με τη συμπεριληψη της διαδικασίας αξιολόγησης σε υφιστάμενους μηχανισμούς διακυβέρνησης είτε μέσω της εφαρμογής νέων διαδικασιών. Το τι θα επιλέξουν εξαρτάται από την εσωτερική οργανωτική τους δομή, καθώς και από το μέγεθος και τους διαθέσιμους πόρους τους.

(116) Από έρευνες<sup>55</sup> προκύπτει ότι απαραίτητη προϋπόθεση για την επίτευξη αλλαγής είναι η μέριμνα της διοίκησης στο ανώτατο επίπεδο. Προκύπτει επίσης ότι η εξασφάλιση της συμμετοχής όλων των ενδιαφερόμενων μερών μιας εταιρείας, ενός οργανισμού ή ενός ιδρύματος συμβάλλει στην αποδοχή και τη σημασία της καθιέρωσης οποιασδήποτε νέας διαδικασίας (τεχνολογικής ή μη)<sup>56</sup>. Ως εκ τούτου, συνιστούμε την εφαρμογή μιας διαδικασίας που θα περιλαμβάνει τη συμμετοχή τόσο σε επιχειρησιακό επίπεδο όσο και σε επίπεδο ανώτατης διοίκησης.

#### Επίπεδο

#### Σχετικές αρμοδιότητες (αναλόγως οργανισμού)

<sup>55</sup> <https://www.mckinsey.com/business-functions/operations/our-insights/secrets-of-successful-change-implementation>

<sup>56</sup> Βλ. για παράδειγμα A. Bryson, E. Barth και H. Dale-Olsen, *The Effects of Organisational change on worker well-being and the moderating role of trade unions*, *ILRRReview*, 66(4), Ιούλιος 2013· Jirjahn, U. και Smith, S.C. (2006). «What Factors Lead Management to Support or Oppose Employee Participation—With and Without Works Councils? Hypotheses and Evidence from Germany's Industrial Relations», 45(4), 650–680· Michie, J. και Sheehan, M. (2003). «Labour market deregulation, “flexibility” and innovation», *Cambridge Journal of Economics*, 27(1), 123–143.

Διοίκηση και διοικητικό συμβούλιο	Η ανώτατη διοίκηση εξετάζει και αξιολογεί την ανάπτυξη, εγκατάσταση ή προμήθεια ΤΝ και λειτουργεί ως όργανο παραπομπής που αξιολογεί όλες τις καινοτομίες και τις χρήσεις ΤΝ όταν προκύπτουν καίριοι προβληματισμοί. Συνεργάζεται με όσους επηρεάζονται από την ενδεχόμενη εισαγωγή συστημάτων ΤΝ (π.χ. εργαζόμενοι) και τους εκπροσώπους τους σε όλα τα στάδια της διεργασίας, μέσω διαδικασιών ενημέρωσης, διαβούλευσης και συμμετοχής.
Τμήμα κανονιστικής συμμόρφωσης/ νομικό τμήμα/ τμήμα εταιρικής ευθύνης	Το αρμόδιο τμήμα παρακολουθεί τη χρήση του καταλόγου αξιολόγησης και την αναγκαία εξέλιξή του με βάση τις τεχνολογικές ή κανονιστικές αλλαγές. Επικαιροποιεί τα πρότυπα ή τις εσωτερικές πολιτικές που αφορούν τα συστήματα ΤΝ και διασφαλίζει ότι η χρήση των συστημάτων αυτών συμμορφώνεται με το υφιστάμενο νομοθετικό και κανονιστικό πλαίσιο, καθώς και με τις αξίες του οργανισμού.
Τμήμα ανάπτυξης προϊόντων και υπηρεσιών ή αντίστοιχο τμήμα	Το τμήμα ανάπτυξης προϊόντων και υπηρεσιών χρησιμοποιεί τον κατάλογο αξιολόγησης για την αξιολόγηση προϊόντων και υπηρεσιών που βασίζονται στην ΤΝ και καταγράφει όλα τα αποτελέσματα. Τα αποτελέσματα εξετάζονται σε επίπεδο διοίκησης, η οποία εγκρίνει εν τέλει τις νέες ή αναθεωρημένες εφαρμογές που βασίζονται στην ΤΝ.
Τμήμα διασφάλισης ποιότητας	Το τμήμα διασφάλισης ποιότητας (ή αντίστοιχο τμήμα) διασφαλίζει και ελέγχει τα αποτελέσματα του καταλόγου αξιολόγησης και παραπέμπει ένα ζήτημα σε υψηλότερα κλιμάκια αν το αποτέλεσμα δεν είναι ικανοποιητικό ή αν διαπιστώνονται απρόβλεπτα αποτελέσματα.
Τμήμα ανθρώπινου δυναμικού	Το τμήμα ανθρώπινου δυναμικού μεριμνά για τον κατάλληλο συνδυασμό ικανοτήτων και την πολυμορφία των χαρακτηριστικών των προγραμματιστών συστημάτων ΤΝ. Μεριμνά για την παροχή κατάλληλου επιπέδου εκπαίδευσης σχετικά με την αξιόπιστη ΤΝ στο εσωτερικό του οργανισμού.
Τμήμα προμηθειών	Το τμήμα προμηθειών μεριμνά ώστε η διαδικασία για την προμήθεια προϊόντων ή υπηρεσιών που βασίζονται στην ΕΕ να περιλαμβάνει έλεγχο της αξιοπιστίας της ΤΝ.
Καθημερινή λειτουργία	Οι προγραμματιστές και οι υπεύθυνοι έργου συμπεριλαμβάνουν τον κατάλογο αξιολόγησης στην καθημερινή εργασία τους και καταγράφουν τα αποτελέσματα της αξιολόγησης.

#### *Χρήση του καταλόγου αξιολόγησης της αξιοπιστίας της ΤΝ*

- (117) Όταν ο κατάλογος αξιολόγησης χρησιμοποιείται στην πράξη συνιστούμε να δίνεται βαρύτητα όχι μόνο στα θέματα που εγείρουν προβληματισμούς αλλά και στις ερωτήσεις που δεν μπορούν να απαντηθούν (εύκολα). Ένα πιθανό πρόβλημα μπορεί να είναι η έλλειψη πολυμορφίας στις δεξιότητες και της ικανότητες της ομάδας που αναπτύσσει και υποβάλλει σε δοκιμές το σύστημα ΤΝ, κι επομένως μπορεί να είναι αναγκαία η συνεργασία με άλλα ενδιαφερόμενα μέρη εντός ή εκτός του οργανισμού. Συνιστάται ιδιαίτερος να καταγράφονται όλα τα αποτελέσματα τόσο από τεχνικής όσο και από διοικητικής άποψης ώστε να διασφαλίζεται ότι η επίλυση των προβλημάτων μπορεί να γίνεται κατανοητή σε όλα τα επίπεδα της δομής διακυβέρνησης.
- (118) Στόχος του καταλόγου αξιολόγησης είναι να καθοδηγεί τους επαγγελματίες της ΤΝ κατά την ανάπτυξη, εγκατάσταση και χρήση της αξιόπιστης ΤΝ. Η αξιολόγηση θα πρέπει να είναι προσαρμοσμένη στη συγκεκριμένη περίπτωση χρήσης με αναλογικό τρόπο. Κατά την πιλοτική φάση ενδέχεται να

αποκαλυφθούν συγκεκριμένοι τομείς ευαίσθητου χαρακτήρα και η ανάγκη για περαιτέρω προδιαγραφές στις περιπτώσεις αυτές θα αξιολογηθεί στο επόμενο στάδιο. Ο κατάλογος αξιολόγησης μπορεί μεν να μην παρέχει συγκεκριμένες απαντήσεις για την επίλυση των ζητημάτων που εγείρονται, ωστόσο ενθαρρύνει τον αναστοχασμό ως προς τα βήματα που μπορούν να συμβάλουν στη διασφάλιση της αξιοπιστίας των συστημάτων TN και στα πιθανά μέτρα που θα πρέπει να ληφθούν προς την κατεύθυνση αυτή.

#### *Σχέση με ισχύουσες νομοθετικές διατάξεις και διαδικασίες*

- (119) Εξίσου σημαντικό για όσους ασχολούνται με την ανάπτυξη, εγκατάσταση και χρήση TN είναι να αναγνωρίζουν ότι υπάρχουν διάφορες ισχύουσες νομοθετικές διατάξεις που προβλέπουν συγκεκριμένες διαδικασίες και απαγορεύουν συγκεκριμένα αποτελέσματα, οι οποίες μπορεί να αλληλεπικαλύπτονται ή να συμπίπτουν με ορισμένα από τα μέτρα που παρατίθενται στον κατάλογο αξιολόγησης. Για παράδειγμα, η νομοθεσία για την προστασία των δεδομένων προβλέπει σειρά νομικών απαιτήσεων που πρέπει να τηρούν όσοι ασχολούνται με τη συλλογή και επεξεργασία δεδομένων προσωπικού χαρακτήρα. Ωστόσο, επειδή η αξιόπιστη TN προϋποθέτει επίσης το δεοντολογικό χειρισμό δεδομένων, εσωτερικές διαδικασίες και πολιτικές που αποβλέπουν στην εξασφάλιση συμμόρφωσης με τη νομοθεσία για την προστασία των δεδομένων ενδέχεται επίσης να διευκολύνουν τη δεοντολογική διαχείριση των δεδομένων και μπορούν, ως εκ τούτου, να συμπληρώνουν τις υφιστάμενες νομικές διαδικασίες. Ωστόσο, η συμμόρφωση με τον παρόντα κατάλογο αξιολόγησης δεν αποτελεί τεκμήριο νομικής συμμόρφωσης ούτε προορίζεται για την παροχή κατευθύνσεων για τη διασφάλιση της συμμόρφωσης με την εφαρμοστέα νομοθεσία. Αντ' αυτού, στόχος του παρόντος καταλόγου αξιολόγησης είναι η παροχή μιας δέσμης συγκεκριμένων ερωτήσεων στους αποδέκτες, σε μια προσπάθεια να διασφαλιστεί ότι η προσέγγισή τους όσον αφορά την ανάπτυξη και την εγκατάσταση της TN είναι προσανατολισμένη προς την αξιόπιστη TN και επιδιώκει να την εξασφαλίσει.
- (120) Ομοίως, πολλοί επαγγελματίες της TN διαθέτουν ήδη υφιστάμενα εργαλεία αξιολόγησης και διεργασίες ανάπτυξης λογισμικού που διασφαλίζουν τη συμμόρφωση και με μη νομοθετικά πρότυπα. Η παρακάτω αξιολόγηση δεν θα πρέπει κατ' ανάγκη να διενεργείται μεμονωμένα, αλλά μπορεί να ενσωματώνεται στις εν λόγω υφιστάμενες πρακτικές.

### **ΚΑΤΑΛΟΓΟΣ ΑΞΙΟΛΟΓΗΣΗΣ ΤΗΣ ΑΞΙΟΠΙΣΤΙΑΣ ΤΗΣ TN (ΠΙΛΟΤΙΚΗ ΕΚΔΟΣΗ)**

#### **1. Ανθρώπινη παρέμβαση και εποπτεία**

##### ***Θεμελιώδη δικαιώματα:***

- ✓ Σε περιπτώσεις χρήσης στις οποίες ενδέχεται να υπάρχουν αρνητικές επιπτώσεις σε θεμελιώδη δικαιώματα, πραγματοποιήσατε εκτίμηση επιπτώσεων στα θεμελιώδη δικαιώματα; Εντοπίσατε και τεκμηριώσατε ενδεχόμενες αντισταθμιστικές ρυθμίσεις μεταξύ των διαφόρων αρχών και δικαιωμάτων;
- ✓ Υπάρχει αλληλεπίδραση του συστήματος TN με διαδικασίες λήψης αποφάσεων από ανθρώπους-τελικούς χρήστες (π.χ. διατύπωση προτεινόμενων ενεργειών ή αποφάσεων, παρουσίαση επιλογών);
  - Στις περιπτώσεις αυτές, υπάρχει κίνδυνος να επηρεάζει το σύστημα TN την ανθρώπινη αυτονομία παρεμβαίνοντας στη διαδικασία λήψης αποφάσεων του τελικού χρήστη με μη προβλεπόμενο τρόπο;
  - Εξετάσατε το κατά πόσο το σύστημα TN θα πρέπει να ενημερώνει τους χρήστες ότι μια

απόφαση, κάποιο περιεχόμενο, μια συμβουλή ή ένα αποτέλεσμα προέκυψε από αλγοριθμική απόφαση;

- Σε περίπτωση που το σύστημα TN περιλαμβάνει διαλογορομπότ (chat-bot) ή σύστημα συνομιλίας, γνωστοποιήθηκε στους ανθρώπους-τελικούς χρήστες ότι αλληλεπιδρούν με μη ανθρώπινο πράκτορα;

#### ***Ανθρώπινη παρέμβαση:***

- ✓ Σε περίπτωση που το σύστημα TN υλοποιείται στο πλαίσιο εργασιακής διαδικασίας, εξετάσατε τον επιμερισμό καθηκόντων μεταξύ του συστήματος TN και των ανθρώπων-εργαζομένων ώστε οι αλληλεπιδράσεις να είναι ουσιαστικές και να υπάρχει κατάλληλη ανθρώπινη εποπτεία και έλεγχος;
  - Το σύστημα TN ενισχύει ή επαυξάνει ανθρώπινες ικανότητες;
  - Προβλέψατε κατάλληλες ασφαλιστικές δικλίδες προκειμένου να αποτραπεί η υπερβολική εμπιστοσύνη στο σύστημα TN ή η υπερβολική εξάρτηση από αυτό στο πλαίσιο των εργασιακών διαδικασιών;

#### ***Ανθρώπινη εποπτεία:***

- ✓ Εξετάσατε ποιο θα ήταν το κατάλληλο επίπεδο ανθρώπινου ελέγχου για το συγκεκριμένο σύστημα TN και τη συγκεκριμένη περίπτωση χρήσης;
  - Μπορείτε να περιγράψετε το επίπεδο ανθρώπινου ελέγχου ή παρέμβασης, αν συντρέχει περίπτωση; Ποιος είναι υπεύθυνος για τον ανθρώπινο έλεγχο και σε ποιες στιγμές ή με ποια εργαλεία γίνεται η ανθρώπινη παρέμβαση;
  - Έχετε θέσει σε εφαρμογή μηχανισμούς και μέτρα που διασφαλίζουν τον εν λόγω δυνητικό ανθρώπινο έλεγχο ή την ανθρώπινη εποπτεία, ή διασφαλίζουν ότι η συνολική ευθύνη για τη λήψη των αποφάσεων ανήκει σε ανθρώπους;
  - Λάβατε μέτρα που καθιστούν δυνατό τον διαχειριστικό έλεγχο και αποκαθιστούν προβλήματα που σχετίζονται με τη διακυβέρνηση της αυτονομίας της TN;
- ✓ Σε περίπτωση ύπαρξης αυτοεκπαιδευόμενου ή αυτόνομου συστήματος TN ή περίπτωσης χρήσης, θέσατε σε εφαρμογή πιο εξειδικευμένους μηχανισμούς ελέγχου και εποπτείας;
  - Τι είδους μηχανισμούς ανίχνευσης και αντιμετώπισης δημιουργήσατε προκειμένου να αξιολογείτε αν μπορεί να προκύψει κάποιο πρόβλημα;
  - Προβλέψατε κάποιο «πλήκτρο διακοπής» ή κάποια διαδικασία ώστε μια λειτουργία να μπορεί να διακοπεί με ασφάλεια αν παραστεί ανάγκη; Με τη διαδικασία αυτή διακόπτεται η εν λόγω διεργασία στο σύνολό της, εν μέρει ή ανατίθεται ο έλεγχος σε άνθρωπο;

## **2. Τεχνική στιβαρότητα και ασφάλεια**

#### ***Ανθεκτικότητα σε επιθέσεις και προστασία:***

- ✓ Αξιολογήσατε πιθανές μορφές επιθέσεων στις οποίες θα μπορούσε να είναι ευάλωτο το σύστημα

TN;

- Ειδικότερα, εξετάσατε διάφορους τύπους και μορφές ευπαθειών, όπως η ρύπανση δεδομένων, οι υλικές υποδομές και οι κυβερνοεπιθέσεις;
- ✓ Εφαρμόσατε μέτρα ή συστήματα για να διασφαλίσετε την ακεραιότητα και την ανθεκτικότητα του συστήματος TN έναντι πιθανών επιθέσεων;
- ✓ Αξιολογήσατε πώς συμπεριφέρεται το σύστημά σας σε απροσδόκητες καταστάσεις και περιβάλλοντα;
- ✓ Εξετάσατε αν και σε ποιο βαθμό το σύστημά σας μπορεί να είναι διπλής χρήσης; Αν ναι, λάβατε κατάλληλα προληπτικά μέτρα για να αποτρέψετε αυτό το ενδεχόμενο (όπως π.χ. να μην δημοσιεύσετε την έρευνα ή να μην εγκαταστήσετε το σύστημα);

***Εφεδρικό σχέδιο και γενική ασφάλεια:***

- ✓ Μεριμνήσατε ώστε το σύστημά σας να διαθέτει επαρκές εφεδρικό σχέδιο σε περίπτωση που αντιμετωπίσει επιθέσεις από αντίπαλες πλευρές ή άλλες απροσδόκητες καταστάσεις (π.χ. διαδικασίες τεχνικής εναλλαγής ή έγκριση από άνθρωπο-χειριστή πριν από την πραγματοποίηση μιας ενέργειας);
- ✓ Λάβατε υπόψη το επίπεδο κινδύνου που εγκυμονεί το σύστημα TN στη συγκεκριμένη περίπτωση χρήσης;
  - Έχετε θέσει σε εφαρμογή κάποια διεργασία για τη μέτρηση και την αξιολόγηση των κινδύνων και της ασφάλειας;
  - Παρείχατε τις απαιτούμενες πληροφορίες σε περίπτωση κινδύνου για τη σωματική ακεραιότητα ανθρώπων;
  - Εξετάσατε το ενδεχόμενο σύναψης ασφαλιστήριου συμβολαίου για την κάλυψη τυχόν ζημίας από το σύστημα TN;
  - Εντοπίσατε τους δυνητικούς κινδύνους για την ασφάλεια που ενέχουν (άλλες) προβλέψιμες χρήσεις της τεχνολογίας, συμπεριλαμβανομένης της τυχαίας ή κακόβουλης κατάχρησής της; Υπάρχει σχέδιο για την άμβλυνση ή τη διαχείριση αυτών των κινδύνων;
- ✓ Εξετάσατε αν υπάρχει πιθανότητα το σύστημα TN να προκαλέσει ζημία ή βλάβη σε χρήστες ή τρίτους; Αν ναι, εξετάσατε τον βαθμό πιθανότητας, την ενδεχόμενη ζημία, το κοινό που θα επηρεαστεί και τον βαθμό βαρύτητας;
  - Εφόσον υπάρχει κίνδυνος το σύστημα TN να προκαλέσει ζημία, εξετάσατε τους κανόνες περί ευθύνης και προστασίας των καταναλωτών, και με ποιον τρόπο τους λάβατε υπόψη;
  - Λάβατε υπόψη τις ενδεχόμενες επιπτώσεις ή τον κίνδυνο για την ασφάλεια του περιβάλλοντος ή των ζώων;
  - Συνεκτιμήσατε στο πλαίσιο της ανάλυσης κινδύνου αν προβλήματα προστασίας ή δικτύου (π.χ. κίνδυνοι κυβερνοπροστασίας) ενέχουν κινδύνους ασφάλειας ή πρόκλησης ζημίας λόγω ακούσιας συμπεριφοράς του συστήματος TN;
- ✓ Εκτιμήσατε τις πιθανές επιπτώσεις αστοχίας του οικείου συστήματος TN που έχει ως αποτέλεσμα την παροχή λανθασμένων αποτελεσμάτων, τη μη διαθεσιμότητα του συστήματός σας ή την παροχή

απαράδεκτων από κοινωνιακής άποψης αποτελεσμάτων από το σύστημά σας (π.χ. πρακτικές που εισάγουν διακρίσεις);

- Καθορίσατε όρια και ρυθμίσεις ώστε τα παραπάνω σενάρια να ενεργοποιήσουν εναλλακτικά/εφεδρικά σχέδια;
- Εκπονήσατε και υποβάλλατε σε δοκιμές τα εφεδρικά σχέδια;

#### **Ακρίβεια**

- ✓ Εξετάσατε ποιο επίπεδο και ποιος ορισμός της ακρίβειας θα απαιτείται στο πλαίσιο του συστήματος TN και της περίπτωσης χρήσης;
  - Εξετάσατε πώς γίνεται η μέτρηση και η διασφάλιση της ακρίβειας;
  - Θέσατε σε εφαρμογή μέτρα που διασφαλίζουν ότι τα δεδομένα που χρησιμοποιούνται είναι ολοκληρωμένα και επικαιροποιημένα;
  - Θέσατε σε εφαρμογή μέτρα που αξιολογούν αν υπάρχει ανάγκη για πρόσθετα δεδομένα, π.χ. για τη βελτίωση της ακρίβειας ή για την εξάλειψη της μεροληψίας;
- ✓ Αξιολογήσατε τη βλάβη που θα μπορούσε να προκληθεί αν το σύστημα TN κάνει ανακριβείς προβλέψεις;
- ✓ Προβλέψατε τρόπους για να μετράται κατά πόσο το σύστημά σας κάνει απαράδεκτα υψηλό αριθμό ανακριβών προβλέψεων;
- ✓ Αν γίνονται ανακριβείς προβλέψεις, θέσατε σε εφαρμογή σειρά μέτρων για την επίλυση του προβλήματος;

#### **Αξιοπιστία και αναπαραγωγιμότητα:**

- ✓ Θέσατε σε εφαρμογή στρατηγική για την παρακολούθηση και τον έλεγχο του κατά πόσο το σύστημα TN ανταποκρίνεται στους στόχους, στους σκοπούς και στις εφαρμογές για τις οποίες προορίζεται;
    - Ελέγξατε αν πρέπει να λαμβάνονται υπόψη συγκεκριμένα πλαίσια ή ειδικές συνθήκες προκειμένου να διασφαλίζεται η αναπαραγωγιμότητα;
    - Θέσατε σε εφαρμογή διεργασίες ή μεθόδους επαλήθευσης που μετρούν και διασφαλίζουν τις διάφορες πτυχές της αξιοπιστίας και της αναπαραγωγιμότητας;
    - Θέσατε σε εφαρμογή διεργασίες που περιγράφουν πότε ένα σύστημα TN θεωρείται ότι αστοχεί σε ορισμένους τύπους ρυθμίσεων;
    - Τεκμηριώσατε με σαφήνεια και υλοποιήσατε επιχειρησιακά αυτές τις διεργασίες για τη δοκιμή και την επαλήθευση της αξιοπιστίας των συστημάτων TN;
- Θέσατε σε εφαρμογή μηχανισμούς ή διαύλους επικοινωνίας με τους (τελικούς) χρήστες ώστε να διαβεβαιώνονται για την αξιοπιστία του συστήματος TN;

### **3. Ιδιωτική ζωή και διακυβέρνηση των δεδομένων**

#### **Σεβασμός της ιδιωτικής ζωής και προστασία των δεδομένων:**



- ✓ Ανάλογα με την περίπτωση χρήσης, δημιουργήσατε μηχανισμούς που επιτρέπουν σε άλλους να επισημαίνουν ζητήματα σχετικά με την ιδιωτική ζωή και την προστασία των δεδομένων όσον αφορά τις διεργασίες συλλογής δεδομένων (τόσο για εκπαίδευση όσο και για λειτουργία) και επεξεργασίας δεδομένων του συστήματος TN;
- ✓ Εξετάσατε τον τύπο και το εύρος των δεδομένων που περιέχουν τα σύνολα δεδομένων σας (π.χ. κατά πόσο περιέχουν δεδομένα προσωπικού χαρακτήρα);
- ✓ Εξετάσατε τρόπους για να αναπτύξετε το σύστημα TN ή να εκπαιδεύσετε το μοντέλο με μηδενική ή ελάχιστη χρήση ενδεχομένως ευαίσθητων δεδομένων ή δεδομένων προσωπικού χαρακτήρα;
- ✓ Ενσωματώσατε μηχανισμούς ειδοποίησης και ελέγχου επί των προσωπικών δεδομένων ανάλογα με την περίπτωση χρήσης (όπως εξασφάλιση συναίνεσης με έγκυρο τρόπο και δυνατότητα ανάκλησής της, όταν συντρέχει περίπτωση);
- ✓ Λάβατε μέτρα για τη βελτίωση της προστασίας της ιδιωτικής ζωής, όπως μέσω κρυπτογράφησης, ανωνυμοποίησης και ομαδοποίησης;
- ✓ Αν υπάρχει υπεύθυνος προστασίας δεδομένων, μεριμνήσατε ώστε να συμμετέχει το πρόσωπο αυτό στη διαδικασία από τα αρχικά της στάδια;

***Ποιότητα και ακεραιότητα δεδομένων:***

- ✓ Ευθυγραμμίσατε το σύστημά σας με τυχόν σχετικά πρότυπα (π.χ. ISO, IEEE) ή ευρέως χρησιμοποιούμενα πρωτόκολλα για τις καθημερινές σας δραστηριότητες διαχείρισης και διακυβέρνησης δεδομένων;
- ✓ Δημιουργήσατε μηχανισμούς εποπτείας για τη συλλογή, αποθήκευση, επεξεργασία και χρήση των δεδομένων;
- ✓ Αξιολογήσατε τον βαθμό στο οποίο έχετε τον έλεγχο της ποιότητας των εξωτερικών πηγών δεδομένων που χρησιμοποιούνται;
- ✓ Θέσατε σε εφαρμογή διεργασίες που διασφαλίζουν την ποιότητα και την ακεραιότητα των δεδομένων σας; Εξετάσατε άλλες διεργασίες; Με ποιον τρόπο επαληθεύετε ότι τα σύνολα δεδομένων σας δεν έχουν θιγεί ή παραβιαστεί;

***Πρόσβαση σε δεδομένα:***

- ✓ Τι πρωτόκολλα, διεργασίες και διαδικασίες ακολουθήθηκαν για τη διαχείριση και τη διασφάλιση της χρηστής διακυβέρνησης των δεδομένων;
  - Εξετάσατε ποιοι έχουν πρόσβαση σε δεδομένα των χρηστών και υπό ποιες συνθήκες;
  - Διασφαλίσατε ότι τα πρόσωπα αυτά είναι ειδικευμένα και ότι είναι αναγκαίο να προσπελάσουν τα δεδομένα, και ότι έχουν τις απαιτούμενες δεξιότητες για να κατανοήσουν τις λεπτομέρειες της πολιτικής για την προστασία των δεδομένων;
  - Μεριμνήσατε για την ύπαρξη μηχανισμού εποπτείας ώστε να καταγράφεται πότε, πού, πώς, από ποιον και για ποιο σκοπό προσπελάστηκαν δεδομένα;

**4. Διαφάνεια**

### ***Ιχνηλασιμότητα:***

- ✓ Θέσατε σε εφαρμογή μέτρα που διασφαλίζουν ιχνηλασιμότητα; Τα μέτρα αυτά μπορεί να περιλαμβάνουν την τεκμηρίωση:
  - των μεθόδων που χρησιμοποιούνται για τον σχεδιασμό και την ανάπτυξη του αλγοριθμικού συστήματος:
    - αν πρόκειται για σύστημα TN που βασίζεται σε κανόνες, θα πρέπει να καταγράφεται η μέθοδος προγραμματισμού ή ο τρόπος δημιουργίας του μοντέλου
    - αν πρόκειται για σύστημα TN που βασίζεται στη μάθηση, θα πρέπει να καταγράφεται η μέθοδος για την εκπαίδευση του αλγορίθμου, όπως, μεταξύ άλλων, ποια δεδομένα εισόδου συλλέχθηκαν και επιλέχθηκαν, και πώς έγινε αυτό
  - των μεθόδων που χρησιμοποιούνται για τη δοκιμή και την επικύρωση του αλγοριθμικού συστήματος:
    - αν πρόκειται για σύστημα TN που βασίζεται σε κανόνες, θα πρέπει να καταγράφονται τα σενάρια ή οι περιπτώσεις που χρησιμοποιούνται για τον έλεγχο και την επικύρωση
    - αν πρόκειται για σύστημα TN που βασίζεται στη μάθηση, θα πρέπει να καταγράφονται πληροφορίες σχετικά με τα δεδομένα που χρησιμοποιούνται για τη δοκιμή και την επικύρωση
  - των αποτελεσμάτων του αλγοριθμικού συστήματος:
    - θα πρέπει να καταγράφονται τα αποτελέσματα του αλγορίθμου ή οι αποφάσεις που λαμβάνει, καθώς και τυχόν άλλες αποφάσεις που προκύπτουν από διάφορες άλλες περιπτώσεις (π.χ. για άλλες υποομάδες χρηστών).

### ***Επεξηγησιμότητα:***

- ✓ Αξιολογήσατε τον βαθμό στον οποίο οι αποφάσεις και, ως εκ τούτου, τα αποτελέσματα που προκύπτουν από το σύστημα TN είναι κατανοητά;
- ✓ Διασφαλίσατε ότι η επεξήγηση των λόγων για τους οποίους το σύστημα έκανε μια συγκεκριμένη επιλογή που οδήγησε σε ορισμένο αποτέλεσμα μπορεί να γίνει κατανοητή από όλους τους χρήστες που ενδέχεται να θέλουν επεξήγηση;
- ✓ Αξιολογήσατε τον βαθμό στον οποίο η απόφαση του συστήματος επηρεάζει τις διεργασίες λήψης αποφάσεων του οργανισμού;
- ✓ Εξετάσατε τους λόγους για τους οποίους το συγκεκριμένο σύστημα εγκαταστάθηκε στον συγκεκριμένο τομέα;
- ✓ Αξιολογήσατε το επιχειρηματικό μοντέλο που αφορά αυτό το σύστημα (π.χ. πώς δημιουργεί αξία για τον οργανισμό);
- ✓ Σχεδιάσατε το σύστημα TN έχοντας εξαρχής ως γνώμονα την ερμηνευσιμότητα;
  - Διερευνήσατε και επιχειρήσατε να χρησιμοποιήσετε το απλούστερο και πλέον ερμηνεύσιμο μοντέλο για την υπό εξέταση εφαρμογή;
  - Εξετάσατε αν μπορείτε να αναλύσετε τα δεδομένα σας για την εκπαίδευση και τη δοκιμή; Μπορείτε να τα τροποποιείτε και να τα επικαιροποιείτε διαχρονικά;

- Εξετάσατε αν μετά την εκπαίδευση και την ανάπτυξη του μοντέλου έχετε επιλογές ως προς την εξέταση της ερμηνευσιμότητας, ή αν έχετε πρόσβαση στην εσωτερική ροή εργασιών του μοντέλου;

#### **Επικοινωνία:**

- ✓ Γνωστοποιήσατε στους (τελικούς) χρήστες —μέσω δήλωσης αποποίησης ευθύνης ή άλλου τρόπου— ότι αλληλεπιδρούν με σύστημα TN και όχι με άλλον άνθρωπο; Έχετε χαρακτηρίσει το σύστημα TN σας ως τέτοιο;
- ✓ Θέσατε σε εφαρμογή μηχανισμούς για την ενημέρωση των χρηστών σχετικά με τους λόγους και τα κριτήρια στα οποία βασίζονται τα αποτελέσματα του συστήματος TN;
  - Η σχετική γνωστοποίηση στους προβλεπόμενους χρήστες γίνεται με σαφή και κατανοητό τρόπο;
  - Δημιουργήσατε διεργασίες που λαμβάνουν υπόψη την αναπληροφόρηση που παρέχουν οι χρήστες και τη χρησιμοποιούν για την προσαρμογή του συστήματος;
  - Γνωστοποιήσατε επίσης δυνητικούς ή θεωρούμενους κινδύνους, όπως η μεροληψία;
  - Ανάλογα με την περίπτωση χρήσης, εξετάσατε επίσης το ζήτημα της επικοινωνίας και της διαφάνειας έναντι άλλων ακροατηρίων, τρίτων ή του ευρέος κοινού;
- ✓ Διευκρινίσατε ποιος είναι ο σκοπός του συστήματος TN και ποιος ή τι μπορεί να ωφεληθεί από το προϊόν/την υπηρεσία;
  - Έχουν προσδιοριστεί και γνωστοποιηθεί με σαφήνεια τα σενάρια χρήσης του προϊόντος, συμπεριλαμβανομένης της εξέτασης και εναλλακτικών μορφών επικοινωνίας, ώστε να διασφαλίζεται ότι είναι κατανοητά και κατάλληλα για τον χρήστη στον οποίο απευθύνονται;
  - Ανάλογα με την περίπτωση χρήσης, λάβατε υπόψη την ανθρώπινη ψυχολογία και τους πιθανούς περιορισμούς, όπως ο κίνδυνος σύγχυσης, σφάλματος επιβεβαίωσης ή γνωστικής κόπωσης;
- ✓ Γνωστοποιήσατε με σαφήνεια τα χαρακτηριστικά, τους περιορισμούς και τις ενδεχόμενες ελλείψεις του συστήματος TN:
  - σε περίπτωση ανάπτυξης: σε οποιονδήποτε το εγκαθιστά σε ένα προϊόν ή μια υπηρεσία;
  - σε περίπτωση εγκατάστασης: στον τελικό χρήστη ή καταναλωτή;

## **5. Πολυμορφία, απαγόρευση των διακρίσεων και δικαιοσύνη**

#### **Αποφυγή αθέμιτης μεροληψίας:**

- ✓ Μεριμνήσατε για την ύπαρξη στρατηγικής ή σειράς διαδικασιών ώστε να αποφεύγεται η δημιουργία ή η ενίσχυση της αθέμιτης μεροληψίας στο σύστημα TN, τόσο όσον αφορά τη χρήση δεδομένων εισόδου όσο και για τον σχεδιασμό των αλγορίθμων;
  - Αξιολογήσατε και συνεκτιμήσατε τους πιθανούς περιορισμούς που προκύπτουν από τη σύνθεση των χρησιμοποιούμενων συνόλων δεδομένων;
  - Λάβατε υπόψη την πολυμορφία και την αντιπροσωπευτικότητα των χρηστών στα δεδομένα; Κάνατε δοκιμές για την ύπαρξη ειδικών πληθυσμών ή προβληματικών περιπτώσεων χρήσης;

- Διερευνήσατε και χρησιμοποιήσατε διαθέσιμα τεχνικά εργαλεία προκειμένου να κατανοήσετε καλύτερα τα δεδομένα, το μοντέλο και τις επιδόσεις;
  - Θέσατε σε εφαρμογή διεργασίες δοκιμής και παρακολούθησης ενδεχόμενων συστημικών σφαλμάτων κατά το στάδιο της ανάπτυξης, της εγκατάστασης και της χρήσης του συστήματος;
  - ✓ Ανάλογα με την περίπτωση χρήσης, μεριμνήσατε για την ύπαρξη μηχανισμού που επιτρέπει σε τρίτους να επισημαίνουν ζητήματα μεροληψίας, διακρίσεων ή μη ικανοποιητικής απόδοσης του συστήματος TN;
    - Εξετάσατε σαφή βήματα και τρόπους επικοινωνίας ως προς το πώς και σε ποιον είναι δυνατό να επισημαίνονται αυτά τα ζητήματα;
    - Εκτός από τους (τελικούς) χρήστες λάβατε υπόψη και άλλους που ενδεχομένως επηρεάζονται έμμεσα από το σύστημα TN;
  - ✓ Εξετάσατε αν υπάρχει πιθανή μεταβλητότητα σε αποφάσεις που λαμβάνονται υπό τις ίδιες συνθήκες;
    - Αν ναι, εξετάσατε σε τι μπορεί να οφείλεται;
    - Αν υπάρχει μεταβλητότητα, δημιουργήσατε μηχανισμό μέτρησης ή εκτίμησης των ενδεχόμενων επιπτώσεων που μπορεί να έχει η εν λόγω μεταβλητότητα στα θεμελιώδη δικαιώματα;
  - ✓ Μεριμνήσατε για την υιοθέτηση κατάλληλου λειτουργικού ορισμού της «δικαιοσύνης» τον οποίο εφαρμόζετε στον σχεδιασμό συστημάτων TN;
    - Ο ορισμός που έχετε υιοθετήσει χρησιμοποιείται ευρέως; Εξετάσατε άλλους ορισμούς προτού επιλέξετε τον συγκεκριμένο;
    - Μεριμνήσατε για ποσοτική ανάλυση ή μέθοδο μέτρησης για την αξιολόγηση και τη δοκιμή του εφαρμοζόμενου ορισμού της δικαιοσύνης;
    - Δημιουργήσατε μηχανισμούς για τη διασφάλιση της δικαιοσύνης στα οικεία συστήματα TN; Εξετάσατε άλλους πιθανούς μηχανισμούς;
- Προσβασιμότητα και καθολικός σχεδιασμός:**
- ✓ Μεριμνήσατε ώστε το σύστημα TN να καλύπτει ευρύ φάσμα ατομικών προτιμήσεων και ικανοτήτων;
    - Αξιολογήσατε αν το σύστημα TN μπορεί να χρησιμοποιηθεί από άτομα με ειδικές ανάγκες ή αναπηρίες ή από άτομα που αντιμετωπίζουν κίνδυνο αποκλεισμού; Πώς ενσωματώθηκε η δυνατότητα αυτή στον σχεδιασμό του συστήματος και πώς επαληθεύεται;
    - Μεριμνήσατε ώστε οι πληροφορίες σχετικά με το σύστημα TN να είναι προσβάσιμες και από χρήστες υποστηρικτικών τεχνολογιών;
    - Επιδιώξατε τη συμμετοχή ή ζητήσατε τη γνώμη της συγκεκριμένης κοινότητας κατά το στάδιο της ανάπτυξης του συστήματος TN;
  - ✓ Λάβατε υπόψη τις επιπτώσεις του συστήματος TN στους πιθανούς χρήστες που αποτελούν το κοινό σας;
    - Η ομάδα που συμμετείχε στη δημιουργία του συστήματος TN είναι αντιπροσωπευτική των

χρηστών που αποτελούν το κοινό-στόχο σας; Είναι αντιπροσωπευτική του ευρύτερου πληθυσμού, λαμβανομένων επίσης υπόψη άλλων ομάδων που μπορεί να επηρεάζονται παρεμπιπτόντως;

- Εξετάσατε αν υπάρχουν πρόσωπα ή ομάδες που μπορεί να επηρεάζονται δυσανάλογα από αρνητικές συνέπειες;
- Λάβατε αναπληροφόρηση από άλλες ομάδες που αντιπροσωπεύουν διαφορετικές καταβολές και βιώματα;

***Συμμετοχή των ενδιαφερόμενων μερών:***

- ✓ Εξετάσατε το ενδεχόμενο μηχανισμού που εξασφαλίζει τη συμμετοχή διαφόρων ενδιαφερόμενων μερών στην ανάπτυξη και τη χρήση του συστήματος TN;
- ✓ Προετοιμάσατε το έδαφος για την εισαγωγή του συστήματος TN στον οργανισμό σας, ενημερώνοντας και συμπεριλαμβάνοντας στη διαδικασία τους εργαζόμενους που επηρεάζονται και τους εκπροσώπους τους εκ των προτέρων;

**6. Κοινωνική και περιβαλλοντική ευημερία**

***Βιώσιμη και περιβαλλοντικά φιλική TN:***

- ✓ Θέσατε σε εφαρμογή μηχανισμούς για τη μέτρηση των περιβαλλοντικών επιπτώσεων της ανάπτυξης, της εγκατάστασης και της χρήσης του συστήματος TN (π.χ. της ενέργειας που καταναλώνουν τα κέντρα δεδομένων, της μορφής ενέργειας που χρησιμοποιούν κ.λπ.);
- ✓ Μεριμνήσατε για τη λήψη μέτρων περιορισμού των περιβαλλοντικών επιπτώσεων του κύκλου ζωής του συστήματος TN σας;

***Κοινωνικές επιπτώσεις:***

- ✓ Αν πρόκειται για σύστημα TN που αλληλεπιδρά απευθείας με ανθρώπους:
  - Εξετάσατε αν το σύστημα TN ενθαρρύνει τους ανθρώπους να αναπτύξουν συναισθηματικούς δεσμούς και ενσυναίσθηση μαζί του;
  - Διασφαλίσατε ότι το σύστημα TN παρέχει σαφείς ενδείξεις ότι η κοινωνική αλληλεπίδρασή του είναι προϊόν προσομοίωσης και ότι δεν ικανοποιεί «δείχνει κατανόηση» ή να «αισθάνεται»;
- ✓ Διασφαλίσατε ότι οι κοινωνικές επιπτώσεις του συστήματος TN είναι απόλυτα κατανοητές; Για παράδειγμα, αξιολογήσατε αν υπάρχει κίνδυνος απώλειας θέσεων εργασίας ή απαξίωσης των δεξιοτήτων του εργατικού δυναμικού; Τι μέτρα έχουν ληφθεί για την αντιστάθμιση αυτών των κινδύνων;

***Κοινωνία και δημοκρατία:***

- ✓ Αξιολογήσατε τις ευρύτερες κοινωνικές επιπτώσεις της χρήσης του συστήματος TN πέρα από το επίπεδο του μεμονωμένου (τελικού) χρήστη, π.χ. στο επίπεδο των ενδιαφερόμενων μερών που ενδεχομένως επηρεάζονται έμμεσα;

**7. Λογοδοσία**

**Ελεγχιμότητα:**

- ✓ Θέσατε σε εφαρμογή μηχανισμούς που διευκολύνουν την ελεγχιμότητα του συστήματος από εσωτερικούς και/ή ανεξάρτητους φορείς, π.χ. διασφαλίζοντας την ιχνηλασιμότητα και την καταγραφή των διεργασιών και των αποτελεσμάτων του συστήματος TN;

**Ελαχιστοποίηση και γνωστοποίηση των αρνητικών επιπτώσεων:**

- ✓ Διενεργήσατε εκτίμηση επικινδυνότητας ή επιπτώσεων για το σύστημα TN στο πλαίσιο της οποίας συνεκτιμήθηκαν τα διάφορα ενδιαφερόμενα μέρη που επηρεάζονται άμεσα και έμμεσα;
- ✓ Θεσπίσατε πλαίσια κατάρτισης και εκπαίδευσης για την ανάπτυξη πρακτικών λογοδοσίας;
  - Ποιοι εργαζόμενοι ή κλάδοι της ομάδας συμμετέχουν; Εκτείνονται και πέρα από το στάδιο της ανάπτυξης;
  - Στο πλαίσιο των εν λόγω προγραμμάτων κατάρτισης διδάσκεται και το δυνητικό νομικό πλαίσιο που διέπει το σύστημα TN;
  - Εξετάσατε το ενδεχόμενο θέσπισης «επιτροπής αξιολόγησης της δεοντολογικής TN» ή παρόμοιου μηχανισμού για την εξέταση του γενικότερου ζητήματος της λογοδοσίας και των δεοντολογικών πρακτικών, συμπεριλαμβανομένων των τυχόν γκρίζων ζωνών;
- ✓ Πέραν των εσωτερικών πρωτοβουλιών ή πλαισίων για την εποπτεία της δεοντολογίας και της λογοδοσίας, υπάρχει κάποιο άλλο είδος εξωτερικής καθοδήγησης ή έχουν θεσπιστεί επίσης διαδικασίες διαχειριστικού ελέγχου;
- ✓ Έχουν προβλεφθεί διαδικασίες μέσω των οποίων τρίτοι (π.χ. προμηθευτές, καταναλωτές, διανομείς/πωλητές) ή εργαζόμενοι μπορούν να αναφέρουν πιθανά τρωτά σημεία, κινδύνους ή συστημικά σφάλματα στο σύστημα ή την εφαρμογή TN;

**Τεκμηρίωση αντισταθμιστικών ρυθμίσεων:**

- ✓ Δημιουργήσατε μηχανισμό για την καταγραφή σχετικών συμφερόντων και αξιών που επηρεάζονται στο πλαίσιο του συστήματος TN και των πιθανών αντισταθμιστικών ρυθμίσεων μεταξύ τους;
- ✓ Ποια διαδικασία χρησιμοποιείτε για τη λήψη αποφάσεων σχετικά με τις εν λόγω αντισταθμιστικές ρυθμίσεις; Μεριμνήσατε για την τεκμηρίωση των αποφάσεων που αφορούν τις αντισταθμιστικές ρυθμίσεις;

**Δυνατότητα έννομης προστασίας:**

- ✓ Θεσπίσατε επαρκές σύνολο μηχανισμών που εξασφαλίζουν έννομη προστασία σε περίπτωση πρόκλησης βλάβης ή δυσμενούς επίπτωσης;
- ✓ Θέσατε σε εφαρμογή μηχανισμούς για την ενημέρωση των (τελικών) χρηστών/τρίτων σχετικά με τις δυνατότητες έννομης προστασίας;

**Καλούμε όλα τα ενδιαφερόμενα μέρη να εφαρμόσουν πιλοτικά αυτόν τον κατάλογο αξιολόγησης και να παράσχουν αναπληροφόρηση ως προς την εφαρμοσιμότητα, την πληρότητα και τη συνάφειά του για την εκάστοτε εφαρμογή ή πεδίο της TN, καθώς και ως προς την αλληλεπικάλυψη ή τη συμπληρωματικότητα του με υφιστάμενες διαδικασίες συμμόρφωσης ή αξιολόγησης. Βάσει της αναπληροφόρησης αυτής, στις**

**Βασικές κατευθύνσεις που προκύπτουν από το κεφάλαιο III:**

- ✓ Έκδοση **καταλόγου αξιολόγησης** της αξιοπιστίας της TN κατά την ανάπτυξη, εγκατάσταση και χρήση TN, και προσαρμογή του εν λόγω καταλόγου στη συγκεκριμένη περίπτωση χρήσης του συστήματος.
- ✓ Σημειωτέον ότι ο κατάλογος αξιολόγησης δεν θα **είναι ποτέ εξαντλητικός**. Η εξασφάλιση αξιοπιστίας TN δεν συνίσταται σε συμπλήρωση τετραγωνιδίων, αλλά σε διαρκή διαδικασία προσδιορισμού απαιτήσεων, αξιολόγησης λύσεων και διασφάλισης βελτιωμένων αποτελεσμάτων σε ολόκληρο τον κύκλο ζωής του συστήματος TN, καθώς και εξασφάλισης της συμμετοχής των ενδιαφερόμενων μερών στην όλη διαδικασία.

**Γ. ΠΑΡΑΔΕΙΓΜΑΤΑ ΕΥΚΑΙΡΙΩΝ ΚΑΙ ΚΑΙΡΙΩΝ ΠΡΟΒΛΗΜΑΤΙΣΜΩΝ ΠΟΥ ΕΓΓΕΙΡΕΙ Η TN**

(121) Στην ακόλουθη ενότητα παρέχουμε παραδείγματα ανάπτυξης και χρήσης της TN που θα πρέπει να ενθαρρύνονται, καθώς και παραδείγματα περιπτώσεων στις οποίες η ανάπτυξη, η εγκατάσταση ή η χρήση της TN μπορεί να αντιβαίνει στις αξίες μας και να εγείρει συγκεκριμένους προβληματισμούς. Θα πρέπει να υπάρχει ισορροπία μεταξύ του τι θα πρέπει και του τι μπορεί να γίνεται με τη χρήση TN, ενώ θα πρέπει να δίνεται κατάλληλη προσοχή και στο τι δεν θα πρέπει να γίνεται με τη χρήση TN.

**1. Παραδείγματα ευκαιριών που προσφέρει η αξιοπιστία TN**

(122) Η αξιοπιστία TN μπορεί να αποτελέσει μια μεγάλη ευκαιρία για τον μετριασμό των πιεστικών προκλήσεων που αντιμετωπίζει η κοινωνία, όπως η γήρανση του πληθυσμού, η αυξανόμενη κοινωνική ανισότητα και η ρύπανση του περιβάλλοντος. Οι δυνατότητες αυτές αντανακλώνται και σε παγκόσμιο επίπεδο, όπως στους στόχους βιώσιμης ανάπτυξης των Ηνωμένων Εθνών<sup>57</sup>. Η ακόλουθη ενότητα εξετάζει τον τρόπο με τον οποίο θα ενθαρρυνθεί μια ευρωπαϊκή στρατηγική για την TN που να αντιμετωπίζει ορισμένες από τις εν λόγω προκλήσεις.

**a. Δράση για το κλίμα και βιώσιμη υποδομή**

(123) Ενώ η αντιμετώπιση της κλιματικής αλλαγής θα πρέπει να αποτελεί ύψιστη προτεραιότητα για τους υπεύθυνους χάραξης πολιτικής σε ολόκληρο τον κόσμο, ο ψηφιακός μετασχηματισμός και η αξιοπιστία TN διαθέτουν μεγάλες δυνατότητες για τη μείωση του αντίκτυπου του ανθρώπου στο περιβάλλον και την επίτευξη αποδοτικής και αποτελεσματικής χρήσης της ενέργειας και των φυσικών πόρων<sup>58</sup>. Η αξιοπιστία TN μπορεί, για παράδειγμα, να συνδυαστεί με τα μαζικά δεδομένα προκειμένου να διαπιστώνονται με μεγαλύτερη ακρίβεια οι ενεργειακές ανάγκες, οδηγώντας σε αποδοτικότερη ενεργειακή υποδομή και κατανάλωση<sup>59</sup>.

<sup>57</sup> <https://sustainabledevelopment.un.org/?menu=1300>

<sup>58</sup> Ορισμένα προγράμματα της ΕΕ στοχεύουν στην ανάπτυξη έξυπνων δικτύων και αποθήκευσης ενέργειας, τα οποία έχουν τη δυνατότητα να συμβάλλουν στην επιτυχή ψηφιακή υποστήριξη της ενεργειακής μετάβασης, μεταξύ άλλων μέσω λύσεων που βασίζονται στην TN και άλλων ψηφιακών λύσεων. Προκειμένου να συμπληρώσει το έργο των εν λόγω μεμονωμένων προγραμμάτων, η Επιτροπή δρομολόγησε την πρωτοβουλία BRIDGE, επιτρέποντας στα εν εξελίξει έργα έξυπνων δικτύων και αποθήκευσης ενέργειας του προγράμματος «Ορίζων 2020» να διαμορφώσουν μια κοινή αντίληψη για διατομεακά ζητήματα: <https://www.h2020-bridge.eu/>.

<sup>59</sup> Βλ. για παράδειγμα, το έργο Encompass: <http://www.encompass-project.eu/>.

(124) Σε τομείς όπως οι δημόσιες συγκοινωνίες, τα συστήματα TN για ευφυή συστήματα μεταφορών<sup>60</sup> είναι δυνατό να χρησιμοποιηθούν για την ελαχιστοποίηση της αναμονής, τη βελτιστοποίηση της δρομολόγησης, την ενίσχυση της ανεξαρτησίας των ατόμων με προβλήματα όρασης<sup>61</sup>, τη βελτιστοποίηση των ενεργειακά αποδοτικών κινητήρων και, κατά συνέπεια, για την ενίσχυση των προσπαθειών απαλλαγής από τις ανθρακούχες εκπομπές και τη μείωση του περιβαλλοντικού αποτυπώματος, για μια κοινωνία πιο φιλική προς το περιβάλλον. Σήμερα, σε ολόκληρο τον κόσμο, ένας άνθρωπος χάνει τη ζωή του κάθε 23 δευτερόλεπτα σε αυτοκινητιστικό δυστύχημα<sup>62</sup>. Τα συστήματα TN θα μπορούσαν να συμβάλουν σημαντικά στη μείωση του αριθμού των θανατηφόρων ατυχημάτων, για παράδειγμα μέσω καλύτερων χρόνων αντίδρασης και καλύτερης τήρησης των κανόνων<sup>63</sup>.

#### b. Υγεία και ευεξία

(125) Οι αξιόπιστες τεχνολογίες TN είναι δυνατό να χρησιμοποιηθούν —και χρησιμοποιούνται ήδη— για να καταστήσουν πιο έξυπνη και πιο στοχευμένη τη θεραπεία και να συμβάλουν στην πρόληψη θανατηφόρων νόσων<sup>64</sup>. Οι ιατροί και οι επαγγελματίες του τομέα της ιατρικής μπορούν δυναμικά να εκτελούν μια πιο ακριβή και λεπτομερή ανάλυση των σύνθετων ιατρικών δεδομένων ενός ασθενούς, ακόμη και πριν από την εκδήλωση κάποιας ασθένειας, και να προσφέρουν προσαρμοσμένη προληπτική θεραπεία<sup>65</sup>. Στο πλαίσιο της γήρανσης του πληθυσμού της Ευρώπης, η TN και η ρομποτική μπορούν να αποτελέσουν πολύτιμα εργαλεία για τη συνδρομή των παρόχων φροντίδας, την υποστήριξη της φροντίδας των ηλικιωμένων<sup>66</sup> και την παρακολούθηση της κατάστασης των ασθενών σε πραγματικό χρόνο, σώζοντας ζωές<sup>67</sup>.

---

<sup>60</sup> Νέες λύσεις βασισμένες στην TN συμβάλουν στην προετοιμασία των πόλεων για το μέλλον της κινητικότητας. Βλ. για παράδειγμα το χρηματοδοτούμενο από την ΕΕ πρόγραμμα Fabulos: <https://fabulos.eu/>.

<sup>61</sup> Βλ. για παράδειγμα, το πρόγραμμα PRO4VIP, το οποίο αποτελεί μέρος της ευρωπαϊκής στρατηγικής «Όραμα 2020» (Vision 2020) για την καταπολέμηση της δυνάμενης να προληφθεί τύφλωσης, ιδίως λόγω του γήρατος. Η κινητικότητα και ο προσανατολισμός αποτέλεσαν τομείς προτεραιότητας του προγράμματος.

<sup>62</sup> <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>.

<sup>63</sup> Για παράδειγμα, το ευρωπαϊκό πρόγραμμα UP-Drive έχει ως στόχο την αντιμετώπιση των προαναφερθεισών προκλήσεων που συνδέονται με τις μεταφορές παρέχοντας συνεισφορές που επιτρέπουν τη σταδιακή αυτοματοποίηση και τη συνεργασία μεταξύ των οχημάτων, διευκολύνοντας ένα ασφαλέστερο, συμπεριληπτικότερο και οικονομικότερο σύστημα μεταφορών. <https://up-drive.eu/>.

<sup>64</sup> Βλ. για παράδειγμα το πρόγραμμα REVOLVER (Repeated Evolution of Cancer): <https://www.healtheuropa.eu/personalised-cancer-treatment/87958/> ή το πρόγραμμα Murab που διεξάγει ακριβέστερες βιοψίες και αποσκοπεί στην ταχύτερη διάγνωση του καρκίνου και άλλων ασθενειών: <https://ec.europa.eu/digital-single-market/en/news/murab-eu-funded-project-success-story>.

<sup>65</sup> Βλ. για παράδειγμα το πρόγραμμα Live INCITE: [www.karolinska.se/en/live-incite](http://www.karolinska.se/en/live-incite). Η συγκεκριμένη κοινοπραξία παρόχων υγειονομικής περίθαλψης καλεί τον κλάδο να αναπτύξει έξυπνες λύσεις TN και άλλες λύσεις ΤΠΕ που επιτρέπουν παρεμβάσεις στον τρόπο ζωής κατά την περιεγχειρητική διαδικασία. Ο στόχος είναι η προσφορά νέων καινοτόμων λύσεων ηλεκτρονικής υγείας που μπορούν να επηρεάσουν τους ασθενείς με εξατομικευμένο τρόπο προκειμένου να προβούν στις απαραίτητες ενέργειες τόσο πριν όσο και μετά από χειρουργική επέμβαση στον τρόπο ζωής τους προκειμένου να βελτιστοποιήσουν το αποτέλεσμα της υγειονομικής περίθαλψης.

<sup>66</sup> Αντικείμενο του χρηματοδοτούμενου από την ΕΕ προγράμματος CARESSES αποτελούν τα ρομπότ για τη φροντίδα των ηλικιωμένων, με έμφαση στην πολιτισμική τους ευαισθησία: προσαρμόζουν τον τρόπο δράσης τους και την ομιλία τους ανάλογα με τον πολιτισμό και τις συνήθειες του ηλικιωμένου ατόμου που βοηθούν: <http://caressesrobot.org/en/project/>. Βλ. επίσης την εφαρμογή TN που ονομάζεται Alfred, ένας εικονικός βοηθός που υποστηρίζει τους ηλικιωμένους ώστε να παραμείνουν δραστήριοι: <https://ec.europa.eu/digital-single-market/en/news/alfred-virtual-assistant-helping-older-people-stay-active>. Επιπλέον, το πρόγραμμα EMPATTICS (EMpowering PATients for a BeTTer Information and improvement of the Communication Systems) θα διερευνήσει και θα προσδιορίσει τον τρόπο με τον οποίο οι επαγγελματίες του τομέα της υγειονομικής περίθαλψης και οι ασθενείς χρησιμοποιούν τεχνολογίες ΤΠΕ, συμπεριλαμβανομένων των συστημάτων TN, για τον σχεδιασμό παρεμβάσεων με ασθενείς και για την παρακολούθηση της εξέλιξης της σωματικής και πνευματικής τους κατάστασης: [www.empattics.eu](http://www.empattics.eu).

<sup>67</sup> Βλ. για παράδειγμα, το πρόγραμμα MyHealth Avatar ([www.myhealthavatar.eu](http://www.myhealthavatar.eu)), το οποίο προσφέρει μια ψηφιακή αναπαράσταση της κατάστασης υγείας ενός ασθενούς. Το ερευνητικό πρόγραμμα ξεκίνησε μια εφαρμογή και μια ηλεκτρονική πλατφόρμα που συλλέγει και παρέχει πρόσβαση σε ψηφιακές πληροφορίες σχετικά με την μακροπρόθεσμη κατάσταση της



(126) Η αξιόπιστη ΤΝ μπορεί επίσης να συνεισφέρει σε ευρύτερη κλίμακα. Για παράδειγμα, μπορεί να εξετάσει και να προσδιορίσει τις γενικές τάσεις στον τομέα της υγειονομικής περίθαλψης και της θεραπείας<sup>68</sup>, οδηγώντας σε πιο πρώιμη διάγνωση ασθενειών, αποδοτικότερη ανάπτυξη φαρμάκων, πιο στοχευμένες θεραπείες<sup>69</sup> και, τελικά, στην προστασία περισσότερων ζωών.

c. Ποιοτική εκπαίδευση και ψηφιακός μετασχηματισμός

(127) Οι νέες τεχνολογικές, οικονομικές και περιβαλλοντικές εξελίξεις συνεπάγονται ότι η κοινωνία θα πρέπει να υιοθετήσει μια περισσότερο προδραστική στάση. Οι κυβερνήσεις, οι επικεφαλής της βιομηχανίας, τα εκπαιδευτικά ιδρύματα και τα συνδικάτα φέρουν την ευθύνη να οδηγήσουν τους πολίτες στη νέα ψηφιακή εποχή, εξασφαλίζοντας ότι διαθέτουν τις κατάλληλες δεξιότητες για να καλύψουν τις μελλοντικές θέσεις εργασίας. Οι τεχνολογίες της αξιόπιστης ΤΝ θα μπορούσαν να συμβάλουν στην ακριβέστερη πρόβλεψη των θέσεων εργασίας και των επαγγελματιών που θα διαταραχθούν από την τεχνολογία, τους νέους ρόλους που θα δημιουργηθούν και τις δεξιότητες που θα απαιτούνται. Κάτι τέτοιο θα μπορούσε να βοηθήσει τις κυβερνήσεις, τα συνδικάτα και τη βιομηχανία να σχεδιάσουν την (επανε)ειδίκευση των εργαζομένων. Θα μπορούσε να προσφέρει επίσης στους πολίτες που ενδεχομένως αισθάνονται ότι απειλούνται από απολύσεις μια πορεία ανάπτυξης σε έναν νέο ρόλο.

(128) Επιπλέον, η ΤΝ μπορεί να αποτελέσει ένα εξαιρετικό εργαλείο για την καταπολέμηση των εκπαιδευτικών ανισοτήτων και τη δημιουργία εξατομικευμένων και προσαρμόσιμων εκπαιδευτικών προγραμμάτων που θα μπορούσαν να βοηθήσουν τους πάντες να αποκτήσουν νέα προσόντα, δεξιότητες και ικανότητες σύμφωνα με τις μαθησιακές ικανότητές τους<sup>70</sup>. Θα μπορούσε επίσης να βελτιώσει τόσο την ταχύτητα μάθησης όσο και την ποιότητα της εκπαίδευσης — από το δημοτικό σχολείο έως το πανεπιστήμιο.

## 2. Παραδείγματα καίριων προβληματισμών που εγείρει η ΤΝ

(129) Ένας καίριος προβληματισμός σε σχέση με την ΤΝ προκύπτει σε περίπτωση παραβίασης κάποιας από τις συνιστώσες της αξιόπιστης ΤΝ. Πολλοί από τους προβληματισμούς που παρατίθενται κατωτέρω εμπνέονται ήδη στο πεδίο εφαρμογής των υφιστάμενων νομικών απαιτήσεων, οι οποίες είναι υποχρεωτικές και θα πρέπει συνεπώς να τηρούνται. Ωστόσο, ακόμη και σε περιπτώσεις όπου έχει αποδειχθεί η συμμόρφωση με τις νομικές απαιτήσεις, ενδέχεται να μην καλύπτουν το πλήρες φάσμα των δεοντολογικών προβληματισμών που μπορεί να προκύψουν. Καθώς η αντίληψή μας για την επάρκεια των κανόνων και των δεοντολογικών αρχών συνεχώς εξελίσσεται και μπορεί να αλλάξει με την πάροδο του χρόνου, ο ακόλουθος μη εξαντλητικός κατάλογος προβληματισμών μπορεί να συντομευθεί, να επεκταθεί, να τροποποιηθεί ή να επικαιροποιηθεί στο μέλλον.

a. Ταυτοποίηση και παρακολούθηση ατόμων με την ΤΝ

---

υγείας σας. Έχει τη μορφή ενός δια βίου συντρόφου υγείας («άβαταρ»). Το MyHealthAvatar προβλέπει επίσης τον κίνδυνο για εμφάνιση εγκεφαλικών επεισοδίων, διαβήτη, καρδιαγγειακών παθήσεων και υπέρτασης.

<sup>68</sup> Βλ. για παράδειγμα, το πρόγραμμα ENRICHME ([www.enrichme.eu](http://www.enrichme.eu)), το οποίο αντιμετωπίζει την προοδευτική μείωση των γνωστικών ικανοτήτων στον γηράσκοντα πληθυσμό. Μια ολοκληρωμένη πλατφόρμα για την αυτόνομη διαβίωση υποβοηθούμενη από το περιβάλλον (AAL) και ένα ρομπότ κινητής υπηρεσίας για μακροπρόθεσμη παρακολούθηση και αλληλεπίδραση θα βοηθήσουν τους ηλικιωμένους να παραμείνουν ανεξάρτητοι και δραστήριοι για μεγαλύτερο χρονικό διάστημα.

<sup>69</sup> Βλ. για παράδειγμα τη χρήση της ΤΝ από τη Sophia Genetics, η οποία αξιοποιεί τα στατιστικά συμπεράσματα, την αναγνώριση προτύπων και τη μηχανική μάθηση για τη μεγιστοποίηση της αξίας των δεδομένων γονιδιωματικής και ραδιομυϊκής: <https://www.sophiagenetics.com/home.html>.

<sup>70</sup> Βλ. για παράδειγμα, το πρόγραμμα MaTHiSiS, που έχει ως στόχο την παροχή λύσεων για μάθηση βάσει επίδρασης (affect-based learning) σε ένα άνετο μαθησιακό περιβάλλον, το οποίο περιλαμβάνει τεχνολογικές συσκευές υψηλής τεχνολογίας και αλγόριθμους: (<http://mathisis-project.eu/>). Βλ. επίσης την πλατφόρμα Watson Classroom της IBM ή την πλατφόρμα της Century Tech.

(130) Η ΤΝ καθιστά δυνατή την ολοένα και πιο αποδοτική ταυτοποίηση μεμονωμένων προσώπων τόσο από δημόσιους όσο και από ιδιωτικούς φορείς. Αξιοσημείωτα παραδείγματα κλιμακούμενης τεχνολογίας ταυτοποίησης ΤΝ είναι η αναγνώριση προσώπου και άλλες ακούσιες μέθοδοι ταυτοποίησης με χρήση βιομετρικών δεδομένων (π.χ. ανίχνευση ψεύδους, αξιολόγηση προσωπικότητας μέσω μικροεκφράσεων και αυτόματη ανίχνευση φωνής). Η ταυτοποίηση των ατόμων αποτελεί ορισμένες φορές το επιθυμητό αποτέλεσμα, σύμφωνα με τις αρχές δεοντολογίας (για παράδειγμα στην ανίχνευση της απάτης, της νομιμοποίησης εσόδων από παράνομες δραστηριότητες ή της χρηματοδότησης της τρομοκρατίας). Ωστόσο, η αυτόματη ταυτοποίηση εγείρει έντονους προβληματισμούς τόσο νομικού όσο και δεοντολογικού χαρακτήρα, καθώς μπορεί να έχει απροσδόκητες επιπτώσεις σε πολλά ψυχολογικά και κοινωνικοπολιτισμικά επίπεδα. Απαιτείται αναλογική χρήση τεχνικών ελέγχου στην ΤΝ για την υποστήριξη της αυτονομίας των Ευρωπαίων πολιτών. Ο σαφής καθορισμός του ενδεχομένου, των περιπτώσεων και των τρόπων χρήσης της ΤΝ για την αυτοματοποιημένη ταυτοποίηση ατόμων και η διαφοροποίηση μεταξύ της ταυτοποίησης ενός ατόμου και της ανίχνευσης και της παρακολούθησης ενός ατόμου, καθώς και μεταξύ της στοχευμένης επιτήρησης και της μαζικής επιτήρησης, θα είναι καθοριστικής σημασίας για την επίτευξη αξιόπιστης ΤΝ. Θα πρέπει να παρέχονται σαφείς εγγυήσεις για την εφαρμογή τεχνολογιών αυτού του είδους στην ισχύουσα νομοθεσία<sup>71</sup>. Όταν η νομική βάση μιας δραστηριότητας αυτού του είδους είναι η «συναίνεση», θα πρέπει να αναπτυχθούν πρακτικά μέσα<sup>72</sup> που καθιστούν δυνατή την έγκυρη και επαληθευμένη συναίνεση για την αυτόματη ταυτοποίηση από τεχνολογία ΤΝ ή από αντίστοιχες τεχνολογίες. Αυτό ισχύει επίσης για τη χρήση «ανώνυμων» προσωπικών δεδομένων που μπορούν να αποκτήσουν ξανά προσωπικό χαρακτήρα.

b. Συγκαλυμμένα συστήματα ΤΝ

(131) Οι άνθρωποι θα πρέπει να γνωρίζουν πάντα αν αλληλεπιδρούν άμεσα με άλλο ανθρώπινο ον ή μηχανή και αποτελεί ευθύνη των επαγγελματιών του τομέα της ΤΝ να το επιτυγχάνουν με αξιοπιστία. Ως εκ τούτου, οι επαγγελματίες του τομέα της ΤΝ θα πρέπει να διασφαλίζουν ότι οι άνθρωποι ενημερώνονται ή μπορούν να ζητούν και να επαληθεύουν το γεγονός ότι αλληλεπιδρούν με ένα σύστημα ΤΝ (για παράδειγμα, εκδίδοντας σαφείς και διαφανείς αποποιήσεις ευθυνών). Θα πρέπει να υπογραμμιστεί ότι υπάρχουν οριακές περιπτώσεις που περιπλέκουν το ζήτημα (π.χ. μια ανθρώπινη φωνή που φιλτράρεται από τεχνολογία ΤΝ). Θα πρέπει να ληφθεί υπόψη ότι η σύγχυση μεταξύ ανθρώπων και μηχανών θα μπορούσε να έχει πολλαπλές συνέπειες όπως οι συναισθηματικοί δεσμοί, η άσκηση επιρροής ή η μείωση της αξίας της ανθρώπινης υπόστασης.<sup>73</sup> Επομένως, η ανάπτυξη ανθρωποειδών ρομπότ<sup>74</sup> θα πρέπει να υποβάλλεται σε προσεκτική δεοντολογική αξιολόγηση.

c. Χαρακτηρισμός πολιτών βάσει ΤΝ κατά παράβαση των θεμελιωδών δικαιωμάτων

(132) Οι κοινωνίες θα πρέπει να μεριμνούν για την προστασία της ελευθερίας και της αυτονομίας όλων των πολιτών. Οποιαδήποτε μορφή χαρακτηρισμού των πολιτών μπορεί να οδηγήσει στην απώλεια της εν λόγω αυτονομίας και να θέσει σε κίνδυνο την αρχή της απαγόρευσης των διακρίσεων. Ο χαρακτηρισμός θα πρέπει να χρησιμοποιείται μόνο εάν υπάρχει σαφής αιτιολόγηση και όταν τα μέτρα είναι αναλογικά και δίκαια. Ο κανονιστικός χαρακτηρισμός των πολιτών (γενική αξιολόγηση της «ηθικής προσωπικότητας» ή της «δεοντολογικής ακεραιότητας») σε όλες τις πτυχές και σε μεγάλη κλίμακα από δημόσιες αρχές ή ιδιωτικούς φορείς θέτει σε κίνδυνο τις εν λόγω αξίες, ιδίως όταν δεν διενεργείται σύμφωνα με τα θεμελιώδη δικαιώματα και όταν χρησιμοποιείται δυσανάλογα και χωρίς έναν οριοθετημένο και κοινοποιημένο νόμιμο

<sup>71</sup> Εν προκειμένω, μπορεί να αναφερθεί το άρθρο 6 του ΓΚΠΔ, το οποίο προβλέπει, μεταξύ άλλων, ότι η επεξεργασία των δεδομένων είναι σύνομη μόνον εάν διαθέτει έγκυρη νομική βάση.

<sup>72</sup> Όπως προκύπτει από τους υφιστάμενους μηχανισμούς για την παροχή συναίνεσης μετά από ενημέρωση στο διαδίκτυο, οι καταναλωτές συνήθως παρέχουν συναίνεση χωρίς ουσιαστικά να εξετάζουν το ζήτημα. Ως εκ τούτου, οι μηχανισμοί δύσκολα μπορούν να θεωρηθούν πρακτικοί.

<sup>73</sup> Madary & Metzinger (2016). Real Virtuality: A Code of Ethical Conduct. Recommendations for Good Scientific Practice and the Consumers of VR-Technology. Frontiers in Robotics and AI, 3(3).

<sup>74</sup> Το ίδιο ισχύει και για τα άβαταρ που βασίζονται στην ΤΝ.

σκοπό.

(133) Επί του παρόντος, ο χαρακτηρισμός των πολιτών —σε μεγάλη ή μικρότερη κλίμακα— χρησιμοποιείται ήδη συχνά αμιγώς περιγραφικά και σε συγκεκριμένους τομείς (π.χ. σχολικά συστήματα, ηλεκτρονική μάθηση και άδειες οδήγησης). Ακόμη και σε αυτές τις πιο περιορισμένες εφαρμογές, θα πρέπει να τίθεται στη διάθεση των πολιτών μια πλήρως διαφανής διαδικασία, συμπεριλαμβανομένης της πληροφόρησης σχετικά με τη διαδικασία, τον σκοπό και τη μεθοδολογία του χαρακτηρισμού. Επισημαίνεται ότι η διαφάνεια δεν μπορεί να αποτρέψει τις διακρίσεις ή να εξασφαλίσει τη δικαιοσύνη και δεν αποτελεί πανάκεια για το πρόβλημα του χαρακτηρισμού. Ιδανικά, θα πρέπει να παρέχεται η δυνατότητα εξαίρεσης από τον μηχανισμό χαρακτηρισμού όταν είναι δυνατό χωρίς αρνητικές συνέπειες —διαφορετικά θα πρέπει να παρέχονται μηχανισμοί για την αμφισβήτηση και την αποκατάσταση των χαρακτηρισμών. Αυτό είναι ιδιαίτερα σημαντικό σε περιπτώσεις όπου υπάρχει ασυμμετρία εξουσίας μεταξύ των μερών. Οι συγκεκριμένες επιλογές εξαίρεσης θα πρέπει να διασφαλίζονται κατά τον σχεδιασμό της τεχνολογίας σε περιπτώσεις όπου αυτό είναι απαραίτητο για τη διασφάλιση της συμμόρφωσης με τα θεμελιώδη δικαιώματα και είναι απαραίτητες σε μια δημοκρατική κοινωνία.

d. Φονικά αυτόνομα οπτικά συστήματα (LAWS)

(134) Επί του παρόντος, ένας άγνωστος αριθμός χωρών και βιομηχανιών ερευνά και αναπτύσσει φονικά αυτόνομα οπτικά συστήματα, από πυραύλους με δυνατότητα επιλεκτικής στόχευσης έως μηχανές μάθησης με γνωστικές ικανότητες που θα είναι σε θέση να αποφασίζουν σε ποιον, πότε και πού θα επιτεθούν χωρίς ανθρώπινη παρέμβαση. Το γεγονός αυτό εγείρει σοβαρότατους δεοντολογικούς προβληματισμούς, όπως το γεγονός ότι θα μπορούσε να οδηγήσει σε έναν άνευ ιστορικού προηγούμενου ανεξέλεγκτο εξοπλιστικό ανταγωνισμό και να δημιουργήσει στρατιωτικά περιβάλλοντα στα οποία ο ανθρώπινος έλεγχος απουσιάζει σχεδόν πλήρως και δεν αντιμετωπίζονται οι κίνδυνοι δυσλειτουργίας. Το Ευρωπαϊκό Κοινοβούλιο ζήτησε την επείγουσα ανάπτυξη μιας κοινής, νομικώς δεσμευτικής θέσης, η οποία θα αφορά τα δεοντολογικά και νομικά ζητήματα ανθρώπινου ελέγχου, εποπτείας, λογοδοσίας και εφαρμογής του διεθνούς δικαίου των ανθρωπίνων δικαιωμάτων, του διεθνούς ανθρωπιστικού δικαίου και των στρατιωτικών στρατηγικών.<sup>75</sup> Ενθυμούμενοι τον στόχο της Ευρωπαϊκής Ένωσης για την προαγωγή της ειρήνης, όπως κατοχυρώνεται στο άρθρο 3 της Συνθήκης για την Ευρωπαϊκή Ένωση, ενστερνιζόμαστε και θα επιδιώξουμε να στηρίξουμε το ψήφισμα του Κοινοβουλίου της 12ης Σεπτεμβρίου 2018 και όλες τις σχετικές προσπάθειες που καταβάλλονται σε σχέση με τα φονικά αυτόνομα οπτικά συστήματα (LAWS).

e. Πιθανοί προβληματισμοί σε πιο μακροπρόθεσμο επίπεδο

(135) Η ανάπτυξη της ΤΝ εξακολουθεί να σχετίζεται με συγκεκριμένους τομείς και απαιτούνται καλά εκπαιδευμένοι άνθρωποι επιστήμονες και μηχανικοί για τον ακριβή καθορισμό των στόχων της. Ωστόσο, με μια παρεκβολή στο μέλλον και δη σε μεγαλύτερο χρονικό ορίζοντα, μπορούν να γίνουν υποθέσεις για ορισμένους καίριους προβληματισμούς σε μακροπρόθεσμο επίπεδο.<sup>76</sup> Στο πλαίσιο μιας προσέγγισης που βασιίζεται στους κινδύνους, οι εν λόγω προβληματισμοί θα πρέπει να λαμβάνονται υπόψη δεδομένου ότι «δεν γνωρίζουμε τι δεν γνωρίζουμε» και με βάση τη θεωρία του «μαύρου κύκνου».<sup>77</sup> Οι σοβαρές επιπτώσεις των συγκεκριμένων προβληματισμών, σε συνδυασμό με την τρέχουσα αβεβαιότητα ως προς αντίστοιχες εξελίξεις, καθιστά αναγκαία την τακτική αξιολόγηση των εν λόγω ζητημάτων.

## **Δ. ΣΥΜΠΕΡΑΣΜΑ**

<sup>75</sup> Ψήφισμα 2018/2752 (RSP) του Ευρωπαϊκού Κοινοβουλίου.

<sup>76</sup> Παρότι ορισμένοι θεωρούν ότι η γενική τεχνητή νοημοσύνη, η τεχνητή συνείδηση, οι τεχνητοί ηθικοί πράκτορες, η υπερευφυΐα ή η μετασημασιολογική ΤΝ μπορεί να αποτελούν παραδείγματα τέτοιων μακροπρόθεσμων προβληματισμών (που επί του παρόντος δεν υφίστανται), άλλοι τους θεωρούν εκτός πραγματικότητας.

<sup>77</sup> Τα γεγονότα τύπου «μαύρου κύκνου» είναι γεγονότα πολύ σπάνια, αλλά με πολύ μεγάλες επιπτώσεις —τόσο σπάνια που μπορεί έχουν περάσει απαρατήρητα. Ως εκ τούτου, η πιθανότητα να συμβούν μπορεί κατά κανόνα μόνο να εκτιμηθεί με πολύ υψηλό βαθμό αβεβαιότητας.

- (136) Το παρόν έγγραφο αποτελεί τις κατευθυντήριες γραμμές δεοντολογίας στον τομέα της ΤΝ τις οποίες εκπόνησε η ομάδα εμπειρογνομόνων υψηλού επιπέδου για την τεχνητή νοημοσύνη (OEYE για την ΤΝ).
- (137) Αναγνωρίζουμε τις θετικές επιπτώσεις που τα συστήματα ΤΝ έχουν ήδη και θα εξακολουθήσουν να έχουν, τόσο από εμπορικής όσο και από κοινωνιακής άποψης. Ωστόσο, μας ενδιαφέρει εξίσου να διασφαλίζεται ότι οι κίνδυνοι και άλλες δυσμενείς επιπτώσεις με τις οποίες σχετίζονται οι τεχνολογίες αυτές αντιμετωπίζονται με κατάλληλο και αναλογικό τρόπο με βάση την εκάστοτε εφαρμογή ΤΝ. Η ΤΝ είναι μια τεχνολογία που είναι και μετασχηματιστική και ανατρεπτική, και η εξέλιξή της τα τελευταία χρόνια διευκολύνεται από τη διαθεσιμότητα τεράστιων ποσοτήτων ψηφιακών δεδομένων, τις μεγάλες τεχνολογικές προόδους στον τομέα της υπολογιστικής ισχύος και της αποθηκευτικής ικανότητας, καθώς και τις σημαντικές επιστημονικές και τεχνικές καινοτομίες όσον αφορά τις μεθόδους και τα εργαλεία της ΤΝ. Τα συστήματα ΤΝ θα συνεχίσουν να επηρεάζουν την κοινωνία και τους πολίτες με τρόπους που δεν μπορούμε ακόμα να φανταστούμε.
- (138) Από την άποψη αυτή, είναι σημαντικό να κατασκευάζονται συστήματα ΤΝ που είναι άξια εμπιστοσύνης, διότι οι άνθρωποι θα μπορέσουν να αποκομίσουν πλήρως και με σιγουριά τα οφέλη της ΤΝ μόνο αν η τεχνολογία, συμπεριλαμβανομένων των διεργασιών και των ανθρώπων πίσω από την τεχνολογία, είναι αξιόπιστη. Ως εκ τούτου, κατά την εκπόνηση αυτών των κατευθυντήριων γραμμών, η αξιόπιστη ΤΝ, αποτέλεσε τη βασική μας φιλοδοξία.
- (139) Η αξιόπιστη ΤΝ έχει τρεις συνιστώσες: 1) θα πρέπει να είναι σύννομη, ήτοι να διασφαλίζει την τήρηση όλων εφαρμοστέων νομοθετικών και κανονιστικών διατάξεων, 2) θα πρέπει να είναι δεοντολογική, ήτοι να διασφαλίζει τη συμμόρφωση με δεοντολογικές αρχές και αξίες και 3) θα πρέπει να είναι στιβαρή, τόσο από τεχνικής όσο και από κοινωνικής άποψης, ώστε να διασφαλίζεται ότι, ακόμη κι όταν υπάρχει καλή πρόθεση, τα συστήματα ΤΝ δεν προκαλούν ακούσια βλάβη. Κάθε συνιστώσα αποτελεί αναγκαία αλλά όχι ικανή συνθήκη για την επίτευξη αξιόπιστης ΤΝ. Υπό ιδανικές συνθήκες, και οι τρεις συνιστώσες συνεργάζονται αρμονικά και αλληλεπικαλύπτονται στη λειτουργία τους. Αν προκύψουν τριβές μεταξύ τους, θα πρέπει να επιχειρούμε να τις ευθυγραμμίζουμε.
- (140) Στο κεφάλαιο I διατυπώσαμε τα θεμελιώδη δικαιώματα και μια αντίστοιχη δέσμη δεοντολογικών αρχών που είναι κρίσιμης σημασίας στο εκάστοτε πλαίσιο της ΤΝ. Στο κεφάλαιο II παραθέσαμε επτά βασικές απαιτήσεις που θα πρέπει να τηρούν τα συστήματα ΤΝ για την πραγμάτωση της αξιόπιστης ΤΝ. Προτείναμε τεχνικές και μη τεχνικές μεθόδους που μπορεί να συμβάλουν στην εφαρμογή τους. Τέλος, στο κεφάλαιο III παρουσιάσαμε έναν κατάλογο αξιολόγησης της αξιόπιστης ΤΝ, ο οποίος μπορεί να συμβάλει στην επιχειρησιακή υλοποίηση των επτά απαιτήσεων. Στην τελευταία ενότητα του εγγράφου παρουσιάσαμε παραδείγματα χρήσιμων ευκαιριών και καίριων προβληματισμών που εγείρουν τα συστήματα ΤΝ, επί των οποίων ελπίζουμε να δώσουμε το έναυσμα για περαιτέρω συζήτηση.
- (141) Η Ευρώπη βρίσκεται σε μοναδική πλεονεκτική θέση χάρη στην έμφαση που δίνει στο να θέτει τον πολίτη στο επίκεντρο των προσπαθειών της. Η έμφαση αυτή είναι γραμμένη στο ίδιο το DNA της Ευρωπαϊκής Ένωσης μέσω των Συνθηκών επί των οποίων οικοδομήθηκε. Το παρόν έγγραφο εντάσσεται στο όραμα της προαγωγής της αξιόπιστης ΤΝ, η οποία θεωρούμε ότι θα πρέπει να αποτελέσει τη βάση επί της οποίας η Ευρώπη θα μπορέσει να εξασφαλίσει ηγετική θέση στα καινοτόμα, προηγμένα συστήματα ΤΝ. Αυτό το φιλόδοξο όραμα θα συμβάλει στην εξασφάλιση της ευημερίας των Ευρωπαίων πολιτών, τόσο σε ατομικό όσο και σε συλλογικό επίπεδο. Στόχος μας είναι η καλλιέργεια μιας νοοτροπίας «αξιόπιστης ΤΝ για την Ευρώπη» στο πλαίσιο της οποίας όλοι θα μπορούν να αποκομίζουν τα οφέλη της ΤΝ με τρόπο που διασφαλίζει τον σεβασμό των θεμελιωδών αξιών μας: των θεμελιωδών δικαιωμάτων, της δημοκρατίας και του κράτους δικαίου.

## **ΓΛΩΣΣΑΡΙΟ**

(142) Το γλωσσάριο αυτό σχετίζεται με τις κατευθυντήριες γραμμές και σκοπός του είναι να συμβάλει στην κατανόηση των όρων που χρησιμοποιούνται στο παρόν έγγραφο.

### **Τεχνητή νοημοσύνη ή συστήματα TN**

(143) Τα συστήματα τεχνητής νοημοσύνης (TN) είναι συστήματα λογισμικού (ή ενδεχομένως και υλισμικού) που σχεδιάζονται από ανθρώπους<sup>78</sup> και, βάσει ενός δεδομένου σύνθετου στόχου, ενεργούν στην υλική ή ψηφιακή διάσταση με το να αντιλαμβάνονται το περιβάλλον τους μέσω της απόκτησης δεδομένων, να ερμηνεύουν τα δομημένα ή αδόμητα δεδομένα που έχουν συλλεχθεί, να προβαίνουν σε συλλογισμούς με βάση τις γνώσεις ή να επεξεργάζονται τις πληροφορίες, που εξάγονται από αυτά τα δεδομένα, και να αποφασίζουν ποια είναι η βέλτιστη ενέργεια (ή οι βέλτιστες ενέργειες) που θα πρέπει να εκτελέσουν για να επιτύχουν τον δεδομένο στόχο. Τα συστήματα TN μπορεί είτε να χρησιμοποιούν συμβολικούς κανόνες είτε να μαθαίνουν ένα αριθμητικό μοντέλο, και μπορεί επίσης να προσαρμόζουν τη συμπεριφορά τους με το να αναλύουν πώς επηρεάζεται το περιβάλλον από τις προηγούμενες ενέργειές τους.

(144) Ως επιστημονικό πεδίο, η TN περιλαμβάνει διάφορες προσεγγίσεις και τεχνικές, όπως η μηχανική μάθηση (συγκεκριμένα παραδείγματα της οποίας είναι η βαθιά μάθηση και η ενισχυτική μάθηση), η μηχανική συλλογιστική (που περιλαμβάνει τον σχεδιασμό, τον προγραμματισμό, την αναπαράσταση και τη συλλογιστική γνώσης, την αναζήτηση και τη βελτιστοποίηση) και η ρομποτική (που περιλαμβάνει έλεγχο, αντίληψη, αισθητήρες και ενεργοποιητές, καθώς και την ενσωμάτωση όλων των άλλων τεχνικών σε κυβερνο-υλικά συστήματα).

(145) Δημοσιεύεται παράλληλα χωριστό έγγραφο που εκπόνησε η ΟΕΥΕ για την TN και στο οποίο εξειδικεύεται περαιτέρω ο ορισμός των *συστημάτων TN* που χρησιμοποιείται για τους σκοπούς του παρόντος εγγράφου, με τίτλο «Ορισμός της TN: κύριες δυνατότητες και επιστημονικοί κλάδοι».

### **Επαγγελματίες του τομέα της TN**

(146) Ως επαγγελματίες του τομέα της TN εννοούμε όλα τα άτομα ή τους οργανισμούς που αναπτύσσουν (συμπεριλαμβανομένης της έρευνας, του σχεδιασμού ή της παροχής δεδομένων), εγκαθιστούν (συμπεριλαμβανομένης της υλοποίησης) ή χρησιμοποιούν συστήματα TN, εξαιρουμένων όσων χρησιμοποιούν συστήματα TN υπό την ιδιότητα του τελικού χρήστη ή καταναλωτή.

### **Κύκλος ζωής συστήματος TN**

(147) Ο κύκλος ζωής ενός συστήματος TN περιλαμβάνει το στάδιο της ανάπτυξης (συμπεριλαμβανομένης της έρευνας, του σχεδιασμού, της παροχής δεδομένων και περιορισμένων δοκιμών), της εγκατάστασης (συμπεριλαμβανομένης της υλοποίησης) και της χρήσης του.

### **Ελεγχιμότητα**

(148) Η ελεγχιμότητα αναφέρεται στη δυνατότητα ενός συστήματος TN να υποβάλλεται σε αξιολόγηση ως προς τους αλγορίθμους, τα δεδομένα και τις διεργασίες σχεδιασμού του συστήματος. Αποτελεί μία από τις επτά απαιτήσεις που θα πρέπει να πληροί ένα αξιόπιστο σύστημα TN. Κάτι τέτοιο δεν σημαίνει απαραίτητα ότι οι πληροφορίες σχετικά με τα επιχειρηματικά μοντέλα και την πνευματική ιδιοκτησία που σχετίζονται με το σύστημα TN θα πρέπει πάντοτε να είναι ευρέως διαθέσιμες. Η διασφάλιση ιχνηλασιμότητας και μηχανισμών καταγραφής από τα αρχικά στάδια σχεδιασμού του συστήματος TN μπορεί να συμβάλει στο να καθίσταται δυνατή η ελεγχιμότητα του συστήματος.

### **Μεροληψία**

---

<sup>78</sup> Οι άνθρωποι σχεδιάζουν συστήματα TN απευθείας, αλλά μπορεί επίσης να χρησιμοποιούν τεχνικές TN για να βελτιστοποιήσουν τον σχεδιασμό τους.

(149) Μεροληψία είναι η τάση προκατάληψης απέναντι ή ενάντια σε ένα πρόσωπο, σε ένα αντικείμενο ή σε μια άποψη. Μεροληψία μπορεί να προκύψει με πολλούς τρόπους στα συστήματα ΤΝ. Για παράδειγμα, σε συστήματα ΤΝ που χρησιμοποιούν δεδομένα, όπως αυτά που παράγονται μέσω της μηχανικής μάθησης, η ύπαρξη σφαλμάτων μεροληψίας στη συλλογή δεδομένων και την εκπαίδευση μπορεί να έχει ως αποτέλεσμα το σύστημα ΤΝ να επιδεικνύει μεροληψία. Στην ΤΝ που βασίζεται στη λογική, όπως π.χ. σε συστήματα που βασίζονται σε κανόνες, μεροληψία μπορεί να προκύψει λόγω του τρόπου με τον οποίο αντιλαμβάνεται ένας μηχανικός γνώσης τους κανόνες που ισχύουν σε ένα συγκεκριμένο περιβάλλον. Μεροληψία μπορεί επίσης να προκύψει λόγω της ηλεκτρονικής μάθησης και της προσαρμογής μέσω της αλληλεπίδρασης. Μπορεί επίσης να προκύψει μέσω της εξατομίκευσης, σε περιπτώσεις που στους χρήστες παρέχονται προτάσεις ή ροές πληροφοριών που είναι προσαρμοσμένες στις προτιμήσεις του χρήστη. Δεν είναι απαραίτητο να σχετίζεται με ανθρώπινη μεροληψία ή με τη συλλογή δεδομένων που γίνεται από ανθρώπους. Μπορεί να προκύψει, για παράδειγμα, μέσω των περιορισμένων πλαισίων εντός των οποίων χρησιμοποιείται ένα σύστημα, περίπτωση κατά την οποία δεν υπάρχει η δυνατότητα να γενικευτεί και σε άλλα πλαίσια. Η μεροληψία μπορεί να είναι καλή ή κακή, εκούσια ή ακούσια. Σε ορισμένες περιπτώσεις, η μεροληψία μπορεί να οδηγήσει σε αποτελέσματα που εισάγουν διακρίσεις και/ή είναι αθέμιτα. Οι περιπτώσεις αυτές αναφέρονται στο παρόν έγγραφο ως αθέμιτη αμεροληψία.

### **Δεοντολογία/Ηθική**

(150) Η ηθική είναι ακαδημαϊκός κλάδος, υποπεδίο της φιλοσοφίας. Σε γενικές γραμμές ασχολείται με ερωτήματα όπως «ποιες πράξεις είναι ορθές;», «ποια είναι η αξία της ανθρώπινης ζωής;», «τι είναι δικαιοσύνη;» ή «τι σημαίνει καλή ζωή;». Η ακαδημαϊκή ηθική περιλαμβάνει τέσσερα κύρια ερευνητικά πεδία: i) την μεταηθική, που αφορά κυρίως τι έννοια έχει και σε τι παραπέμπει η κανονιστική πρόταση, καθώς και το ζήτημα του πώς μπορούν να προσδιοριστούν οι σχετικές τιμές αλήθειας (εφόσον υπάρχουν), ii) την κανονιστική ηθική, που αποτελεί το πρακτικό μέσο για τον καθορισμό του ηθικού πράττειν μέσω της εξέτασης των προτύπων για το ορθό και το εσφαλμένο και της απόδοσης μιας αξίας σε συγκεκριμένες πράξεις, iii) την περιγραφική ηθική, στόχος της οποίας είναι η εμπειρική διερεύνηση της ηθικής συμπεριφοράς και των πεποιθήσεων των ανθρώπων, και iv) την εφαρμοσμένη ηθική, που αφορά το τι είμαστε υποχρεωμένοι (ή μας επιτρέπεται) να κάνουμε σε μια συγκεκριμένη (συχνά καινοφανή) κατάσταση ή σε ένα συγκεκριμένο πεδίο (συχνά χωρίς ιστορικό προηγούμενο) πιθανών τρόπων ανάληψης δράσης. Η εφαρμοσμένη ηθική ασχολείται με πραγματικές καταστάσεις, στις οποίες καλούμαστε να λαμβάνουμε αποφάσεις υπό την πίεση του χρόνου και, συχνά, με περιορισμένη ορθολογικότητα. Η δεοντολογία της ΤΝ θεωρείται γενικά παράδειγμα εφαρμοσμένης ηθικής, και εστιάζει στα κανονιστικά ζητήματα που εγείρει η ανάπτυξη, η υλοποίηση και η χρήση της ΤΝ.

(151) Κατά την εξέταση ζητημάτων δεοντολογίας χρησιμοποιούνται συχνά οι όροι «ηθικός» και «δεοντολογικός». Ο όρος «ηθικός» αναφέρεται σε συγκεκριμένα, πραγματολογικά πρότυπα συμπεριφοράς, στα ήθη και στις συμβάσεις που απαντώνται σε συγκεκριμένους πολιτισμούς, ομάδες ή άτομα σε δεδομένη χρονική στιγμή. Ο όρος «δεοντολογικός» αναφέρεται στην αξιολογική εξέταση των εν λόγω συγκεκριμένων ενεργειών και συμπεριφορών από συστηματική, ακαδημαϊκή σκοπιά.

### **Δεοντολογική ΤΝ**

(152) Στο παρόν έγγραφο, ο όρος δεοντολογική ΤΝ αναφέρεται στην ανάπτυξη, εγκατάσταση και χρήση ΤΝ που διασφαλίζει τη συμμόρφωση με τους δεοντολογικούς κανόνες, συμπεριλαμβανομένων των θεμελιωδών δικαιωμάτων ως ειδικών ηθικών δικαιωμάτων, τις δεοντολογικές αρχές και τις σχετικές βασικές αξίες. Είναι η δεύτερη από τις τρεις συνιστώσες που είναι αναγκαίες για την επίτευξη αξιόπιστης ΤΝ.

### **Ανθρωποκεντρική ΤΝ**

(153) Η ανθρωποκεντρική προσέγγιση της ΤΝ επιδιώκει να διασφαλίσει ότι οι ανθρώπινες αξίες έχουν κεντρικό ρόλο στον τρόπο με τον οποίο αναπτύσσονται, εγκαθίστανται, χρησιμοποιούνται και παρακολουθούνται τα συστήματα ΤΝ, μέσω της εξασφάλισης σεβασμού για τα θεμελιώδη δικαιώματα, συμπεριλαμβανομένων

όσων κατοχυρώνονται στις Συνθήκες της Ευρωπαϊκής Ένωσης και στον Χάρτη των Θεμελιωδών Δικαιωμάτων της ΕΕ, τα οποία ενώνει το γεγονός ότι αναφέρονται σε μια κοινή βάση εδραιωμένη στον σεβασμό της ανθρώπινης αξιοπρέπειας, που θέλει τον άνθρωπο να απολαμβάνει μια μοναδική και αναπαλλοτρίωτη ηθική υπόσταση. Αυτό συνεπάγεται επίσης ότι λαμβάνεται υπόψη το φυσικό περιβάλλον και τα υπόλοιπα έμβια όντα που αποτελούν μέρος του ανθρώπινου οικοσυστήματος, καθώς και μια βιώσιμη προσέγγιση που καθιστά δυνατή την ευημερία των μελλοντικών γενεών.

### **Δοκιμές τύπου «Red Teaming»**

(154) «Red Teaming» ονομάζεται η πρακτική στο πλαίσιο της οποίας μια ανεξάρτητη ομάδα (red team) «προκαλεί» έναν οργανισμό να βελτιώσει την αποτελεσματικότητά του αναλαμβάνοντας τον ρόλο ή την οπτική γωνία μιας αντίπαλης πλευράς. Χρησιμοποιείται ιδίως για τον εντοπισμό και την αντιμετώπιση πιθανών ευπαθειών ασφάλειας.

### **Αναπαραγωγικότητα**

(155) Η αναπαραγωγικότητα περιγράφει κατά πόσον ένα πείραμα TN παρουσιάζει την ίδια συμπεριφορά όταν επαναλαμβάνεται υπό τις ίδιες συνθήκες.

### **Στιβαρή TN**

(156) Η στιβαρότητα ενός συστήματος TN περιλαμβάνει τόσο την τεχνική στιβαρότητά του (καταλληλότητα για το δεδομένο πλαίσιο, όπως ο τομέας εφαρμογής ή η φάση του κύκλου ζωής), καθώς και τη στιβαρότητά του από κοινωνική άποψη (μέριμνα ώστε το σύστημα TN να λαμβάνει δεόντως υπόψη το πλαίσιο και το περιβάλλον στο οποίο λειτουργεί το σύστημα). Αυτό είναι κρίσιμης σημασίας ώστε να διασφαλίζεται ότι, ακόμη κι όταν υπάρχει καλή πρόθεση, δεν προκαλείται ακούσια βλάβη. Η στιβαρότητα είναι η τρίτη από τις τρεις συνιστώσες που είναι αναγκαίες για την επίτευξη αξιόπιστης TN.

### **Ενδιαφερόμενα μέρη**

(157) Ως ενδιαφερόμενα μέρη εννοούμε όσους ερευνούν, αναπτύσσουν, σχεδιάζουν, εγκαθιστούν ή χρησιμοποιούν TN, καθώς και όσους επηρεάζονται (άμεσα ή έμμεσα) από την TN —συμπεριλαμβανομένων, ενδεικτικά, εταιρειών, οργανισμών, ερευνητών, δημόσιων υπηρεσιών, ιδρυμάτων, οργανώσεων της κοινωνίας των πολιτών, κυβερνήσεων, ρυθμιστικών αρχών, κοινωνικών εταίρων, ατόμων, πολιτών, εργαζομένων και καταναλωτών.

### **Ιχνηλασιμότητα**

(158) Η ιχνηλασιμότητα ενός συστήματος TN αναφέρεται στην ικανότητα παρακολούθησης των διεργασιών δεδομένων, ανάπτυξης και εγκατάστασης του συστήματος, κατά κανόνα μέσω καταγεγραμμένου αναγνωριστικού.

### **Εμπιστοσύνη**

(159) Στη βιβλιογραφία παρατίθεται ο ακόλουθος ορισμός: «Ως εμπιστοσύνη νοείται: 1) ένα σύνολο συγκεκριμένων πεποιθήσεων που σχετίζονται με την καλή προαίρεση, την ικανότητα, την ακεραιότητα και την προβλεψιμότητα (πεποιθήσεις εμπιστοσύνης), 2) η προθυμία κάποιου να εξαρτηθεί από κάποιον άλλο σε μια κατάσταση που ενέχει κινδύνους (πρόθεση εμπιστοσύνης), ή 3) ο συνδυασμός των παραπάνω.»<sup>79</sup> Παρότι η «εμπιστοσύνη» δεν είναι συνήθως μια ιδιότητα που αποδίδεται σε μηχανές, το παρόν έγγραφο επιδιώκει να τονίσει πόσο σημαντικό είναι να μπορεί να υπάρχει εμπιστοσύνη όχι μόνο ως προς το ότι τα συστήματα TN είναι σύννομα, δεοντολογικά και στιβαρά, αλλά και ως προς το ότι η εμπιστοσύνη αυτή μπορεί να αποδοθεί και σε όλους τους ανθρώπους και τις διεργασίες που εμπλέκονται στον κύκλο ζωής του συστήματος TN.

---

<sup>79</sup> Siau, K., Wang, W. (2018), Building Trust in Artificial Intelligence, Machine Learning, and Robotics, *CUTTER BUSINESS TECHNOLOGY JOURNAL* (31), S. 47–53.

## **Αξιόπιστη ΤΝ**

(160) Η αξιόπιστη ΤΝ έχει τρεις συνιστώσες: 1) θα πρέπει να είναι σύννομη, ήτοι να διασφαλίζει την τήρηση όλων των εφαρμοστέων νομοθετικών και κανονιστικών διατάξεων, 2) θα πρέπει να είναι δεοντολογική, ήτοι να διασφαλίζει τον σεβασμό και τη συμμόρφωση με δεοντολογικές αρχές και αξίες και 3) θα πρέπει να είναι στιβαρή, τόσο από τεχνικής όσο και από κοινωνικής άποψης, διότι, ακόμη κι όταν υπάρχει καλή πρόθεση, τα συστήματα ΤΝ μπορεί να προκαλέσουν ακούσια βλάβη. Η αξιόπιστη ΤΝ δεν αφορά μόνο την αξιοπιστία του ίδιου του συστήματος ΤΝ, αλλά περιλαμβάνει και την αξιοπιστία όλων των διεργασιών και των παραγόντων που αποτελούν μέρος του κύκλου ζωής του συστήματος.

## **Ευάλωτα πρόσωπα και ομάδες**

(161) Δεν υπάρχει ένας συμφωνημένος ή κοινά αποδεκτός νομικός ορισμός των ευάλωτων προσώπων, λόγω της ετερογένειας που τα χαρακτηρίζει. Τα χαρακτηριστικά που καθιστούν ευάλωτο ένα πρόσωπο ή μια ομάδα εξαρτώνται συχνά από το εκάστοτε πλαίσιο. Ρόλο μπορεί να παίζουν προσωρινά γεγονότα στη ζωή ενός ατόμου (όπως η παιδική ηλικία ή μια ασθένεια), παράγοντες της αγοράς (όπως η ασυμμετρία πληροφόρησης ή η ισχύς στην αγορά), οικονομικοί παράγοντες (όπως η φτώχεια), παράγοντες που συνδέονται με την ταυτότητα (όπως το φύλο, η θρησκεία ή το πολιτισμικό υπόβαθρο) ή άλλοι παράγοντες. Ο Χάρτης των Θεμελιωδών Δικαιωμάτων της ΕΕ περιλαμβάνει υπό το άρθρο 21 για την απαγόρευση των διακρίσεων τους ακόλουθους λόγους διακρίσεων, οι οποίοι, μεταξύ άλλων μπορούν να αποτελούν σημείο αναφοράς: διακρίσεις λόγω φύλου, φυλής, χρώματος, εθνοτικής καταγωγής ή κοινωνικής προέλευσης, γενετικών χαρακτηριστικών, γλώσσας, θρησκείας ή πεποιθήσεων, πολιτικών ή άλλων φρονημάτων, ιδιότητας μέλους εθνικής μειονότητας, περιουσίας, γέννησης, αναπηρίας, ηλικίας και γενετήσιου προσανατολισμού. Άλλες νομικές πράξεις διέπουν τα δικαιώματα συγκεκριμένων ομάδων, πέραν όσων παρατίθενται ανωτέρω. Ο κατάλογος αυτός δεν είναι εξαντλητικός και μπορεί να τροποποιείται κατά καιρούς. Ευάλωτη ομάδα είναι μια ομάδα προσώπων που είναι ευάλωτα λόγω ενός ή περισσότερων κοινών χαρακτηριστικών τους.



**Το παρόν έγγραφο εκπονήθηκε από τα μέλη της ομάδας εμπειρογνομόνων υψηλού επιπέδου για την TN**

τα οποία παρατίθενται ακολούθως κατ' αλφαβητική σειρά

Pekka Ala-Pietilä, πρόεδρος ΟΕΥΕ για την TN AI Finland, Huhtamaki, Sanoma	Pierre Lucas Orgalim – Europe's technology industries
Wilhelm Bauer Fraunhofer	Ieva Martinkenaite Telenor
Urs Bergmann – Συνεισηγητής Zalando	Thomas Metzinger – Συνεισηγητής JGU Mainz & European University Association
Mária Bielíková Slovak University of Technology της Μπρατισλάβα	Catelijne Muller ALLAI Netherlands & EOKE
Cecilia Bonefeld-Dahl – Συνεισηγήτρια DigitalEurope	Markus Noga SAP
Yann Bonnet ANSSI	Barry O'Sullivan, αντιπρόεδρος ΟΕΥΕ για την TN University College Cork
Loubna Bouarfa OKRA	Ursula Pacht Ευρωπαϊκό Γραφείο Ενώσεων Καταναλωτών (ΕΓΕΚ)
Stéphan Brunessaux Airbus	Nicolas Petit – Συνεισηγητής Πανεπιστήμιο της Λιέγης
Raja Chatila IEEE Initiative Ethics of Intelligent/Autonomous Systems & πανεπιστήμιο της Σορβόνης	Christoph Peylo Bosch
Mark Coeckelbergh Πανεπιστήμιο της Βιέννης	Iris Plöger BDI
Virginia Dignum – Συνεισηγήτρια Πανεπιστήμιο της Umeå	Stefano Quintarelli Garden Ventures
Luciano Floridi Πανεπιστήμιο της Οξφόρδης	Andrea Renda College of Europe Faculty & Κέντρο Μελετών Ευρωπαϊκής Πολιτικής (CEPS)
Jean-Francois Gagné – Συνεισηγητής Element AI	Francesca Rossi IBM
Chiara Giovannini Ευρωπαϊκή Ένωση για τον Συντονισμό της Εκπροσώπησης των Καταναλωτών στην Τυποποίηση (ANEC)	Cristina San José Ομοσπονδία Ευρωπαϊκών Τραπεζών (EBF)
Joanna Goodey Οργανισμός Θεμελιωδών Δικαιωμάτων	George Sharkov Digital SME Alliance
Sami Haddadin Σχολή ρομποτικής και νοημοσύνης των μηχανών του Μονάχου (MSRM)	Philipp Slusallek Γερμανικό Κέντρο Έρευνας για την TN (DFKI)
Gry Hasselbalch The thinkdotank DataEthics & πανεπιστήμιο της Κοπεγχάγης	Françoise Soulié Fogelman Σύμβουλος σε θέματα TN
Fredrik Heintz Πανεπιστήμιο του Λινσέπινγκ	Saskia Steinacker – Συνεισηγήτρια Bayer
Fanny Hidvegi Access Now	Jaan Tallinn Ambient Sound Investment
Eric Hilgendorf Πανεπιστήμιο του Würzburg	Thierry Tingaud STMicroelectronics
Klaus Höckner Hilfsgemeinschaft der Blinden und Sehschwachen	Jakob Uszkoreit Google
Mari-Noëlle Jégo-Laveissière Orange	Aimee Van Wynsberghe – Συνεισηγήτρια TU Delft
Leo Kärkkäinen Nokia Bell Labs	Thiébaud Weber Ευρωπαϊκή Συνομοσπονδία Συνδικαλιστικών Οργανώσεων (CES)
Sabine Theresia Köszegi TU Wien	Cecile Wendling AXA
Robert Kroplewski Δικηγόρος & σύμβουλος της πολωνικής κυβέρνησης	Karen Yeung – Συνεισηγήτρια Πανεπιστήμιο του Μπέρμιγχαμ
Elisabeth Ling RELX	

Οι Urs Bergmann, Cecilia Bonefeld-Dahl, Virginia Dignum, Jean-François Gagné, Thomas Metzinger, Nicolas Petit, Saskia Steinacker, Aimee Van Wynsberghe και Karen Yeung εργάστηκαν ως εισηγητές του παρόντος εγγράφου.

Ο Pekka Ala-Pietilä είναι πρόεδρος της ΟΕΥΕ για την ΤΝ. Ο Barry O'Sullivan είναι αντιπρόεδρος και συντονιστής του δεύτερου παραδοτέου της ΟΕΥΕ για την ΤΝ. Η Nozha Boujemaα, αντιπρόεδρος έως την 1η Φεβρουαρίου 2019 και συντονίστρια του πρώτου παραδοτέου, συνεισέφερε και στο περιεχόμενο του παρόντος εγγράφου.

Η Nathalie Smuha παρείχε υποστήριξη σε επίπεδο επιμέλειας.