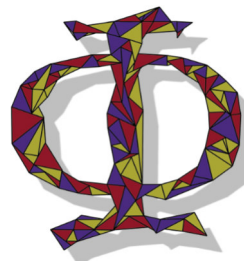




ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΕΘΝΙΚΟΝ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟΝ
ΠΑΝΕΠΙΣΤΗΜΙΟΝ ΑΘΗΝΩΝ

ΦΙΛΟΣΟΦΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΦΙΛΟΣΟΦΙΑΣ



ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΦΙΛΟΣΟΦΙΑ»
ΕΙΔΙΚΕΥΣΗ: ΕΦΑΡΜΟΣΜΕΝΗ ΗΘΙΚΗ

Μαρία Ι. Σπανού

ΘΕΜΑ ΕΡΓΑΣΙΑΣ

«Η εφαρμογή της Τεχνητής Νοημοσύνης στην Ψυχική Υγεία: Οι ηθικές διαστάσεις της χρήσης συστημάτων Τ.Ν. στις τηλεφωνικές γραμμές συμβουλευτικής υποστήριξης».

Διπλωματική Εργασία για την απόκτηση Διπλώματος Μεταπτυχιακών Σπουδών στην ειδίκευση: «Εφαρμοσμένη Ηθική».

Επιβλέπων:

Γεώργιος Κωστελέτος
Διδάκτωρ Φιλοσοφίας, Τμήμα Φιλοσοφίας ΕΚΠΑ

Μέλη Τριμελούς Εξεταστικής Επιτροπής:

Αλκιβιάδης Γούναρης Διδάκτωρ Φιλοσοφίας, Τμήμα Φιλοσοφίας ΕΚΠΑ	Γεώργιος Κωστελέτος Διδάκτωρ Φιλοσοφίας, Τμήμα Φιλοσοφίας ΕΚΠΑ	Ευάγγελος Πρωτοπαπαδάκης Διδάκτωρ Φιλοσοφίας, Τμήμα Φιλοσοφίας ΕΚΠΑ
--	--	---

ΑΘΗΝΑ, 2024

Εθνικόν και Καποδιστριακόν Πανεπιστήμιον Αθηνών, Φιλοσοφική Σχολή, Τμήμα
Φιλοσοφίας, Μαρία Σπανού © [2024] - Με την επιφύλαξη παντός δικαιώματος.

ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΕΘΝΙΚΟΝ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟΝ ΠΑΝΕΠΙΣΤΗΜΙΟΝ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΦΙΛΟΣΟΦΙΑΣ
ΦΙΛΟΣΟΦΙΚΗ ΣΧΟΛΗ

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ «ΦΙΛΟΣΟΦΙΑ»
ΕΙΔΙΚΕΥΣΗ: ΕΦΑΡΜΟΣΜΕΝΗ ΗΘΙΚΗ

Μαρία Ι. Σπανού

ΘΕΜΑ ΕΡΓΑΣΙΑΣ

«Η εφαρμογή της Τεχνητής Νοημοσύνης στην Ψυχική Υγεία: Οι ηθικές διαστάσεις της χρήσης των συστημάτων Τ.Ν. στις τηλεφωνικές γραμμές βοήθειας και συμβουλευτικής υποστήριξης».

Διπλωματική Εργασία για την απόκτηση Διπλώματος Μεταπτυχιακών Σπουδών στην ειδίκευση: «Εφαρμοσμένη Ηθική».

Επιβλέπων:

Γεώργιος Κωστελέτος
Διδάκτωρ Φιλοσοφίας, Τμήμα Φιλοσοφίας ΕΚΠΑ

Μέλη Τριμελούς Εξεταστικής Επιτροπής:

Αλκιβιάδης Γούναρης Διδάκτωρ Φιλοσοφίας, Τμήμα Φιλοσοφίας ΕΚΠΑ	Γεώργιος Κωστελέτος Διδάκτωρ Φιλοσοφίας, Τμήμα Φιλοσοφίας ΕΚΠΑ	Ευάγγελος Πρωτοπαπαδάκης Διδάκτωρ Φιλοσοφίας, Τμήμα Φιλοσοφίας ΕΚΠΑ
--	--	---

ΑΘΗΝΑ, 2024

Στον αγαπημένο μου G.

Ευχαριστίες

Η παρούσα διπλωματική εργασία πραγματοποιήθηκε στο πλαίσιο φοίτησης μου στο ΠΜΣ Φιλοσοφίας με κατεύθυνση την Εφαρμοσμένη Ηθική. Η ολοκλήρωσή της δε θα ήταν εφικτή χωρίς την πολύτιμη συμβολή του Διδάκτορα και υποψήφιου Μεταδιδάκτορα Φιλοσοφίας κυρίου Γεώργιου Κωστελέτου, ένας εκ των διδασκόντων του μαθήματος «Ηθική και Τ.Ν.» και ακολούθως, επιβλέπων κατά την εκπόνηση της εργασίας αυτής. Η συνεχής υποστήριξη και η ενθάρρυνση από πλευράς του, η έμπνευση, η καθοδήγησή προς νέους τρόπους σκέψης και η συμβολή του στην προσωπική και ακαδημαϊκή μου εξέλιξη, αποτελούν ελάχιστους μόνο από τους λόγους της απέραντης ευγνωμοσύνης μου.

Επιπλέον, το σεβασμό, την εκτίμησή και τις ευχαριστίες μου θα ήθελα να εκφράσω στον Διδάκτορα Φιλοσοφίας κύριο Αλκιβιάδη Γούναρη, επίσης διδάσκων του μαθήματος «Ηθική και Τ.Ν.», για τη μετάδοση της γνώσης, την εξαιρετική παρουσίαση του αντικειμένου της Τ.Ν. και την ξεχωριστή εμπειρία των μαθημάτων που, μαζί με τον κύριο Κωστελέτο, δημιούργησαν και προσέφεραν σε ένα περιβάλλον που διευρύνει τη σκέψη και ενθαρρύνει τον επαναπροδιορισμό.

Περιεχόμενα

Περίληψη	1
Εισαγωγή.....	2
Συντομογραφίες και ακρωνύμια.....	4
ΚΕΦΑΛΑΙΟ 1^ο: ΣΥΜΒΟΥΛΕΥΤΙΚΗ ΥΠΟΣΤΗΡΙΞΗ ΚΑΙ ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ.	5
1.α. Τεχνητή Νοημοσύνη: Ο ορισμός και η ιστορική εξέλιξή της μέχρι και σήμερα.	5
1.β. Η Συμβουλευτική υποστήριξη στην Ψυχική υγεία.....	9
1.γ. Η ενσωμάτωση της Τ.Ν. στο πεδίο της Συμβουλευτικής σήμερα	15
ΚΕΦΑΛΑΙΟ 2^ο: ΤΑ ΟΦΕΛΗ ΚΑΙ ΤΑ ΕΠΙΣΤΗΜΟΛΟΓΙΚΑ - ΜΕΘΟΔΟΛΟΓΙΚΑ ΠΡΟΒΛΗΜΑΤΑ ΤΗΣ ΕΜΠΛΟΚΗΣ ΤΗΣ Τ.Ν. ΣΤΗΝ ΣΥΜΒΟΥΛΕΥΤΙΚΗ ΚΑΙ ΣΤΗΝ ΨΥΧΙΚΗ ΥΓΕΙΑ.....	23
2.α. Τα οφέλη και τα επιστημολογικά - μεθοδικά προβλήματα βάσει των σημερινών δεδομένων.	23
2.β. Οι παράγοντες που επηρεάζουν την αξιολόγηση των αποτελεσμάτων	29
2.γ. Οι ενδεχόμενες μελλοντικές εξελίξεις στον τομέα της ψυχικής υγείας.....	34
ΚΕΦΑΛΑΙΟ 3^ο: Η ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΣΤΙΣ ΤΗΛΕΦΩΝΙΚΕΣ ΓΡΑΜΜΕΣ ΒΟΗΘΕΙΑΣ ΚΑΙ ΣΥΜΒΟΥΛΕΥΤΙΚΗΣ ΥΠΟΣΤΗΡΙΞΗΣ.	39
3.α. Οι τηλεφωνικές γραμμές βοήθειας και συμβουλευτικής υποστήριξης.	39
3.β. Οφέλη και προβληματισμοί γύρω από την ενσωμάτωση της Τ.Ν. στις τηλεφωνικές γραμμές βοήθειας και συμβουλευτική υποστήριξη.	43
3.γ. Η θεραπευτική σχέση	52
3.δ. Οι δεξιότητες του συμβούλου κατά τη συμβουλευτική διαδικασία.	59
3.ε. Περίπτωση μελέτης (case study).....	67
ΚΕΦΑΛΑΙΟ 4^ο: Η ΕΞΕΤΑΣΗ ΜΕΣΩ ΗΘΙΚΗΣ	75
4.α. Οι τέσσερις αρχές της Βιοηθικής.	75
4.α.1. Απόρρητο και ιδιωτικότητα.....	77
4.α.2. Έρευνα και εφαρμογή	86
4.α.3. Ασφάλεια του χρήστη.....	91
4.β. Αξιολόγηση υπό το πρίσμα του Ωφελιμισμού και της Δεοντοκρατίας.	99
ΕΠΙΛΟΓΟΣ – ΣΥΜΠΕΡΑΣΜΑΤΑ.....	111
ΒΙΒΛΙΟΓΡΑΦΙΑ	114

Περίληψη

Τα τελευταία χρόνια, η Τεχνητή Νοημοσύνη σταδιακά εδραιώνεται στον τομέα της ψυχικής υγείας, παρέχοντας ισχυρά εργαλεία που δημιουργούν νέες δυνατότητες για την πρόληψη, τη διάγνωση και τη θεραπεία των ψυχικών διαταραχών. Ως αποτέλεσμα, η είσοδός της στο κομμάτι συμβουλευτικής υποστήριξης συγκεκριμένα, είναι ολοένα και πιο δυναμική, με τα θεραπευτικά chatbot, τα οποία ήδη χρησιμοποιούνται από χιλιάδες χρήστες ανά τον κόσμο, να μεταμορφώνουν ριζικά την εμπειρία της συμβουλευτικής διαδικασίας.

Στην παρούσα διατριβή, θα συζητήσουμε τους ηθικούς προβληματισμούς που εγείρει η ριζική αυτή αλλαγή, εστιάζοντας συγκεκριμένα, στο ενδεχόμενο χρήσης των θεραπευτικών chatbot στις τηλεφωνικές γραμμές βοήθειας και συμβουλευτικής υποστήριξης, που μέχρι σήμερα στελεχώνονται από ανθρώπους επαγγελματίες. Ειδικότερα, θα εξετάσουμε τις ηθικές προκλήσεις που το ενδεχόμενο αυτό δημιουργεί, τόσο λόγω της επίδρασης του στο βαθμό ποιότητας των παρεχόμενων υπηρεσιών, όσο και λόγω του επαναπροσδιορισμού που μπορεί να επιφέρει στις ζωτικές έννοιες της θεραπευτικής αλληλεπίδρασης και σχέσης. Παράλληλα, θα μελετήσουμε το κατά πόσο η χρήση των συστημάτων της Τ.Ν. στην ψυχική υγεία είναι συμβατή με τις αρχές της Ηθικής, προβαίνοντας στην επισκόπηση του φαινομένου μέσα από τις τέσσερις αρχές της βιοηθικής και δύο εκ των κυριότερων ρευμάτων της ηθικής φιλοσοφίας, συγκεκριμένα του Ωφελιμισμού και της Δεοντοκρατίας.

Λέξεις κλειδιά: Τεχνητή Νοημοσύνη, Συνομιλητική Τ.Ν., chatbot, ψυχική υγεία, Συμβουλευτική, γραμμές βοήθειας, Ηθική της Τ.Ν., Βιοηθική.

Abstract

In recent years, Artificial Intelligence is gradually established in the field of mental health, providing powerful tools that create new possibilities for the prevention, the diagnosis and the treatment of mental disorders. As a result, its entry into the field of counselling support in particular, is increasingly dynamic, through therapeutic chatbots that, by being already used by thousands of users around the world, are radically transform the experience of the counselling process.

In the present thesis, we will discuss the ethical concerns raised by this radical change, focusing specifically on the potential use of the therapeutic chatbots in helplines and counselling support lines, which until now are staffed by human professionals. In particular, we will consider the ethical challenges that this possibility creates, both because of its effect on the quality of services provided, and because of the redefinition it can bring about to the vital concepts of the therapeutic interaction and relationship. In parallel, we will study whether the use of A.I. systems in mental health is compatible with the principles of Ethics, by proceeding to the overview of the phenomenon through the four principles of bioethics and two of the main streams of moral philosophy, namely Utilitarianism and Deontology.

Keywords: Artificial Intelligence, Conversational AI, chatbot, mental health, Counselling, hotlines, AI Ethics, Bioethics

Εισαγωγή

Στόχος της παρούσας διατριβής είναι να αναδείξει τις ηθικές διαστάσεις της ενσωμάτωσης της Τεχνητής Νοημοσύνης στον τομέα της ψυχικής φροντίδας και ειδικότερα, στις τηλεφωνικές γραμμές βοήθειας και συμβουλευτικής υποστήριξης. Το κύριο ερώτημα αφορά στο εάν η ενσωμάτωση αυτή θα μπορούσε να είναι αποτελεσματική και να συμβάλλει στην βελτίωση της ποιότητας των συμβουλευτικών υπηρεσιών, ούσα ευθυγραμμισμένη με τις ηθικές εκείνες αρχές και τους ηθικούς κανόνες που διέπουν την εφαρμογή της συμβουλευτικής διαδικασίας.

Για να απαντήσουμε στο ερώτημα αυτό, στο πρώτο κεφάλαιο, θα ξεκινήσουμε από την ανάπτυξη των εννοιών των δύο κύριων πεδίων που θα μας απασχολήσουν, αυτών της Τ.Ν. και της Συμβουλευτικής. Αφού εξηγήσουμε τις βασικές έννοιες που τα συνθέτουν, θα εισέλθουμε στους τρόπους με τους οποίους τα δύο αυτά πεδία συνδυάζονται τα τελευταία χρόνια, εστιάζοντας συγκεκριμένα στα εργαλεία της Τ.Ν. που εφαρμόζονται σήμερα, για την επίτευξη της ψυχικής φροντίδας.

Εστιάζοντας περαιτέρω στα εν λόγω εργαλεία, στο δεύτερο κεφάλαιο, θα περιγράψουμε τους τρόπους με τους οποίους αυτά επιφέρουν βελτιώσεις στην παρεχόμενη φροντίδα, αυξάνοντας την πρόσβαση, την ακριβέστερη και την ταχύτερη διάγνωση και ανίχνευση των ψυχικών διαταραχών. Ταυτόχρονα όμως, θα αναφερθούμε στις επιστημολογικές και μεθοδολογικές αδυναμίες που χαρακτηρίζουν τη χρήση και εφαρμογή τους. Πέρα από το παραπάνω, θα επισημάνουμε τους παράγοντες εκείνους βάσει των οποίων δύναται να προβούμε στην αξιολόγηση της χρήσης αυτής και της εφαρμογής, ενώ κλείνοντας το δεύτερο κεφάλαιο, θα αναφερθούμε στις ενδεχόμενες μελλοντικές εξελίξεις στον τομέα της ψυχικής υγείας, όπως αυτές προβλέπεται να λάβουν χώρα, χάριν της ασταμάτητης προόδου που η Τ.Ν. της προφέρει.

Συνεχίζοντας στο τρίτο κεφάλαιο, θα επικεντρώσουμε την προσοχή μας στις τηλεφωνικές γραμμές βοήθειας και συμβουλευτικής υποστήριξης, περιγράφοντας αρχικά, τον τρόπο που αυτές υλοποιούνται αυτή τη στιγμή και εξετάζοντας στη συνέχεια, τα οφέλη που θα μπορούσε να επιφέρει στη λειτουργία τους η ενσωμάτωση των συστημάτων της Τ.Ν., καθώς και τους προβληματισμούς που αυτή θα μπορούσε να εγείρει. Συγκεκριμένα, το ζητούμενο του τρίτου κεφαλαίου είναι η εξέταση του εάν ένα chatbot ψυχικής υγείας, που παρέχει συμβουλευτική υποστήριξη, θα μπορούσε να ανταποκριθεί επαρκώς στις απαιτήσεις μιας θεραπευτικής σχέσης και η λειτουργία του να ευθυγραμμιστεί με τις δεξιότητες εκείνες, που απαιτούνται αντίστοιχα, από πλευράς ενός ανθρώπου συμβούλου. Για την επίτευξη αυτού, επιχειρείται η παρουσίαση ενός υποθετικού περιστατικού σε μία γραμμή βοήθειας και συμβουλευτικής υποστήριξης. Με τη βοήθεια αυτού, επιδιώκεται η κατανόηση στην πράξη, του εάν η ενσωμάτωση ενός chatbot σε μία γραμμή θα μπορούσε να βελτιώσει την αποτελεσματικότητά της, είτε ως εργαλείο στα χέρια του θεραπευτή, είτε ως αυτόνομο σύστημα που διαχειρίζεται μια περίπτωση κλήσης χωρίς την ανθρώπινη παρέμβαση.

Στο τέταρτο και τελευταίο κεφάλαιο της διατριβής, το θέμα της εργασίας μας προσεγγίζουμε μέσω της Ηθικής. Αρχικά, προβαίνουμε στην εξέτασή του μέσω των τεσσάρων αρχών της Βιοηθικής, δίνοντας έμφαση στα κυρίαρχα ζητήματα της ιδιωτικότητας, της έρευνας, του σχεδιασμού και της εφαρμογής, καθώς και της ασφάλειας του χρήστη, θέματα που προκύπτουν κατά την ενσωμάτωση και χρήση ενός συστήματος της Τ.Ν. στον τομέα της ψυχικής υγείας και τα οποία άλλοτε υπηρετούν και άλλοτε διακυβεύουν τις θεμελιώδεις αυτές αρχές. Έπειτα, προσεγγίζουμε το θέμα μας μέσω των φιλοσοφικών ρευμάτων της δεοντοκρατίας και του ωφελιμισμού, σε μια προσπάθεια να διαπιστώσουμε το κατά πόσο οι φιλοσοφικές αυτές θεωρίες δύναται να παρέχουν κατευθύνσεις γύρω από τους ηθικούς κινδύνους που προκύπτουν και να

δρομολογήσουν τη σύνθεση των ηθικών κανονιστικών πλαισίων που θα πλαισιώσουν την επιστράτευση της Τ.Ν. για την επίτευξη της ψυχικής ευημερίας του ανθρώπου. Η διατριβή ολοκληρώνεται με τα συμπεράσματα της μελέτης, όπως αυτά προκύπτουν από την υπάρχουσα βιβλιογραφία και αρθρογραφία που χρησιμοποιήθηκε. Η κύρια σκέψη στην οποία και καταλήγουμε είναι πως παρά τις σημαντικές και ανεκτίμητες προόδους που η Τ.Ν. μπορεί να επιφέρει στο κομμάτι της φροντίδας του ανθρώπινου ψυχισμού, η ενσωμάτωσή της σε αυτό και η ευρεία διάθεση των εφαρμογών της στο κοινό αν μη τι άλλο, πρέπει να έπεται της αναγνώρισης και αποτροπής των σημαντικών ηθικών κινδύνων που εμπεριέχει.

Συντομογραφίες και ακρωνύμια

ΓΚΠΔ	Γενικός Κανονισμός Προστασίας Δεδομένων
ΔΜΣ	Δείκτη Μάζας Σώματος
ΜΚΟ	Μη Κυβερνητική Οργάνωση
ΚΕΘΕΑ	Κέντρο Θεραπείας Εξαρτημένων Ατόμων
T.N.	Τεχνητή Νοημοσύνη
ΟΚΑΝΑ	Οργανισμός κατά των ναρκωτικών
Π.Ο.Υ.	Παγκόσμιος Οργανισμός Υγείας
ΠΦΥ	Πρωτοβάθμια Φροντίδα Υγείας
A.I.	Artificial Intelligence
AFI	Acceptance Facilitating Interventions
ARM	Agnew Relationship Measure
BACP	British Association of Counselling and Psychotherapy
CAI	Conversational AI
CBT	Cognitive Behavioral Therapy
DSM-5	Diagnostic and Statistical Manual of Mental Disorders
DTA	Digital Therapeutic Alliance
EDVAC	Electronic Discrete Variable Automatic Computer
eHealth	Electronic Health
EMA	Ecological Momentary Assessment
EMI	Ecological Momentary Interventions
FRT	Facial recognition Technology
GDPR	General Data Protection Regulation
GOFAI	Good Old Fashioned AI systems
GPS	Global Positioning System
IBI	Internet-based intervention
ICD-11	International Classification of Diseases 11th Revision
IoT	Internet of Things
LLM	Large Language Model
MARS	Mobile Application Rating Scale
mHealth	Mobile Health
NLP	Natural language processing
NLU	Natural Language Understanding
PHQ-9	Nine-Item Patient Health Questionnaire
RDoC	Research Domain Criteria
Snarc	Stochastic Neural Analog Reinforcement Calculator
WAI	Working Alliance Inventory

Κεφάλαιο 1^ο: Συμβουλευτική υποστήριξη και Τεχνητή Νοημοσύνη.

1.α. Τεχνητή Νοημοσύνη: Ο ορισμός και η ιστορική εξέλιξή της μέχρι και σήμερα.

Σύμφωνα με τον ορισμό του John MacCarthy (1956), ως *Τεχνητή Νοημοσύνη* (T.N.) εννοείται η επιστήμη και η μηχανική κατασκευής έξυπνων μηχανών.¹ Σε έτερο ορισμό, που συναντούμε στο βιβλίο με τίτλο *The Cambridge Handbook of Artificial Intelligence*, οι Frankish & Rumsey ορίζουν ως T.N. την διεπιστημονική προσέγγιση για την κατανόηση, τη μοντελοποίηση και την αναπαραγωγή της νοημοσύνης και των γνωστικών διαδικασιών, μέσα από την επίκληση διαφόρων υπολογιστικών, μαθηματικών, λογικών, μηχανικών, ακόμη και βιολογικών αρχών και συσκευών.²

Ο όρος Τεχνητή Νοημοσύνη εισήχθη από τον Αμερικανό μαθηματικό John McCarthy,³ στο πλαίσιο συνεδρίου ερευνητών στο Dartmouth College (1956), με θέμα τη δυνατότητα των υπολογιστών στην προσομοίωση της ανθρώπινης νοημοσύνης. Στο συνέδριο αυτό, πραγματοποιήθηκε επισήμως η έναρξη του προγράμματος της T.N., όταν οι Allen Newell και Herbert Simon παρουσίασαν τη *Μηχανή λογικής θεωρίας* (*The Logic Theory Machine*), ένα πολύπλοκο σύστημα επεξεργασίας πληροφοριών, ικανό να ανακαλύπτει αποδείξεις για τα θεωρήματα της συμβολικής λογικής.⁴ Μια στιγμή ιστορική, για την οποία το έδαφος είχαν προλειαίνει, αρκετά χρόνια νωρίτερα, ερευνητές όπως ο Άγγλος μαθηματικός Alan Matheson Turing⁵ και ο Ούγγρος μαθηματικός John von Neumann.⁶

¹ Pavel Hamet, Johanne Tremblay, “Artificial intelligence in medicine,” *Metabolism Clinical Experimental* 69 (Apr., 2017): 37, <http://dx.doi.org/10.1016/j.metabol.2017.01.011>.

² Άλκης Γούναρης, Γεώργιος Κωστέλετος, *Τεχνητή Νοημοσύνη και Ηθική, Ενότητα 1^η: Εισαγωγική Συνάντηση* [πανεπιστημιακές σημειώσεις]. ΕΚΠΑ, Τμήμα Φιλοσοφίας, ΠΜΣ Φιλοσοφίας: Εφαρμοσμένη Ηθική, χειμερινό εξάμηνο 2022-2023, Αθήνα.

³ Ο John McCarthy (1927-2011) ήταν ο εφευρέτης της γλώσσας προγραμματισμού *Lisp*. Θεωρώντας πως η επίλυση των συλλογιστικών προβλημάτων προϋποθέτει την κατανόησή τους, υποστήριξε με συνέπεια μια μεθοδολογία έρευνας που χρησιμοποιεί λογικές τεχνικές για την διατύπωσή τους. Βλ. Richmond Thomason, “Logic and Artificial Intelligence,” *Stanford Encyclopedia of Philosophy*, accessed August 18, 2023, <https://plato.stanford.edu/entries/logic-ai/#jmc>.

⁴ Βλ. σχετικά στο Allen Newell and Herbert Simon, “The logic theory machine: A complex information processing system,” *IRE Transactions on Information Theory*, vol. 2, no. 3, (Sep. 1956): 61, doi: 10.1109/TIT.1956.1056797.

⁵ Στο έργο του Turing (1912-1954) και κυρίως, σε όσα αυτός ανέλυσε στο μνημειώδες άρθρο του με τίτλο “Computing, Machinery and Intelligence” (1950), βασίστηκε η Δοκιμασία Turing (*Turing Test*). Αυτή, εισηγήθηκαν μεταγενέστεροι του Turing ερευνητές, ως κριτήριο αξιολόγησης των συστημάτων T.N., σχετικά με τη δυνατότητα των τελευταίων να επιδεικνύουν στοιχεία ανθρώπινης νόησης. Βλ. σχετικά στο: Selmer Bringsjord, Naveen Sundar Govindarajulu, “Artificial Intelligence,” *Stanford Encyclopedia of Philosophy*, accessed July 30, 2023, <https://plato.stanford.edu/archives/fall2022/entries/artificial-intelligence/>.

⁶ Το 1945, ο von Neumann παρουσίασε το προσχέδιο του ψηφιακού υπολογιστή γενικής χρήσης με την ονομασία *EDVAC* (*Electronic Discrete Variable Automatic Computer*), θέτοντας έτσι τις βάσεις για την ανάπτυξη αλγορίθμων που μηχανοποιούν την ανθρώπινη σκέψη και λογική. Βλ. Jack B. Copeland, “The Modern History of Computing,” *Stanford Encyclopedia of Philosophy*, accessed August 20, 2023, <https://plato.stanford.edu/entries/computing-history/>.

Το έργο των Simon και Newell και η από μέρους τους εισήγηση της *Υπόθεσης Συστήματος Φυσικών Συμβόλων (Physical Symbol System Hypothesis)*, αποτέλεσε την απαρχή του υπο-πεδίου της *Συμβολικής Τ.Ν. (Symbolic A.I. ή GOFAI)*, βαθιά εμπνευσμένο από την πυθαγόρεια φιλοσοφία και τον αριθμητικό αναγωγισμό που αυτή εισήγαγε, βάσει της άποψης πως οι αριθμοί και συγκεκριμένα, η *Μονάς* και η *Διάς*⁷ συνθέτουν τον *Κόσμο*.⁸ Πιο συγκεκριμένα, η Συμβολική Τ.Ν. βασίζεται στην αντίληψη πως η νοήμον συμπεριφορά δεν είναι παρά ένας λογικός χειρισμός νοητικών συμβόλων (δηλαδή ιδεών). Έτσι, δύναται να αναπαρασταθεί από οποιοδήποτε φυσικό σύστημα συμβόλων, όπως ένα κλασσικό υπολογιστικό πρόγραμμα, που θα μπορούσε να επεξεργαστεί σύμβολα με τη μορφή *χαρακτήρων*, εφόσον προγραμματιστεί αναλόγως για αυτό (*αλγοριθμοποίηση* της συμπεριφοράς).⁹ Στόχος της Συμβολικής Τ.Ν. λοιπόν, είναι η μελέτη της φύσης της ανθρώπινης νοημοσύνης και η μετέπειτα προσπάθεια για την αναπαραγωγή αυτής σε υπολογιστές, μέσα από τη χρήση συμβολικών αναπαραστάσεων για την επεξεργασία των πληροφοριών και την επίλυση των προβλημάτων.¹⁰

Η προσέγγιση της Συμβολικής Τ.Ν. επικράτησε μέχρι και το 1980, όπου καταγράφηκε μία σημαντική στροφή στη *Συνδετική Τ.Ν. (Connectionist A.I. ή Connectionism)*. Πρόκειται για το υπο-πεδίο εκείνο, κατά το οποίο, η συμπεριφορά ενός νοήμονος όντος γίνεται αντιληπτή ως το αποτέλεσμα συγκεκριμένων εγκεφαλικών διεργασιών. Έτσι, στόχο της αποτελεί η αναπαράσταση της λειτουργίας των βιολογικών δομών του ανθρώπινου εγκεφάλου, που υποστηρίζουν τη νοητική διαδικασία, μέσω της χρήσης τεχνητών νευρωνικών δικτύων.¹¹ Η Συνδετική Τ.Ν. πρωτοστάτησε τη δεκαετία του 1980 και συνεχίζει να πρωτοστατεί μέχρι και σήμερα, με την ανάπτυξη των συστημάτων *Machine learning* και *Deep learning*¹² να προσφέρουν σημαντικά αποτελέσματα σε πολλά πεδία. Η αρχή της ωστόσο, ως υπο-πεδίο της Τ.Ν., σημειώθηκε αρκετά χρόνια νωρίτερα και συγκεκριμένα το 1943, όταν οι McCulloch και Pitts παρουσίασαν ένα μοντέλο τεχνητών νευρώνων με τη δυνατότητα να μαθαίνει και να προσομοιώνει λογικές πράξεις, με τους Minsky και Edmonds να ακολουθούν, παρουσιάζοντας το 1951, το πρώτο νευρωνικό δίκτυο με την ονομασία *Snarc (Stochastic Neural Analog Reinforcement Calculator)*.¹³

⁷ Η *Μονάς* είναι ο Θεός δηλ. το πρωταρχικό Ον και η *Διάς*, που προέρχεται από αυτήν, γεννά τους υπόλοιπους αριθμούς. Βλ. Γεώργιος Κωστελέτος, *Τεχνητή Νοημοσύνη και Ηθική, Ενότητα 2^η: Το πρόγραμμα της Τεχνητής Νοημοσύνης* [πανεπιστημιακές σημειώσεις]. ΕΚΠΑ, Τμήμα Φιλοσοφίας, ΠΜΣ Φιλοσοφίας: Εφαρμοσμένη Ηθική, χειμερινό εξάμηνο 2022-2023, Αθήνα

⁸ *Αυτόθι*.

⁹ *Αυτόθι*: Κατερίνα Γεωργούλη, *Τεχνητή Νοημοσύνη. Μια εισαγωγική προσέγγιση* (Αθήνα: Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών, 2015), 14, <https://repository.kallipos.gr/bitstream/11419/3381/3/%CE%A4%CE%B5%CF%87%CE%BD%CE%B7%CF%84%CE%AE%20%CE%9D%CE%BF%CE%B7%CE%BC%CE%BF%CF%83%CF%8D%CE%BD%CE%B7.pdf>.

¹⁰ Γεωργούλη, *Τεχνητή Νοημοσύνη*, 14.

¹¹ Κωστελέτος, *Τεχνητή Νοημοσύνη και Ηθική, Ενότητα 2^η* [πανεπιστημιακές σημειώσεις], ΕΚΠΑ.

¹² *Αυτόθι*. Τα συστήματα *Machine learning* είναι *Τεχνητά Νευρωνικά Δίκτυα* σε υλισμικό επίπεδο, τα οποία είναι ικανά στην εκμάθηση και ανίχνευση μοτίβων (*patterns*) μέσα σε μεγάλα σύνολα δεδομένων. Αντίστοιχα, τα συστήματα που αποκαλούνται *Deep learning* είναι τεχνητά νευρωνικά δίκτυα, που περιλαμβάνουν επιπλέον, ένα στρώμα νευρώνων εισόδου των πληροφοριών, ένα στρώμα νευρώνων εξόδου των αποτελεσμάτων της επεξεργασίας και αρκετά ενδιάμεσα στρώματα νευρώνων, με αποτέλεσμα η αρχιτεκτονική τους να είναι περίπλοκη.

¹³ Φίλιππος Ζηκόπουλος, «Η χρήση της τεχνητής νοημοσύνης και της τεχνολογίας στην πρόληψη και στην βελτίωση της υγείας» (Διπλωματική διατριβή, Πανεπιστήμιο Δυτικής Αττικής, 2021), 15-17.

Έκτοτε, οι στιγμές εκείνες που σηματοδότησαν την ιστορική εξέλιξη της Τ.Ν. ήταν αρκετές. Κατά τα έτη 1971-1977 δημιουργήθηκαν τα πρώτα συστήματα της Τ.Ν. (DENDRAL, MYCIN και Prospector), ενώ την ίδια περίοδο, ολοκληρώθηκε η δημιουργία της γλώσσας λογικού προγραμματισμού με την ονομασία *PROLOG* και αναπτύχθηκαν οι εξελικτικοί αλγόριθμοι, με τη μετέπειτα έρευνα και βιβλιογραφία να εστιάζει στη βελτίωση και εξέλιξη της μέχρι τότε γνώσης. Την περίοδο 1980-1990, τα Τεχνητά Νευρωνικά Δίκτυα αναγεννήθηκαν και δημιουργήθηκαν υπολογιστικά συστήματα, με τη δυνατότητα προσαρμογής στο περιβάλλον τους (π.χ. ρομπότ), καθώς και εφαρμογές με τη δυνατότητα μάθησης μέσω της εμπειρίας (λ.χ. διαδικτυακές μηχανές αναζήτησης).¹⁴ Μέχρι σήμερα η ανάπτυξη είναι ταχεία, με μια από τις πρόσφατες εξελίξεις να αφορά στην εφαρμογή της *Παραγωγικής Τ.Ν. (Generative AI)*, που από το 2014, είναι αποτελεσματική στο να παράγει διαφόρων τύπων πρωτότυπο περιεχόμενο, όπως κείμενο, γραφικά, μουσική και βίντεο μέσα σε λίγα μόλις δευτερόλεπτα. Ένα δημοφιλές δείγμα εφαρμογής της Παραγωγικής Τ.Ν. είναι τα λεγόμενα *chatbot* (όπως το DALL-E, το ChatGPT και το Bard), που όντας εκπαιδευμένα σε μεγάλα σύνολα δεδομένων, καθίστανται χρήσιμα εργαλεία για την ικανοποίηση ποικίλων αναγκών (λ.χ. εξυπηρέτηση πελατών, δημιουργία τέχνης, δημιουργία φαρμακευτικών ενώσεων προς δοκιμή κ.ο.κ.).¹⁵

Κάνοντας μια αναδρομή στο παρελθόν, εύκολα διαπιστώνουμε πως ο αρχικός στόχος της κατασκευής έξυπνων μηχανών που θα διαθέτουν διανοητικές ικανότητες όμοιες με του ανθρώπου, έχει διαφοροποιηθεί, λαμβάνοντας μια πιο πρακτική μορφή και στοχεύοντας στην κατασκευή συστημάτων που θα μπορέσουν να εστιάσουν σε συγκεκριμένες εργασίες. Ένας από τους ορισμούς που το αποτυπώνει αυτό, ανήκει στους Rich & Knight (1990), που αναφέρουν πως η «Τεχνητή Νοημοσύνη είναι η μελέτη του πώς να κάνουμε τους υπολογιστές ικανούς να κάνουν πράγματα στα οποία προς το παρόν οι άνθρωποι τα καταφέρνουν καλύτερα».¹⁶ Έτσι, σημειώθηκε η σταδιακή μετακίνηση από τη στόχευση της «Ισχυρής Τ.Ν.» στην επικύρωση μιας «Ασθενούς Τ.Ν.»,¹⁷ όπου κύριο αντικείμενο είναι η κατασκευή συστημάτων που αναπαράγουν την ανθρώπινη συμπεριφορά και αποτελούν χρήσιμα εργαλεία για τον άνθρωπο, χωρίς να έχει ιδιαίτερη σημασία ο τρόπος που αυτό επιτυγχάνεται.¹⁸

Στην μέχρι τώρα εξέλιξη της Τ.Ν., σημαντική υπήρξε η επιρροή διαφόρων επιστημονικών κλάδων. Μεταξύ αυτών η φιλοσοφία, με την αριστοτελική θεωρία να χαρτογραφεί τη Λογική και παρέχει μία μέθοδο συλλογιστικής διαδικασίας, την οποία διάφοροι στοχαστές όπως ο Gottfried Wilhelm Leibniz, καθώς και μηχανικοί όπως ο Charles Babbage προσπάθησαν να μηχανοποιήσουν τους αιώνες που ακολούθησαν. Επιπλέον, η νευρολογία, της οποίας τα πορίσματά για τη λειτουργία του ανθρώπινου νευρικού συστήματος, η Τ.Ν. επεδίωξε να μιμηθεί και να αναπαραγάγει και η

¹⁴ *Αυτόθι*.

¹⁵ George Lawton, “What is generative AI? Everything you need to know,” Techtargget: Enterprise AI, accessed August 20, 2023, <https://www.techtargget.com/searchenterpriseai/definition/generative-AI>.

¹⁶ Γεωργούλη, *Τεχνητή Νοημοσύνη*, 14.

¹⁷ Η «Ισχυρή Τ.Ν.» (*Strong AI*) στοχεύει στη δημιουργία συστημάτων που διαθέτουν όλες τις νοητικές δυνατότητες που διαθέτει και ο άνθρωπος, ενώ στόχος της «Ασθενούς Τ.Ν.» (*Weak AI*) είναι η δημιουργία συστημάτων που εμφανίζουν ή αλλιώς, «φαίνεται ότι διαθέτουν» πλήρως το σύνολο των νοητικών χαρακτηριστικών του ανθρώπου. Βλ. Bringsjord, Govindarajulu, “Artificial Intelligence.”

¹⁸ Αθανάσιος Πρωτόπαπας, *Εισαγωγή στην θεωρία και Μεθοδολογία της Γνωσιακής Επιστήμης* [πανεπιστημιακές σημειώσεις], σελ 8. ΕΚΠΑ, Τμήμα Ιστορίας και Φιλοσοφίας της Επιστήμης, 2021, Αθήνα, http://users.uoa.gr/~aprotopapas/CV/pdf/Protopapas_cogsci-notes_2011.pdf

γλωσσολογία, που παρείχε στην Τ.Ν. τη δυνατότητα κατανόησης της γλώσσας και αναπαράστασης γλωσσικά αποδεκτών προτάσεων. Ακόμη, οι επιστήμες των υπολογιστών και της μηχανικής παρείχαν στην Τ.Ν. τα αναγκαία εργαλεία για την ανάπτυξη της (λ.χ. λογισμικά συστήματα, γλώσσες προγραμματισμού), ενώ σημαντική υπήρξε και η επιρροή της Ψυχολογίας και κυρίως του συμπεριφορισμού, που ως ρεύμα, με ηγέτη τον John Watson, έλαβε χώρα κατά τις δεκαετίες από το 1920 έως το 1960. Σύμφωνα με τη συμπεριφοριστική θεωρία, ο ανθρώπινος εγκέφαλος αναγνωρίζεται ως μία συσκευή που επεξεργάζεται πληροφορίες. Βασικό της αξίωμα είναι το συμπέρασμα πως η ανθρώπινη συμπεριφορά είναι αποτέλεσμα εκμάθησης, που προκύπτει από την αλληλεπίδραση μεταξύ των ερεθισμάτων που προέρχονται από το εξωτερικό περιβάλλον και των τρόπων που το άτομο αποκρίνεται σε αυτά. Έτσι, η άποψη πως η γνωστική διαδικασία θα πρέπει να μοιάζει με πρόγραμμα υπολογιστή, αποτέλεσε το πρόσφορο έδαφος για την υπολογιστική μοντελοποίηση της διαδικασίας της μάθησης από τους ερευνητές.¹⁹

Σήμερα, τα συστήματα της Τ.Ν. μπορούν να χρησιμοποιηθούν ποικιλοτρόπως και σε διάφορους τομείς, όπως η υγεία, η τέχνη, το περιβάλλον ή η εκπαίδευση και αν μη τι άλλο, η συμβολή τους είναι καθοριστική για τη βελτίωση των δυνατοτήτων ή και της ποιότητας των υπηρεσιών που παρέχονται. Ωστόσο, παράλληλα με τη συνεχή πρόοδο των συστημάτων της Τ.Ν. που σημειώνεται, εγείρονται σοβαρά ηθικά ερωτήματα σχετικά με το πώς θα αυτά πρέπει να λειτουργήσουν, τους κινδύνους που ελλοχεύουν και το πώς αυτοί δύναται να προβλεφθούν και να ελεγχθούν. Οι παραπάνω προβληματισμοί σταδιακά οδήγησαν στην ανάπτυξη του υπο-πεδίου της Εφαρμοσμένης Ηθικής που αποκαλείται *Ηθική της Τ.Ν. (AI Ethics)* καθώς και του υπο-πεδίου έρευνας της Τ.Ν. με την ονομασία *Ηθική των Μηχανών (Machine Ethics)*.²⁰

Την Ηθική της Τ.Ν. θεμελιώνει η αντίληψη πως ο σχεδιασμός των συστημάτων της Τ.Ν. είναι ηθικά συναφής με τη χρήση τους και επομένως, πέρα από την υπεύθυνη χρήση απαιτείται και ο υπεύθυνος σχεδιασμός τους. Ως πεδίο εστιάζει στα ηθικά ζητήματα που εγείρονται από την ανάπτυξη και χρήση των συστημάτων της Τ.Ν. Τα ερωτήματα αυτά σχετίζονται, μεταξύ άλλων, με τις έννοιες της ιδιωτικότητας, της χειραγώγησης, της αδιαφάνειας, της μεροληψίας και της ανεργίας και προκύπτουν άμεσα ή μακροπρόθεσμα, από την εφαρμογή των συστημάτων της Τ.Ν. ως *αντικείμενα*, τα οποία κατασκευάζονται και χρησιμοποιούνται από τον άνθρωπο.²¹

Παράλληλα, κύριος στόχος της Ηθικής των Μηχανών είναι η δημιουργία αμερόληπτων συστημάτων και η διασφάλιση ότι αυτά, θα συμπεριφέρονται με ηθικούς όρους κατά την αλληλεπίδρασή τους με τον άνθρωπο. Για τους εμπνευστές του συγκεκριμένου αυτού ερευνητικού πεδίου, Suzan και Michael Anderson, η διαφορά από τον κλάδο της Ηθικής Τ.Ν. είναι μικρή, αλλά σημαντική και αφορά στην εστίαση της Ηθικής των Μηχανών στα συστήματα που αναπτύσσονται με αλγορίθμους βαθιάς μάθησης. Αυτό, γιατί τα συστήματα αυτά χειρίζονται μεγάλους όγκους δεδομένων και δε δύναται να ελεγχθούν λόγω της εξαιρετικά σύνθετης αρχιτεκτονικής τους, γεγονός που δημιουργεί σημαντικά εμπόδια στην επίτευξη της ηθικής συμπεριφοράς τους.²²

¹⁹ Ζηκόπουλος, «Η χρήση της τεχνητής νοημοσύνης», 12-13.

²⁰ Vincent Müller, “Ethics of Artificial Intelligence and Robotics,” Stanford Encyclopedia of Philosophy, accessed August 21, 2023, <https://plato.stanford.edu/entries/ethics-ai/>.

²¹ *Αυτόθι*.

²² Michael Anderson, Suzan L. Anderson, Alkis Gounaris, George Kosteletos, “Towards Moral Machines: A Discussion with Michael Anderson and Susan Leigh Anderson,” *Conatus* 6, no. 1 (2021): 177-179, <http://dx.doi.org/10.12681/cjp.26832>.

Γίνεται αντιληπτό από τα παραπάνω πως, η ηθική πλαισίωση της ανάπτυξης και εφαρμογής των συστημάτων της Τ.Ν. αποτελεί μια παράμετρο ζωτικής σημασίας, που δε θα μπορούσαμε να παραβλέψουμε, εξετάζοντας τις εφαρμογές της σε όλους τους τομείς και ιδίως, στον τομέα της Ψυχικής υγείας και συγκεκριμένα της *Συμβουλευτικής*, που θα μας απασχολήσει στην παρούσα διατριβή. Αυτό γιατί, η παράβλεψη των ηθικών αρχών και αξιών που θα διέπουν την ανάπτυξη και την εφαρμογή των εν λόγω συστημάτων στη συμβουλευτική διαδικασία, θα μπορούσε να διακυβεύσει την ασφάλεια κατά τη χρήση, καθώς και θεμελιώδη ανθρώπινα δικαιώματα, όπως η ελευθερία και η ισότητα.

Τα θέματα αυτά θα μας απασχολήσουν στις σελίδες που ακολουθούν, ξεκινώντας από την επόμενη ενότητα, στην οποία θα εξετάσουμε το γνωστικό πεδίο της Συμβουλευτικής. Συγκεκριμένα, θα δούμε τον ορισμό και τους στόχους της, καθώς και τις δεξιότητες που απαιτούνται από τους λειτουργούς της, ώστε αυτή να είναι αποτελεσματική. Επιπλέον, θα δούμε τους τρόπους με τους οποίους η συμβουλευτική διαδικασία εφαρμόζεται μέχρι και σήμερα, όπου η Τ.Ν. σταδιακά μα δυναμικά, ενσωματώνεται σε αυτήν. Άλλοτε, εξάπτοντας τον ενθουσιασμό για τα όσα μπορεί να της προσφέρει και άλλοτε, εγείροντας ερωτήματα για το εάν θα μπορέσει ποτέ να ευθυγραμμιστεί με τα ουσιώδη χαρακτηριστικά της.

1.β. Η Συμβουλευτική υποστήριξη στην Ψυχική υγεία

Όπως αναγράφεται στο *Λεξικό Νέας Ελληνικής Γλώσσας* του Μπαμπινιώτη (2002), ο όρος *συμβουλευτική*, ετυμολογικά αποτελείται από την πρόθεση *συν*, που σημαίνει μαζί και το ουσιαστικό *βουλή*, που δηλώνει τη σκέψη και την απόφαση ή το ρήμα *βουλεύομαι*, που σημαίνει σκέφτομαι, κάνω σχέδια ή αποφασίζω μετά από σύσκεψη με κάποιον άλλο.²³ Στην ουσία, ο όρος δηλώνει τη διαδικασία κατά την οποία, η συζήτηση βοηθά ένα άτομο να κατανοήσει την κατάσταση που βιώνει και να βρει τρόπους να αντιμετωπίσει τυχόν δυσκολίες που σχετίζονται με αυτήν. Δύο άνθρωποι δηλαδή, συνσκέπτονται με στόχο την έκφραση, την αναγνώριση και την επίλυση μιας δυσκολίας.

Η συμβουλευτική απευθύνεται σε ανθρώπους που δεν πάσχουν από κάποια σοβαρή ψυχική διαταραχή, βιώνουν όμως μια αρνητική κατάσταση και αισθάνονται την ανάγκη καθοδήγησης από ειδικό.²⁴ Η στιγμή που απευθύνονται σε αυτόν, περιλαμβάνει συναισθήματα σύγχυσης, καθώς και την επιθυμία επίλυσης των όποιων προβλημάτων, σε ένα περιβάλλον σχέσης, το οποίο θα είναι περισσότερο πειθαρχημένο και εμπιστευτικό από το φιλικό και ίσως, τους στιγματίσει λιγότερο από τις σχέσεις που λαμβάνουν χώρα στο παραδοσιακό ιατρικό ή ψυχιατρικό πλαίσιο.²⁵

Οι ορισμοί που κατά καιρούς έχουν διατυπωθεί για τη συμβουλευτική είναι αρκετοί. Μεταξύ άλλων, η Βρετανική Εταιρεία Συμβουλευτικής (British Association of Counselling and Psychotherapy - BACP) το 1984, όρισε ως συμβουλευτική «την

²³ Μαρία Μαλικιώση - Λοΐζου, «Η Συμβουλευτική Ψυχολογία στην Ελλάδα σήμερα», *Hellenic Journal of Psychology*, Vol. 8 (Ιαν. 2011): 267, <file:///C:/Users/estia/Downloads/Malikiosi-Loizou8cHellenicJournalofPsychology.pdf>.

²⁴ *Αυτόθι*.

²⁵ Ευφροσύνη Κατσικονούρη, Ευθυμία Μανίτσα, «Δεξιότητες συμβουλευτικής 'αντίδοτο' στην κρίση», *Ερευνώντας τον κόσμο του παιδιού* 13 (2014): 178; John McLeod, *Εισαγωγή στην συμβουλευτική* (Αθήνα: Μεταίχμιο, 2003), 30.

εργασία με άτομα και με σχέσεις, η οποία μπορεί να συνδέεται με την ανάπτυξη, την υποστήριξη κρίσεων, να είναι ψυχοθεραπευτική, καθοδηγητική ή να έχει στόχο την επίλυση προβλημάτων». Ο σκοπός της είναι να δώσει στον συμβουλευόμενο την ευκαιρία «να εξερευνήσει, να ανακαλύψει και να αποσαφηνίσει τους τρόπους που θα του παράσχουν μια ζωή πιο ικανοποιητική και δημιουργική».²⁶ Παρόμοιος είναι και ο ορισμός της Αμερικανικής Ένωσης Συμβουλευτικής, που την ορίζει ως την επαγγελματική εκείνη σχέση που ενδυναμώνει άτομα, ομάδες ή και οικογένειες ώστε να επιτύχουν ψυχική υγεία, ευεξία, στόχους εκπαιδευτικούς και επαγγελματικούς.²⁷

Σε συνέχεια των παραπάνω, δεν θα μπορούσαμε να παραλείψουμε τον ορισμό του Carl Rogers, που από πολλούς έχει χαρακτηριστεί ως ο «πατέρας» της συμβουλευτικής. Ο Rogers λοιπόν, ορίζει τη συμβουλευτική ως τη δομημένη σχέση, εντός της οποίας ένας άνθρωπος δύναται να καταρρίψει τις συναισθηματικές του άμυνες και τις παγιωμένες του αλήθειες και έτσι, να εξερευνήσει ελεύθερα αυτό που τον απασχολεί. Η εξερεύνηση αυτή, παρέχει τη δυνατότητα της κατανόησης του ίδιου του εαυτού, καθώς τα μοτίβα της συμπεριφοράς περνούν στη σφαίρα του συνειδητού και γίνονται αντικείμενο κριτικής σκέψης και εξέτασης. Ως αποτέλεσμα, το άτομο επαναπροσδιορίζει τον εαυτό του και υπό το πρίσμα αυτού του νέου προσανατολισμού, αναπτύσσει την προσωπικότητά του.²⁸

Ως γνωστό πεδίο, η συμβουλευτική εξελίχθηκε κατά τη διάρκεια του εικοστού αιώνα. Αναζητώντας τις ρίζες της έννοιας όμως, φτάνουμε στον αρχαίο φιλόσοφο Σωκράτη και στη διαλεκτική μέθοδο που αυτός εφάρμοσε. Τα λεγόμενά του και ο τρόπος θεώρησης του εαυτού από τον ίδιο, επηρέασαν τη σύγχρονη σκέψη και αποτέλεσαν αναπόσπαστο μέρος της συμβουλευτικής θεωρίας. Σε τέτοιο βαθμό, που δικαίως κατατάσσεται από πολλούς, ως μέλος της συμβουλευτικής κοινότητας.²⁹

Ο Σωκράτης, πρώτος επικεντρώθηκε στην εξέταση και ανάλυση του εαυτού στο πλαίσιο μιας προσωποκεντρικής θεώρησης του κόσμου.³⁰ Ο δελφικός χρησμός *Γνώθι σαυτόν* συνίσταται στην αυτογνωσία και στην κατανόηση των ορίων του εαυτού, τα οποία δεν σχετίζονται μονάχα με τους περιορισμούς της ανθρώπινης σοφίας, αλλά και τους γνωσιακούς, ηθικούς και ψυχολογικούς περιορισμούς του κάθε ανθρώπου.³¹ Ο Σωκράτης μας παρακινεί στην αντίληψη του ανώτερου ορίου γνώσης που μπορεί ο καθένας από εμάς να αγγίξει με βάση τις ικανότητες του. Αυτό, ως τρόπος φροντίδας και βελτίωσης, θα επιτευχθεί μέσα από την αυτοεξέταση και την ανακάλυψη της αλήθειας που ο καθένας κρύβει μέσα του.³²

Είναι σημαντικό να αναφερθεί στο σημείο αυτό πως στα κείμενα που αφορούν στη συμβουλευτική διαδικασία, άλλοτε συναντούμε τον όρο *Συμβουλευτική* και άλλοτε τον όρο *Συμβουλευτική Ψυχολογία*. Γεννάται λοιπόν το ερώτημα του εάν οι δύο όροι

²⁶ «Τι είναι η Συμβουλευτική», Συμβουλευτικό Κέντρο Φοιτητών ΑΣΠΑΙΤΕ, πρόσβαση στις 07 Ιουλίου 2023, <https://skf.aspete.gr/ti-einai-i-symvouleftiki/>.

²⁷ “20/20: Consensus definition of Counseling,” American Counseling Association, accessed June 26, 2023, <https://www.counseling.org/about-us/about-aca/20-20-a-vision-for-the-future-of-counseling/consensus-definition-of-counseling>.

²⁸ Carl R. Rogers, “Counseling,” *Review of Educational Research* 15, no. 2 (Apr. 1945): 155, <https://doi.org/10.2307/1168689>.

²⁹ Μαλικιώση - Λοΐζου, «Η Συμβουλευτική Ψυχολογία στην Ελλάδα σήμερα», 268.

³⁰ *Αυτόθι*.

³¹ Λαμπρινή Μανίκα, «Το πρόσωπο του Σωκράτη μέσα από την οπτική του Γρηγόρη Βλαστού & του Olof Gigon» (Διπλωμ. εργασία, ΑΠΘ, 2013), 51.

³² *Αυτόθι*, 37.

συμπίπτουν ή εκφράζουν κάτι το διαφορετικό. Η αλήθεια είναι πως όσο διεισδύουμε σε αυτούς, παρατηρούμε πως συνδέονται απόλυτα σε ότι αφορά τις μεθόδους που εφαρμόζουν, τους σκοπούς και τις αρχές που τους διέπουν. Αυτός είναι και ο λόγος για τον οποίο συχνά και προς αποφυγή της παρερμηνείας που προκύπτει από το διπλό όρο, αναφέρεται ο όρος *Συμβουλευτική*, ο οποίος καλύπτει το γνωστικό αντικείμενο στο σύνολό του. Η μόνη διαφορά συναντάται στο είδος της επαγγελματικής εκπαίδευσης ενός συμβούλου.³³ Και είναι σίγουρο πως αυτή απαιτείται, καθώς η άσκηση της συμβουλευτικής δεν περιορίζεται στην απλή συζήτηση ή στην απλή παροχή συμβουλών.

Είναι γεγονός πως για πολλά χρόνια, η αλληλεπίδραση μεταξύ συμβούλου και συμβουλευόμενου θεωρείτο πως ανήκει αποκλειστικά στη σφαίρα της τέχνης και πως ήταν απρόσιτη από την επιστημονική μεθοδολογία. Τα ευρήματα, ωστόσο, διαφόρων μελετών έδειξαν ξεκάθαρα πως το αδόμητο υλικό της συνέντευξης μπορεί κάλλιστα να αναλυθεί αντικειμενικά, μιας και κατά τη διάρκεια της συμβουλευτικής εξελίσσεται μία σαφής και ορισμένη διαδικασία από την πλευρά των συμμετεχόντων.³⁴

Μεταξύ άλλων, οι μελέτες των Snyder και Lewis, επιβεβαίωσαν πως οι δηλώσεις ενός προβλήματος κυριαρχούν στις αρχικές συναντήσεις, η αναγνώριση και η ενόραση αυξάνονται στο μέσο της διαδικασίας, ενώ οι επιλογές και οι αποφάσεις λαμβάνουν χώρα στις τελικές συνεδρίες. Αντίστοιχα, παρατηρείται διαφοροποίηση των συναισθημάτων, καθώς τα αρνητικά συναισθήματα των αρχικών φάσεων της συμβουλευτικής, γίνονται αμφίθυμα στη συνέχεια και μετατρέπονται σε θετικά καθώς ολοκληρώνεται η διαδικασία.³⁵ Η ιδιαιτερότητα της κάθε φάσης απαιτεί από το σύμβουλο τον κατάλληλο χειρισμό. Αυτό μπορεί να γίνει μόνο μέσα από την καλλιέργεια των αντίστοιχων δεξιοτήτων.

Κατά τον Rogers, υπάρχουν τρία βασικά χαρακτηριστικά που κάθε σύμβουλος θα πρέπει να πληροί. Πρώτον, να είναι αυθεντικός και αληθινός με τον άνθρωπο που έχει απέναντί του.³⁶ Αυτό μπορεί να περικλείει τις δεξιότητες της αυτό-αποκάλυψης και της ανατροφοδότησης. Η αυτο-αποκάλυψη αφορά στο μοίρασμα προσωπικών βιωμάτων ή και συναισθημάτων από το σύμβουλο. Έρευνες δείχνουν πως κάτι τέτοιο ενθαρρύνει το συμβουλευόμενο να κάνει το ίδιο και πως γενικότερα, οι συμβουλευόμενοι αντικρύζουν θετικά ένα σύμβουλο που αποκαλύπτει πλευρές της προσωπικότητάς του. Σχετική με την αυτο-αποκάλυψη είναι και η ανατροφοδότηση, που έχει να κάνει με την περιγραφή στον συμβουλευόμενο, του τρόπου με τον οποίο τον βλέπουν ή αισθάνονται κοντά του οι άλλοι. Με τον τρόπο αυτό, θα μπορέσει να αντιληφθεί τον αντίκτυπο που έχει στον κόσμο γύρω του.³⁷

Δεύτερον, ένας σύμβουλος οφείλει να είναι ικανός στην άνευ όρων αποδοχή χωρίς να κρίνει ή να αξιολογεί τον άνθρωπο που του απευθύνεται. Οφείλει δηλαδή, να παραμερίσει την τάση του για διαγνώσεις και επαγγελματικές αξιολογήσεις και να εγκαταλείψει κάθε προσπάθεια καθοδήγησης του συμβουλευόμενου στη λήψη της

³³ Σοφία Κώστα, Χρήστος Ράπτης, «Η Συμβουλευτική στις μονάδες υγείας: Τεχνικές για τους διοικητικούς» (Πτυχιακή εργασία, TEI Καλαμάτας, 2003), 13-14.

³⁴ Rogers, "Counseling," 157.

³⁵ *Αυτόθι*.

³⁶ Lisa Ede, "Is Rogerian Rhetoric Really Rogerian?," *Rhetoric Review* 3, no. 1 (Sep. 1984): 44, <http://www.jstor.org/stable/465729>.

³⁷ Μαρία Μαλκιώση - Λοΐζου. *Συμβουλευτική Ψυχολογία* (Αθήνα: Ελληνικά Γράμματα, 2006), 294-295.

όποιας απόφασης.³⁸ Όπως αναφέρει ο Rogers, η αλληλεπίδραση μεταξύ του ατόμου και του συμβούλου, είναι «διαίτερα επιτρεπτική».³⁹ Με τη φράση αυτή, υπερθεματίζει το στοιχείο της αποδοχής και της ελευθερίας που, χάρη στις δεξιότητες του συμβούλου, εισέρχονται στη συμβουλευτική διαδικασία και ανοίγουν το δρόμο προς την αλλαγή. Θα λέγαμε πως η αποδοχή αυτή, δηλώνει το σεβασμό στην προσωπικότητα του ατόμου και αν μη τι άλλο, δημιουργεί ένα περιβάλλον στο οποίο μπορεί με ελευθερία να εκφραστεί και να μοιραστεί τις πιο βαθιές του σκέψεις.⁴⁰ Κατά την βράβειυσή του από την Αμερικανική Ψυχολογική Εταιρεία το 1973,⁴¹ ο Rogers δήλωσε χαρακτηριστικά πως κάθε άνθρωπος διαθέτει τη δυνατότητα κατανόησης και αλλαγής των στάσεων, της συμπεριφοράς του και της γενικότερης αντίληψης που έχει για τον εαυτό του. Η δυνατότητα αυτή δύναται να αξιοποιηθεί μόνο σε ένα περιβάλλον σχέσης, όπου παρέχεται η διευκόλυνση της έκφρασης και η άνευ όρων αποδοχή. Τη ζωτική δύναμη της ικανότητας του ατόμου για εποικοδομητική δράση, συμπληρώνει, μπορεί να την αντιληφθεί μόνο ο θεραπευτής εκείνος, που είναι δεκτικός σε οποιοδήποτε αποτέλεσμα επέλθει και σε οποιαδήποτε απόφαση τελικώς ληφθεί.⁴²

Το τρίτο χαρακτηριστικό που διακρίνει ο Rogers, αποτελεί ίσως τη σημαντικότερη δεξιότητα που ένας αποτελεσματικός σύμβουλος μπορεί να διαθέτει. Είναι η ικανότητα να *εν-συναισθάνεται*, δηλαδή να αντικρύζει τον κόσμο μέσα από τα μάτια του συμβουλευόμενου, σε ένα πλαίσιο βαθιάς κατανόησης των στάσεων που αυτός διατηρεί, ενόσω εξερευνά βήμα-βήμα τις επικίνδυνες περιοχές που αρνείται στη συνείδησή του.⁴³

Είναι γεγονός πως η απόδοση του όρου *ενσυναίσθηση* έχει προκαλέσει σύγχυση, πράγμα που διαφαίνεται στους διαφορετικούς τρόπους με τους οποίους έχει διατυπωθεί (*εμβίωση, εμπάθεια, εναίσθηση, ενσυναίσθητη κατανόηση*). Θεωρείται πως προέρχεται από τη λέξη *εμπάθεια*, που ετυμολογικά προκύπτει από τις λέξεις *έν* και *πάθος* και δηλώνει την κατάσταση συγκίνησης στην οποία βρίσκεται κανείς. Η ερμηνεία αυτή διατηρήθηκε από τους αγγλοσάξωνες, σε αντίθεση με την ελληνική γλώσσα, όπου η λέξη *εμπάθεια* εκφράζει την αντιπάθεια απέναντι σε κάποιον.⁴⁴

Ο αγγλικός όρος *empathy* εισήχθη στην ιατρική κοινότητα κατόπιν αναφοράς του Αμερικανού ψυχιάτρου Southard (1876-1920), ως χαρακτηριστικό της σχέσης ανάμεσα στον ιατρό και τον ασθενή και μέσο διευκόλυνσης του ιατρικού έργου. Μέχρι τότε, την επάρκεια ενός γιατρού καθόριζαν οι ιατρικές γνώσεις, ενώ έκτοτε και μέχρι σήμερα, η φροντίδα της υγείας χαρακτηρίζεται από μια προσέγγιση ολιστική και ανθρωποκεντρική. Η προσέγγιση αυτή, περιλαμβάνει την ουσιαστική επικοινωνία ως συστατικό μια σχέσης εμπιστοσύνης και συνεργασίας με τον ασθενή και κριτήριο αποτελεσματικότητας και ποιότητας των παρεχόμενων υπηρεσιών.⁴⁵

Σύμφωνα με τον ορισμό που δίνει ο Rogers (1959) λοιπόν, η ενσυναίσθηση είναι η ικανότητα ενός ανθρώπου να αντιλαμβάνεται με ακρίβεια τα συναισθήματα ενός

³⁸ Ede, "Is Rogerian Rhetoric Really Rogerian?", 44.

³⁹ Rogers, "Counseling," 155.

⁴⁰ Ede, "Is Rogerian Rhetoric Really Rogerian?", 43-45.

⁴¹ Βραβείο Διακεκριμένης Επαγγελματικής Συνεισφοράς.

⁴² Ede, "Is Rogerian Rhetoric Really Rogerian?", 43-45.

⁴³ Αυτόθι, 44.

⁴⁴ Ευαγγελία Κοσμά, «Η Σχέση της Ενσυναίσθησης και της Ενεργητικής Ενσυναισθητικής Ακρόασης της γονικής σχέσης με την Ψυχοπαθολογία» (Διπλ. Εργασία, ΕΚΠΑ, 2020), 23.

⁴⁵ Αυτόθι, 24.

άλλου, όπως και τις αιτίες προέλευσης αυτών, «σαν να» τα βιώνει ο ίδιος, χωρίς ωστόσο ποτέ να παραβλέπει το γεγονός πως τα συναισθήματα αυτά δεν είναι δικά του⁴⁶. Αν η τελευταία αυτή ιδιότητα χαθεί, τότε δεν πρόκειται για κατάσταση ενσυναίσθησης, αλλά ταύτισης.⁴⁷ Το τελευταίο, θα προκαλούσε σύγχυση στον σύμβουλο και αδυναμία παροχής ουσιαστικής βοήθειας στον συμβουλευόμενο.

Σύμφωνα με τον McLeod, η ενσυναίσθηση είναι δείγμα της ικανότητας του συμβούλου να εισέρχεται χωρίς προκαταλήψεις στον ιδιωτικό κόσμο της αντίληψης του συμβουλευόμενου και να αποκτά οικειότητα με αυτόν. Με άλλα λόγια, να βιώνει προσωρινά όσα βιώνει κι εκείνος στη ζωή του, να κινείται μέσα στη σκέψη του χωρίς να κρίνει, να αισθάνεται έννοιες για τις οποίες ελάχιστα γνωρίζει και να κοιτάζει άφοβα όσα εκείνος φοβάται να δει. Έτσι, ο σύμβουλος διαδραματίζει το ρόλο του συντρόφου σε ένα ταξίδι εσωτερικό και καθοδηγεί, χωρίς να αποκαλύπτει νωρίτερα όσα διακρίνει. Αυτό άλλωστε, μας λέει ο McLeod, θα ήταν ιδιαίτερα απειλητικό.⁴⁸

Το πλαίσιο μέσα στο οποίο η ενσυναίσθηση αναπτύσσεται προϋποθέτει την προσεκτική παρακολούθηση και την ενεργητική ακρόαση από τον σύμβουλο καθ' όλη τη διάρκεια της συμβουλευτικής διαδικασίας. Η αίσθηση ότι τον ακούν με προσοχή, που διαφαίνεται είτε σε λεκτικό είτε σε μη λεκτικό επίπεδο επικοινωνίας, παρέχει στο συμβουλευόμενο την ενθάρρυνση και την απαραίτητη αίσθηση ασφάλειας, ώστε να ξεδιπλώσει τις σκέψεις και τα συναισθήματά του. Η προσοχή αυτή φανερώνεται μέσα από σημάδια συμπεριφοράς του συμβούλου, όπως είναι η στάση του σώματος (εστιασμένη, αλλά όχι απειλητική), η οπτική επαφή (σταθερή, αλλά όχι επίμονη), οι εκφράσεις του προσώπου (ως μη λεκτική συμμετοχή), η λεκτική ακολουθία ή και η σιωπή (στο πλαίσιο συντονισμού της συζήτησης), όλα στο κατάλληλο μέτρο και την κατάλληλη στιγμή της συνεδρίας.⁴⁹

Οι δεξιότητες που προαναφέρθηκαν, ως αποτέλεσμα εκπαίδευσης και επαγγελματικού προσανατολισμού, είναι σημαντικές. Ωστόσο, δεν είναι αυτές που τελικά θα καθορίσουν την επιθυμητή ποιότητα και αποτελεσματικότητα της θεραπευτικής σχέσης. Η ποιοτική στάθμη, υποστηρίζει ο McLeod, σχετίζεται με μια πτυχή του επαγγελματία, πιο προσωπική και συνάμα επαγγελματική. Αυτή αφορά στο προσωπικό

⁴⁶ Από πολλούς ερευνητές στο πεδίο της Ψυχολογίας, η ενσυναίσθηση αντιμετωπίζεται ως ένα πολυδιάστατο χαρακτηριστικό, λ.χ. βλέπε: Mark H. Davis, "A Multidimensional Approach to Individual Differences in Empathy," *JSAS Catalog of Selected Documents in Psychology* 10, p. 85 (Jan., 1980): 1-19, https://www.uv.es/~friasnav/Davis_1980.pdf; Mark H. Davis, "Measuring individual differences in empathy: Evidence for a multidimensional approach," *Journal of Personality and Social Psychology* 44, no. 1 (Jan., 1983): 113-126, https://www.researchgate.net/publication/215605832_Measuring_individual_differences_in_empathy_Evidence_for_a_multidimensional_approach. Από κάποιους άλλους, η ενσυναίσθηση αντιμετωπίζεται ως μονοδιάστατο χαρακτηριστικό. Βλ. Nathan R. Spreng, Margaret C. McKinnon, Raymond A. Mar, Brian Levine, "The Toronto Empathy Questionnaire: scale development and initial validation of a factor-analytic solution to multiple empathy measures," *Journal of Personality Assessment* 91, no. 1 (Jan., 2009): 62-71, <https://pubmed.ncbi.nlm.nih.gov/19085285/>. Για μία επισκόπηση των τρόπων μέτρησης της ενσυναίσθησης στο πεδίο της Πειραματικής Ψυχολογίας, βλ. Ioanna Zioga, George Kosteletos, Evangelos D. Protopapadakis, Christos Papageorgiou, Konstantinos Kontoangelos, Charalabos Papageorgiou, "The Fox and the Lion: Investigating Associations between Empathy and Emotion Perspective-Taking in Aesop's Fables," *Psychology* 13, no. 4 (April, 2022): 482-513, <https://doi.org/10.4236/psych.2022.134033>.

⁴⁷ John McLeod, *The Counsellor's Workbook: Developing a Personal Approach* (England: Open University Press: The McGraw-Hill Companies, 2nd ed, 2010), 83.

⁴⁸ *Αυτόθι*, 83.

⁴⁹ Μαρία Μαλικιώση - Λοϊζου, *Η συμβουλευτική ψυχολογία στην εκπαίδευση. Από την θεωρία στην πράξη* (Αθήνα: Ελληνικά Γράμματα, 2008), 126.

του σύστημα αξιών και στον κώδικα ηθικής, που προσδιορίζουν την ταυτότητα του ως σύμβουλο και τον τρόπο αλληλεπίδρασής του με τους ανθρώπους.⁵⁰

Η σπουδαιότητα αυτή, συνεχίζει ο McLeod, έγκειται στο ότι η συμβουλευτική διαδικασία αποτελεί μια ηθικά ευαίσθητη προσπάθεια. Στη διάρκεια αυτής, ο συμβουλευόμενος καλείται να κάνει επιλογές και είναι ζωτικής σημασίας, ο σύμβουλος να γνωρίζει τις αξίες που ο ίδιος εκπροσωπεί και να έχει αποσαφηνίσει τις έννοιες και τις αρχές στις οποίες βασίζεται η καθοδήγηση που παρέχει. Η έλλειψη σταθερότητας στην ηθική στάση που τηρεί, δημιουργεί αδυναμία στη μετάδοση της σημασίας ζωτικών εννοιών ή και στην κατανόηση της ηθικής του ανθρώπου που έχει απέναντί του. Η προσωπική φιλοσοφία λοιπόν, που αντικατοπτρίζεται σε όλα τα κομμάτια της θεραπείας, καθορίζει την αποτελεσματικότητά της και καθιστά τη θεραπευτική σχέση μοναδική.⁵¹

Η συμβουλευτική υποστήριξη δύναται να παρέχεται δια ζώσης, επώνυμα και κατόπιν συνεννόησης μεταξύ σύμβουλου και συμβουλευόμενου. Ωστόσο, τις τελευταίες δεκαετίες η είσοδος της τεχνολογίας και στο πεδίο αυτό, έχει δώσει τη δυνατότητα της εξ αποστάσεως συμβουλευτικής υποστήριξης, που παρέχεται μέσω ηλεκτρονικής αλληλογραφίας (e-mail), διαδικτυακής συνομιλίας (online chat), τηλεδιάσκεψης (video conference) κ.ο.κ. Ταυτόχρονα, οι τηλεφωνικές γραμμές βοήθειας, που έχουν στελεχωθεί με το κατάλληλο προσωπικό και αναλόγως με το θέμα προς συζήτηση, έχουν δώσει τη δυνατότητα αναζήτησης κατευθύνσεων σε πραγματικό χρόνο, ανώνυμα και χωρίς προγραμματισμό, συχνά σε εικοσιτετράωρη βάση.

Στη συζήτηση αυτή, περί των τρόπων πραγμάτωσης της συμβουλευτικής διαδικασίας, εισέρχεται ολοένα και πιο δυναμικά η Τεχνητή Νοημοσύνη Τ.Ν. Εξετάζοντας τις ραγδαίες εξελίξεις και τις σχετικές έρευνες των τελευταίων ετών, διαπιστώνεται η ήδη υπάρχουσα αποτελεσματικότητα της χρήσης της στον τομέα της ψυχικής υγείας γενικότερα, ενώ η συνεχής βελτίωση των αλγορίθμων υπόσχεται πολλά παραπάνω σε ότι αφορά τα στάδια της πρόληψης, της διάγνωσης, της θεραπείας και της ευρύτερης βελτίωσης στην εξατομικευμένη φροντίδα των ασθενών.

Ωστόσο, εγείρονται σημαντικά ερωτήματα γύρω από το εάν τα συστήματα της Τ.Ν. θα μπορέσουν ποτέ ουσιαστικά να ευθυγραμμιστούν με τα χαρακτηριστικά εκείνα της Συμβουλευτικής που την καθιστούν μοναδική. Αυτό, γιατί στο πλαίσιο της συμβουλευτικής διαδικασίας, οι χειρισμοί ενός συμβούλου υποκινούνται από τις γνώσεις και δεξιότητες που έχει αποκτήσει μέσω της εκπαίδευσης, κυρίως όμως από σημαντικές αξίες που διέπουν την ίδια τη διαδικασία και την προσέγγισή του στο πλαίσιο της ανθρώπινης σχέσης που αναπτύσσεται. Έννοιες ζωτικές, όπως η αποδοχή, η ενσυναίσθηση, η αυθεντικότητα ή και η ηθική, χτίζουν τη θεραπευτική σχέση, η οποία στη συνέχεια θα αποτελέσει το πρότυπο συσχέτισης με τον κόσμο. Εντός αυτής και της ασφάλειας με την οποία περικλείει τον ευάλωτο ψυχισμό του συμβουλευόμενου, θα επέλθει ο επαναπροσδιορισμός του εαυτού και η αλλαγή στον τρόπο θέασης του κόσμου.

Ίσως η προσαρμοστικότητα των συστημάτων Τ.Ν. και οι πολλές δυνατότητες που διαθέτουν, να επιτύχουν εύκολα την προσομοίωση της διαδικασίας σε επίπεδο πρακτικό. Ωστόσο, τίθεται το ερώτημα του εάν η απλή αναπαραγωγή κάποιων αρχών θα μπορέσει να εκπληρώσει, επί της ουσίας, προϋποθέσεις όπως η ικανότητα του *βουλεύεσθαι*, η αυθεντικότητα, ο σεβασμός ή η ελευθερία. Με άλλα λόγια, είναι άξιο

⁵⁰ McLeod, *The Counsellor's Workbook*, 151-152.

⁵¹ *Αυτόθι*, 151-152.

διερεύνησης το εάν η Τ.Ν. θα καταφέρει, είτε ως εργαλείο στα χέρια ενός ανθρώπινου θεραπευτή να διατηρήσει αναλλοίωτη τη θεραπευτική σχέση, είτε ως αυτόνομο σύστημα να θεμελιώσει η ίδια μια θεραπευτική σχέση, ικανή να αποτελέσει το εφελτήριο της αλλαγής.

Τα ερωτήματα αυτά θα μας απασχολήσουν έντονα στις σελίδες που ακολουθούν, ξεκινώντας από τους τρόπους με τους οποίους αυτή τη στιγμή, τα συστήματα της Τ.Ν. εφαρμόζονται στο γνωστικό πεδίο της Συμβουλευτικής. Οι τρόποι αυτοί παρουσιάζονται στην ενότητα που ακολουθεί.

1.γ. Η ενσωμάτωση της Τ.Ν. στο πεδίο της Συμβουλευτικής σήμερα

Είναι αλήθεια πως η κατανόηση του τρόπου με τον οποίο η Τ.Ν. συμβάλλει στην πραγμάτωση της συμβουλευτικής διαδικασίας δεν θα μπορούσε να μελετηθεί ξέχωρα από τη γενικότερη ενσωμάτωση της στον τομέα της ψυχικής υγείας. Τα τελευταία χρόνια ένα νέο πεδίο, αυτό της ψηφιακής ψυχικής υγείας, βελτιώνεται συνεχώς και ήδη, ιατρικά ιδρύματα σε όλο τον κόσμο αναπτύσσουν αλγόριθμους Τ.Ν. για τα τμήματά τους, ενόσω οι κυβερνήσεις των κρατών επενδύουν μεγάλα χρηματικά ποσά στο όλο εγχείρημα.⁵²

Πιο συγκεκριμένα, στον τομέα της ψυχικής υγείας τα συστήματα της Τ.Ν. στοχεύουν στην προσομοίωση της ανθρώπινης ικανότητας κατά την ανάλυση και την κατανόηση σύνθετων ιατρικών δεδομένων. Αυτό που τα διαφοροποιεί από τις παραδοσιακές τεχνολογίες, είναι η ικανότητα συλλογής και επεξεργασίας μεγάλου όγκου δεδομένων. Οι αλγόριθμοι που χρησιμοποιούνται είναι αυτοί της μηχανικής μάθησης, που έχουν τη δυνατότητα να εκπαιδεύονται από τα δεδομένα αυτά και οι αλγόριθμοι βαθιάς μάθησης, που μπορούν να τα επεξεργάζονται και να εκτελούν εργασίες με βάση μια συγκεκριμένη εντολή που έχει οριστεί από το χρήστη. Έτσι, οι αλγόριθμοι είναι σε θέση να αναγνωρίζουν μοτίβα συμπεριφοράς και δημιουργώντας τη δική τους λογική, να ανταποκρίνονται σε έναν συγκεκριμένο στόχο.⁵³

Οι κύριοι στόχοι γύρω από τους οποίους κινούνται τα συστήματα της Τ.Ν. είναι η πρόληψη, η ανίχνευση στοιχείων συμπεριφοράς που δηλώνουν διαταραχές στον ψυχισμό, η πρόβλεψη, η διάγνωση, η θεραπεία και η ευρύτερη βελτίωση της εξατομικευμένης φροντίδας ψυχικής υγείας. Για την επίτευξη των σκοπών αυτών, η Τ.Ν. εφαρμόζεται με τρεις, κυρίως, τρόπους. Αυτοί, είναι τα μοντέλα προσωπικής ανίχνευσης ή αλλιώς, ο ψηφιακός φαινότυπος (*Personal sensing or digital phenotyping*), ο νευρογλωσσικός προγραμματισμός NLP ή αλλιώς, η επεξεργασία της φυσικής γλώσσας κλινικών κειμένων και περιεχομένου των μέσων κοινωνικής δικτύωσης (*Natural language processing of clinical texts and social media content - NLP*) και επιπλέον, τα *chatbot*, λογισμικά συστήματα, που όντας ικανά να μιμηθούν την ανθρώπινη συνομιλία, αναπαράγουν την αλληλεπίδραση μέσω γραπτού ή προφορικού λόγου.⁵⁴

⁵² Simon D'Alfonso, "AI in mental health," *Current Opinion in Psychology* 36 (June 2020): 112, <https://doi.org/10.1016/j.copsyc.2020.04.005>; Ζηκόπουλος, «Η χρήση της τεχνητής νοημοσύνης», 21-22.

⁵³ Ζηκόπουλος, «Η χρήση της τεχνητής νοημοσύνης», 21-22.

⁵⁴ D'Alfonso, "AI in mental health," 112.

Πιο συγκεκριμένα, ο ψηφιακός φαινότυπος λειτουργεί μέσω της επεξεργασίας των δεδομένων (*δεδομένα αισθητήρων και χρήσης*), που συλλέγονται από τις προσωπικές ψηφιακές συσκευές και της μετέπειτα εξαγωγής πληροφοριών από αυτά, σχετικά με την ανθρώπινη συμπεριφορά.⁵⁵ Το σκεπτικό πίσω από τη δημιουργία του είναι η υπόθεση πως, τα ψηφιακά αποτυπώματα ενός ατόμου μπορούν να μας δώσουν σημαντικές πληροφορίες σχετικά με τον ψυχισμό του και τους τρόπους που αυτός μεταβάλλεται.⁵⁶

Η πιο χρήσιμη από τις συσκευές, που όλοι χρησιμοποιούμε, είναι το κινητό smartphone, με το 60% περίπου του παγκόσμιου πληθυσμού να πραγματοποιεί καθημερινή ενεργή χρήση και να έχει, μέσω αυτού, πρόσβαση στο διαδίκτυο.⁵⁷ Η μεθοδολογία των σχετικών ερευνών που διεξάγονται, περιλαμβάνει τη συλλογή δεδομένων μέσω ενσωματωμένων αισθητήρων στα κινητά smartphone, όπως το επιταχυνσιόμετρο, το GPS, το γυροσκόπιο κ.ο.κ., τα οποία στη συνέχεια, υποβάλλονται σε επεξεργασία από αλγόριθμους μηχανικής μάθησης (*αλγόριθμοι παλινδρόμησης, αλγόριθμοι ταξινόμησης και αλγόριθμοι μηχανής διανύσματος υποστήριξης*), που ανιχνεύουν στοιχεία των ψυχικών διαταραχών, βάσει των κλιμάκων που τις ορίζουν (λ.χ. *PHQ-9*⁵⁸) και εκτιμούν το βαθμό της σοβαρότητάς τους.⁵⁹

Έχοντας ενσωματωμένους αισθητήρες, ένα smartphone δίνει τη δυνατότητα συσχετίσεων ανάμεσα στις διάφορες μεταβλητές και την προσωπικότητα του ατόμου, οδηγώντας έτσι, σε συμπεράσματα γύρω από τις ατομικές διαφορές στα ψυχολογικά χαρακτηριστικά. Έτσι, η μελέτη των μεταβλητών της επικοινωνιακής συμπεριφοράς μπορεί να διαθέσει σημαντικές πληροφορίες σχετικά με την κοινωνική συμπεριφορά και δικτύωση του χρήστη. Για παράδειγμα, σχετική έρευνα των Montag et al. έδειξε ότι οι μεταβλητές κλήσης συνδέονται με την εξωστρέφεια, ενώ η μείωση του αριθμού αυτών, όπως και των γραπτών μηνυμάτων που αποστέλλονται, συσχετίζονται με υποτροπές σχιζοφρένειας.⁶⁰ Παρόμοια έρευνα έδειξε πως η δραστηριότητα πληκτρολόγησης στην οθόνη αφής (π.χ. κύλιση, διπλό κλικ) ή και το μοτίβο του κωδικού εισόδου, μπορεί να παρέχει ενδείξεις ψυχικής υγείας (λ.χ. καταθλιπτική συμπεριφορά), χωρίς την ανάγκη γλωσσικής ανάλυσης του περιεχομένου των χαρακτήρων που εισάγονται.⁶¹ Το ίδιο μπορεί να συμβεί και βάσει δεδομένων για το περιβάλλον του χρήστη (π.χ. τα επίπεδα του θορύβου, της θερμοκρασίας και του

⁵⁵ *Αυτόθι*.

⁵⁶ Christian Montag, Cornelia Sindermann, Harald Baumeister, “Digital phenotyping in psychological and medical sciences: a reflection about necessary prerequisites to reduce harm and increase benefits,” *Current Opinion in Psychology* 36 (Apr. 2020): 19, <https://doi.org/10.1016/j.copsyc.2020.03.013>.

⁵⁷ *Αυτόθι*, 19.

⁵⁸ Πρόκειται για την κλίμακα κατάθλιψης που χρησιμοποιείται στο *Nine-Item Patient Health Questionnaire [PHQ-9]*, ένα ερωτηματολόγιο που αναπτύχθηκε το 1999 από τον Spitzer. Βλ. σχετικώς: “Patient Health Questionnaire [PHQ-9],” *DatAnalysis*, accessed August 23, 2023, <https://datanalysis.net/questionnaires/patient-health-questionnaire-phq/>; «Online Ερωτηματολόγιο Κατάθλιψης PHQ-9», *Psychiatros.gr*, προσπελάστηκε στις 23 Αυγούστου 2023, <https://www.psychiatros.gr/online-%ce%b5%cf%81%cf%89%cf%84%ce%b7%ce%bc%ce%b1%cf%84%ce%bf%ce%bb%cf%8c%ce%b3%ce%b9%ce%bf-%ce%ba%ce%b1%cf%84%ce%bb%ce%b9%cf%88%ce%b7%cf%82-phq-9/>.

⁵⁹ Βλ. Mohammed T. Masud, Mohammed A. Mamun, K. Thapa, D.H. Lee, Mark D. Griffiths, S.H. Yang, “Unobtrusive monitoring of behavior and movement patterns to detect clinical depression severity level via smartphone,” *Journal of Biomedical Informatics* 103, no. 103371 (March 2020): 3, <https://doi.org/10.1016/j.jbi.2019.103371>.

⁶⁰ *Αυτόθι*, 20; D’Alfonso, “AI in mental health,” 112.

⁶¹ D’Alfonso, “AI in mental health,” 112.

φωτισμού) ή το παγκόσμιο σήμα θέσης (GPS) που μπορεί να συνδεθεί με την κατάθλιψη, εφόσον εντοπίσει πως ένας χρήστης παραμένει τις περισσότερες ώρες εντός της οικίας.⁶²

Πέρα από τα παραπάνω, τα smartphones αποτελούν ένα αποτελεσματικό μέσο για την εφαρμογή της *Οικολογικής Στιγμιαίας Αξιολόγησης (Ecological Momentary Assessment - EMA)*, που αφορά στην στιγμιαία παρακολούθηση της ψυχικής υγείας, μέσα από την αποστολή ερωτηματολογίων στο χρήστη για την επιτόπου αξιολόγηση του πλαισίου στο οποίο βρίσκεται, της συμπεριφοράς και της ψυχολογικής του διάθεσης. Η μέθοδος αυτή επιτρέπει την παρακολούθηση του ατόμου σε πραγματικό χρόνο, μέσα από τις σκέψεις και τα συναισθήματα που εκφράζει. Επιπλέον, πέρα από τις απλές αξιολογήσεις, η εφαρμογή της *Οικολογικής Στιγμιαίας Παρέμβασης (Ecological Momentary Interventions - EMI)* επιτρέπει στιγμιαίες ψυχολογικές παρεμβάσεις ή προτροπές συμπεριφοράς στον χρήστη μέσω της κινητής του συσκευής, συχνά διαμορφωμένες βάσει των απαντήσεων που έδωσε στα προαναφερόμενα ερωτηματολόγια.⁶³

Με παρόμοιους τρόπους, τα ψηφιακά αποτυπώματα στα μέσα κοινωνικής δικτύωσης, που αφορούν σε δραστηριότητες όπως τα σχόλια, τα Likes, οι ενημερώσεις κατάστασης και ο αριθμός φωτογραφιών που δημοσιεύονται, χρησιμοποιούνται για την εξαγωγή πληροφοριών γύρω από την ψυχολογική κατάσταση του χρήστη και την πιθανή μελλοντική του συμπεριφορά. Επιπλέον, όπως υποστηρίζει η πρόσφατη μελέτη των Eichstaedt et al., η ανάλυση των δεδομένων στα μέσα κοινωνικής δικτύωσης δύναται να χρησιμοποιηθεί ως σύστημα έγκαιρης ανίχνευσης της συναισθηματικής διαταραχής. Όπως και στην περίπτωση των smartphones, οι αλγόριθμοι της T.N. επεξεργάζονται τα δεδομένα που συλλέγονται, με τρόπους όπως η συσχέτιση της φυσικής δραστηριότητας με τα χαρακτηριστικά των ψυχικών διαταραχών, η καταγραφή προτύπων συμπεριφοράς ή η ανάλυση του γλωσσικού περιεχομένου, εξάγοντας συμπεράσματα για την ψυχική κατάσταση του χρήστη.⁶⁴

Με το παραπάνω σκεπτικό, τα τελευταία χρόνια, οι εταιρείες τεχνολογίας σπεύδουν να εισέλθουν στον τομέα του ψηφιακού φαινότυπου. Χαρακτηριστικά, το Facebook ανακοίνωσε τη χρήση της T.N. για την ανίχνευση ενδείξεων αυτοκτονικού ιδεασμού στις αναρτήσεις και τα ζωντανά βίντεο των χρηστών, ειδοποιώντας, σε περίπτωση εντοπισμού, τις αρμόδιες αρχές ή παρακινώντας τον χρήστη στην αναζήτηση εξειδικευμένης βοήθειας. Ομοίως, η εταιρεία Mindstrong Health στην Καλιφόρνια, έχει αναπτύξει πλατφόρμα παρακολούθησης των συνθηκών των χρηστών κινητών τηλεφώνων, με στόχο την ανίχνευση συμπεριφορών που φανερώνουν την ύπαρξη κατάθλιψης, ενώ η Sharecare στην Ατλάντα, διαθέτει εφαρμογή ανάλυσης των επιπέδων στρες των χρηστών κατά τη διάρκεια των κλήσεων. Μάλιστα η τελευταία, μετά τον τερματισμό της κλήσης αποστέλλει μήνυμα στο χρήστη, περιγράφοντάς του την εντύπωση που αποκόμισε βάσει της καταγραφής (π.χ. «φαινόσουν αγχωμένος»)⁶⁵.

Σε συνδυασμό με τα παραπάνω, οι έρευνες γύρω από τον ψηφιακό φαινότυπο στοχεύουν στην ενσωμάτωση συστημάτων της T.N. που θα μπορούν να παρέχουν συναφή και εξατομικευμένη φροντίδα, μέσα από την παροχή πληροφοριών σε ασθενείς

⁶² Montag, Sindermann, Baumeister, “Digital phenotyping” 20.

⁶³ D’Alfonso, “AI in mental health,” 113.

⁶⁴ Montag, Sindermann, Baumeister, “Digital phenotyping” 19.

⁶⁵ Natasha Singer, “How companies scour our digital lives for clues to our health,” *New York Times*, July 10, 2018, <https://www.nytimes.com/2018/02/25/technology/smartphones-mental-health.html>.

ή/και επαγγελματίες, σχετικά με την εξέλιξη της ψυχικής υγείας. Για παράδειγμα, αυτοματοποιημένες ειδοποιήσεις στην περίπτωση επιδείνωσης των συμπτωμάτων ως ένδειξη κινδύνου υποτροπής ή πληροφορίες που φανερώνουν βελτίωση, όπως η αύξηση της επικοινωνιακής δραστηριότητας του χρήστη που αναρρώνει από ένα επεισόδιο κατάθλιψης ή ο πιο ενεργός τρόπος ζωής που αντικατοπτρίζεται στο σήμα GPS.⁶⁶ Τέτοιου είδους συστήματα, όσο περισσότερα θα εξοικειώνονται με το συγκεκριμένο χρήστη, τόσο θα εξελίσσονται και θα βελτιώνουν τις αποκρίσεις τους με την πάροδο του χρόνου.⁶⁷

Πέρα βέβαια, από τα smartphones και τα social media, στο επίκεντρο των ερευνών βρίσκονται και άλλες συσκευές, όπως τα έξυπνα ρολόγια (smartwatches), που ερευνούν τη σχέση μεταξύ της ψυχικής ασθένειας και της σωματικής δραστηριότητας που καταγράφεται από αισθητήρες γεωεντοπισμού και επιτάχυνσης.⁶⁸ Γενικότερα, από τη δεκαετία του 1990, έχει αναδυθεί ο όρος *Διαδίκτυο των Πραγμάτων* (*Internet of Things - IoT*), που περιλαμβάνει ένα δίκτυο επικοινωνίας μεταξύ των διαφορετικών ψηφιακών συσκευών (οικιακών, αυτοκινήτου κ.ο.κ.), που συνδεδεμένες μεταξύ τους ανταλλάσσουν δεδομένα και επιτρέπουν τη συνεχή και σε πραγματικό χρόνο εξαγωγή πληροφοριών σχετικά με τη σωματική και ψυχική υγεία ενός ατόμου και τις σχετικές συμπεριφορές του.⁶⁹

Η τεχνολογία IoT υιοθετείται με ταχείς ρυθμούς από τον κλάδο της υγειονομικής περίθαλψης παγκοσμίως. Τα πορίσματα των σχετικών ερευνών των τελευταίων ετών, επισημαίνουν πως τα ψηφιακά αποτυπώματα από πηγές όπως το αυτοκίνητο, το ντους, ή και το πληκτρολόγιο του υπολογιστή, συμβάλλουν στην αποτελεσματική παρακολούθηση των συνθηκών υγείας του ασθενούς, στην ανίχνευση συμπτωμάτων, στην παρακολούθηση της εξέλιξης της νόσου και στη διευκόλυνση των γιατρών στη διαχείριση της ιατρικής θεραπείας.⁷⁰ Προτάσεις, όπως η δημιουργία συστημάτων απομακρυσμένης παρακολούθησης των ζωτικών σημείων σε πραγματικό χρόνο μέσα από φορητούς αισθητήρες, ανίχνευσης ορισμένων παραμέτρων των εγκεφαλικών κυμάτων και μια ποικιλία εφαρμογών που ενεργοποιούνται από το IoT, αν μη τι άλλο, δημιουργούν έναν θετικό μετασχηματισμό της υγειονομικής περίθαλψης προς τους ψυχικά ασθενείς.⁷¹

Πέρα από τα ψηφιακά αποτυπώματα, η υπόθεση ότι η γλώσσα που χρησιμοποιούμε και ο τρόπος που εκφραζόμαστε συνδέονται με τις ψυχολογικές μας καταστάσεις, αποτέλεσε αφορμή για την εμφάνιση σχετικών ερευνών. Στο πλαίσιο αυτών, μέσα από τα συστήματα επεξεργασίας της φυσικής γλώσσας (Natural language processing - NLP) και της ανάλυσης του ήχου μέσω T.N., οι ερευνητές προέβησαν στην καταγραφή των κλινικών συνεντεύξεων, καθώς και του περιεχομένου των μηνυμάτων που ανταλλάσσονται και των σχολίων που αναρτώνται στα μέσα κοινωνικής δικτύωσης ή

⁶⁶ Montag, Sindermann, Baumeister, “Digital phenotyping” 20.

⁶⁷ D’Alfonso, “AI in mental health,” 113.

⁶⁸ *Αυτόθι*, 112.

⁶⁹ Isabel De la Torre Díez, Susel Góngora Alonso, Sofiane Hamrioui, Eduardo Motta Cruz, Lola Morón Nozaleda, Manuel A. Franco, “IoT-Based Services and Applications for Mental Health in the Literature,” *Journal of Medical Systems* 43, no. 1 (Dec. 2018): 1-2, [10.1007/s10916-018-1130-3](https://doi.org/10.1007/s10916-018-1130-3) ; D’Alfonso, “AI in mental health,” 112.

⁷⁰ De la Torre Díez et al., “IoT-Based Services,” 3; D’Alfonso, “AI in mental health,” 112.

⁷¹ De la Torre Díez et al., “IoT-Based Services,” 1-2.

στα διαδικτυακά φόρουμ. Τα τελευταία, αν και δεν έχουν κλινικό χαρακτήρα, αποτελούν πλούσιες πηγές προς ανάλυση.⁷²

Οι έρευνες έδειξαν πως χαρακτηριστικά στον τρόπο έκφρασης, όπως το φτωχό λεξιλόγιο, η σημασιολογική ασυνέπεια και η μειωμένη συντακτική πολυπλοκότητα είναι δείκτες σοβαρής ψυχικής ασθένειας, ιδιαίτερα της σχιζοφρένειας. Πέρα από τη φυσική γλωσσική επεξεργασία του κειμένου, η ανάλυση του ήχου της φωνής (π.χ. ένταση, ύψος, τονισμός) μπορεί επίσης να αναλυθεί με τρόπο ώστε να παρέχει πληροφορίες σχετικά με την ψυχική υγεία.⁷³

Όπως διαπιστώνεται, η ποσοτικοποίηση των χαρακτηριστικών αυτών μέσω τεχνικών NLP και η μετέπειτα χρήση τους σε μοντέλα μηχανικής μάθησης, δύναται να προβλέψουν με ακρίβεια και νωρίτερα από τις παραδοσιακές προσεγγίσεις, την εμφάνιση ψυχικής διαταραχής. Αυτό έκανε ερευνητική ομάδα των Bedi et al., που χρησιμοποίησε τη μεταβλητή της σημασιολογικής συνοχής και της συντακτικής πολυπλοκότητας και κατάφερε να προβλέψει με ακρίβεια 100% την μετέπειτα εμφάνιση ψύχωσης.⁷⁴ Η ίδια ομάδα, εφαρμόζοντας την ίδια μέθοδο σε μεγαλύτερο σύνολο δεδομένων,⁷⁵ προέβλεψε με ακρίβεια 83% την έναρξη ψύχωσης και με ακρίβεια 72% κατάφερε να διακρίνει, με βάση την ομιλία, τα υγιή άτομα από εκείνα που εμφάνισαν ψύχωση προσφάτως.⁷⁶

Μέχρι στιγμής, το μεγαλύτερο μέρος των ερευνών έχει επικεντρωθεί στην κατάθλιψη, με τους ερευνητές να επιδιώκουν τον εντοπισμό των συμπεριφορών που την εκδηλώνουν, βάσει του πορίσματος ότι οι γλωσσικοί προγνωστικοί της παράγοντες περιλαμβάνουν συναισθηματικές (λύπη), διαπροσωπικές (μοναξιά, εχθρότητα) και γνωστικές διεργασίες (ενασχόληση με τον εαυτό, μηρυκασμός). Αν και τέτοιου είδους αυτοματοποιημένα εργαλεία δεν κάνουν κάτι περισσότερο από τον μακροχρόνιο συσχετισμό μεταξύ γλωσσικών και ψυχικών διαταραχών, η ανάπτυξη και η συνεχής βελτίωσή τους φαίνεται να προσφέρουν ένα ακριβές και άμεσο μέσο ανίχνευσης και πρόβλεψης.⁷⁷

Παράλληλα με τους τρόπους που μόλις περιγράφηκαν, τα τελευταία χρόνια η χρήση της *Συνομιλητικής Τ.Ν. (Conversational AI - CAI)* αυξάνεται συνεχώς. Τα chatbot συγκεκριμένα, παρουσιάζουν σημαντικές ευκαιρίες στον τομέα της ψυχικής υγείας και ειδικότερα της συμβουλευτικής, ενώ σε συνδυασμό με τις δυνατότητες που προσφέρουν ο ψηφιακός φαινότυπος και η επεξεργασία φυσικής γλώσσας NLP,

⁷² D'Alfonso, "AI in mental health," 113.

⁷³ *Αυτόθι*.

⁷⁴ Βλ. Gillinder Bedi, Facundo Carrillo, Guillermo Cecchi, Diego Fernández Slezak, Mariano Sigman, Natália Mota, Sidarta Ribeiro, Daniel Javitt, Mauro Copelli, Cheryl Corcoran, "Automated analysis of free speech predicts psychosis onset in high-risk youths," *npj Schizophrenia* 1, no. 15030 (Aug. 2015): 1-7, <https://doi.org/10.1038/npjischz.2015.30>.

⁷⁵ Cheryl M. Corcoran, Facundo Carrillo, Diego Fernández-Slezak, Gillinder Bedi, Casimir Klim, Daniel C. Javitt, Carrie E. Bearden, Guillermo A. Cecchi, "Prediction of psychosis across protocols and risk cohorts using automated language analysis," *World Psychiatry* 17, no. 1 (Feb. 2018): 67-75, <https://doi.org/10.1002/wps.20491>.

⁷⁶ D'Alfonso, "AI in mental health," 113.

⁷⁷ *Αυτόθι*.

επιδιώκουν μια ολοκληρωμένη προσέγγιση, καλύπτοντας όλα τα στάδια από την πρόληψη έως και την θεραπεία.⁷⁸

Με τον όρο *chatbot*, εννοούνται τα λογισμικά προγράμματα που μιμούμενα την ανθρώπινη συνομιλία, αλληλοεπιδρούν με τους χρήστες (προφορικά ή γραπτά) σε ένα πλαίσιο επικοινωνίας που κυμαίνεται από απλές απαντήσεις, βασισμένες σε κανόνες και αντιστοιχίσεις λέξεων-κλειδιών, έως εξελιγμένες τεχνικές NLP, που παρέχουν βασικές πληροφορίες γύρω από την ψυχική υγεία και αναπτύσσουν θεραπευτικές τεχνικές.⁷⁹

Τα τελευταία χρόνια, τη λειτουργία των chatbot εμπλουτίζει η *Παραγωγική Τ.Ν.* (*Generative AI*), ένα πεδίο που εισήχθη το 2014 από τον Αμερικανό επιστήμονα της Πληροφορικής, Ian Goodfellow. Με το συνδυασμό αλγορίθμων βαθιάς μάθησης και τη χρήση πολλαπλών τεχνικών κωδικοποίησης, η Παραγωγική Τ.Ν. έχει τη δυνατότητα μετατροπής ακατέργαστων χαρακτήρων σε πρωτότυπο περιεχόμενο (εικόνες, κείμενο, ήχος κ.ο.κ.), ως απόκριση σε μία σχετική προτροπή που δόθηκε σε απλή γλώσσα από τον χρήστη. Οι δυνατότητες χρήσης του συγκεκριμένου πεδίου είναι πολλές και κυμαίνονται από την απλή εξυπηρέτηση των πελατών μιας εταιρείας μέχρι και προτάσεις νέων φαρμακευτικών ενώσεων προς δοκιμή. Η δε ευκολία στη χρήση καθώς και πρωτοποριακές ανακαλύψεις όπως το chatGPT, καθιστούν την Παραγωγική Τ.Ν. ολοένα και πιο προσιτή για τους χρήστες ανά τον κόσμο.⁸⁰

Επιπλέον, από το 2018 έχει εισαχθεί ο όρος *Large Language Model (LLM)*, ένας τύπος αλγορίθμου της Τ.Ν., που χρησιμοποιώντας τη μέθοδο της βαθιάς μάθησης και αναλύοντας μεγάλους όγκους δεδομένων, επιτυγχάνει την παραγωγή κειμένου που προσομοιάζει τον ανθρώπινο τρόπο ομιλίας. Τα *LLM* προβαίνουν στη συλλογή και επεξεργασία δεδομένων, καθώς και στη μετέπειτα χρήση αυτών, για την εκπαίδευση των αλγορίθμων και των νευρωνικών δικτύων στην πρόβλεψη των λέξεων που θα απαρτίσουν μία πρόταση του χρήστη. Επιπλέον, μέσα από την επεξεργασία φυσικής γλώσσας NLP, δύναται να δημιουργήσουν ή να μεταφράσουν ένα κείμενο, να συνομιλήσουν με το χρήστη ή ακόμη και να αναλύσουν τα συναισθήματά του.⁸¹

Σε ότι αφορά τον τομέα της ψυχικής υγείας και τα θεραπευτικά chatbot, σκοπός αυτών είναι να αναγνωρίσουν πρότυπα συμπεριφοράς και να παρέχουν συμβουλευτική υποστήριξη ψυχικής υγείας, εκπαιδύοντας τα άτομα σε νέες τεχνικές και δεξιότητες που θα εφαρμόσουν στην καθημερινή ζωή. Παράλληλα, ο εντοπισμός διαταραχών της ψυχικής υγείας σε πρώιμα στάδια (με τη βοήθεια του ψηφιακού φαινοτύπου) δύναται να συντελέσει τόσο στη πρόληψη, όσο και στην καλύτερη ποιότητας της φροντίδας μέσα από την εξατομίκευση της θεραπείας.⁸² Η μεγαλύτερη δυνατοτήτά των chatbot έγκειται στην ανά πάσα στιγμή διαθεσιμότητά τους για επικοινωνία και στην προσβασιμότητα που παρέχουν σε ευάλωτα άτομα, που για διάφορους λόγους, δεν

⁷⁸ Jana Sedlakova, Manuel Trachsel, “Conversational Artificial Intelligence in Psychotherapy: A New Therapeutic Tool or Agent?” *The American Journal of Bioethics* 23, no. 5 (Apr. 2022): 4, <https://doi.org/10.1080/15265161.2022.2048739>.

⁷⁹ D’Alfonso, “AI in mental health,” 114.

⁸⁰ “What is generative AI? Everything you need to know.”

⁸¹ «Τι Είναι τα Large Language Models και Πώς Λειτουργούν;» Big Blue Data Academy, accessed August 22, 2023, <https://bigblue.academy/gr/large-language-models>.

⁸² Sedlakova, Trachsel, “Conversational Artificial Intelligence in Psychotherapy,” 4.

λαμβάνουν θεραπεία από τις παραδοσιακές υπηρεσίες ψυχικής υγείας (λ.χ. λόγω κινητικής αναπηρίας ή συναισθηματικής δυσφορίας στο γραφείο ενός θεραπευτή).⁸³

Το πρώτο chatbot που γνωρίζουμε είναι το ELIZA (1966), η δημιουργία του οποίου στόχευε στην προσομοίωση ενός Ροτζεριανού ψυχοθεραπευτή.⁸⁴ Έκτοτε, τα παραδείγματα είναι πολλά και δεν εστιάζουν μονάχα στον τομέα της ψυχικής υγείας. Ανάμεσά τους το chatbot PARRY (1972) που προσομοίωσε έναν ασθενή με σχιζοφρένεια, το Jabberwacky (1988) που δημιουργήθηκε για ψυχαγωγικούς σκοπούς, το A.L.I.C.E. (1995) που διέθετε περισσότερες από 41.000 κατηγορίες γνώσεων, το SmarterChild (2001) που απαντούσε σε ερωτήματα σχετικά τον καιρό, τις ειδήσεις της ημέρας κ.ο.κ., η Siri (2011) που έδινε στους χρήστες τη δυνατότητα να πραγματοποιήσουν διάφορες ενέργειες (λ.χ. να ενεργοποιήσουν ένα ξυπνητήρι, ή να κάνουν μία κλήση), μέχρι την εμφάνιση του Chat GPT το 2015, με την αρχική μορφή να εξελίσσεται στο σημερινό Chat GPT- 4, ένα εργαλείο ευέλικτο και προσαρμοστικό, που εισέρχεται ολοένα και περισσότερο στην καθημερινότητα των χρηστών ανά τον κόσμο.⁸⁵

Σήμερα, τρία από τα σημαντικότερα είδη θεραπευτικών chatbot ψυχικής υγείας είναι το Woebot,⁸⁶ το Wysa⁸⁷ και το Tess.⁸⁸ Το Woebot, μέσα από σύντομες συνομιλίες και καθημερινή παρακολούθηση των χρηστών, μπορεί να βοηθήσει στη μείωση των συμπτωμάτων του στρες, της κατάθλιψης και του άγχους. Η θεραπευτική προσέγγιση που εφαρμόζει είναι η *γνωσιακή συμπεριφορική θεραπεία (Cognitive Behavioral Therapy - CBT)*, ενώ σε σχετική μελέτη που διεξήχθη ως προς την αποτελεσματικότητά του, διαπιστώθηκε πως η χρήση του για περίοδο δύο εβδομάδων, είναι αρκετή ώστε να επιφέρει σημαντική μείωση της κατάθλιψης και του άγχους σε περίπου 10 εκατομμύρια φοιτητές των ΗΠΑ, που βιώνουν εξουθενωτικό άγχος και κατάθλιψη.⁸⁹

Παρομοίως, μελέτες για το chatbot Wysa, που απευθύνεται σε καταθλιπτικούς ασθενείς, έδειξαν σημαντική βελτίωση στα συμπτώματα μείζονος κατάθλιψης, ενώ το 67,7% αυτών που το χρησιμοποίησαν, χαρακτήρισε την εμπειρία της εφαρμογής χρήσιμη και ενθαρρυντική.⁹⁰ Επιπλέον, έρευνες σχετικά με το chatbot Tess, που σχεδιάστηκε για να προσφέρει εξατομικευμένες συνομιλίες με βάση τα συναισθήματα και τις ανησυχίες που εκφράζονται για την ψυχική υγεία των συμμετεχόντων, έδειξε βελτιώσεις στην κατάθλιψη και το άγχος σε μια ομάδα φοιτητών κολεγίου, μειώνοντας

⁸³ *Αυτόθι*; D'Alfonso, "AI in mental health," 114.

⁸⁴ D'Alfonso, "AI in mental health," 114.

⁸⁵ Ευάγγελος Φωτίου, «Η χρήση του Chat GPT – Προκλήσεις & ευκαιρίες: Εφαρμογές στην εκπαίδευση, στις επιχειρήσεις και στην κοινωνία» (Διπλωματική διατριβή, Πανεπιστήμιο Μακεδονίας & Δημοκρίτειο Πανεπιστήμιο Θράκης, 2023), 12-14.

⁸⁶ "Adult Mental Health," Woebot Health, accessed July 9, 2023, <https://woebothealth.com/adult-mental-health/>.

⁸⁷ "Mental health that meets people where they are," Wysa, accessed July 9, 2023, <https://www.wysa.io/>.

⁸⁸ "Bringing mental health to everyone," Cass, accessed July 9, 2023, <https://www.x2ai.com/>.

⁸⁹ Kathleen Kara Fitzpatrick, Alison Darcy, Molly Vierhile, "Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial," *JMIR Mental Health* 4, no. 2 (Apr.-Jun. 2017): 1,9, doi: [10.2196/mental.7785](https://doi.org/10.2196/mental.7785).

⁹⁰ Becky Inkster, Shubhankar Sarda, Vinod Subramanian, "An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study," *JMIR Mhealth Uhealth* 6, no. 11 (Nov. 2018): 1 & 10, doi: [10.2196/12106](https://doi.org/10.2196/12106).

σημαντικά τα συμπτώματα, μετά από απεριόριστη πρόσβαση στην εφαρμογή για διάστημα δύο έως τεσσάρων εβδομάδων.⁹¹

Αν και τα chatbot δύναται να λειτουργήσουν χωρίς την ανθρώπινη παρέμβαση, στην ουσία οι αποκρίσεις τους βασίζονται σε προγραμματισμένους κανόνες και είναι αποτέλεσμα συνεχούς εκπαίδευσης των αλγορίθμων μηχανικής και βαθιάς μάθησης που χρησιμοποιούνται. Πάνω σε αυτό, οι προσπάθειες βελτίωσης τους εντείνονται συνεχώς, ενώ γενικότερα η δημιουργία περισσότερων διαδικτυακών chatbot που θα παρέχουν συμβουλευτική υποστήριξη μέσω Τ.Ν. αποτελεί στόχο. Σε αυτό, μεταξύ άλλων, συντελεί η διαρκώς αυξανόμενη ανάγκη των ανθρώπων για καθοδήγηση από ειδικό και χωρίς αμφιβολία, οι δυνατότητες των εργαλείων που αναλύσαμε είναι πράγματι ελπιδοφόρες.

Αυτό γιατί, στον τομέα της υγείας, η αποτελεσματικότητα της θεραπείας προϋποθέτει την ακριβή και σωστή αξιολόγηση της κατάστασης ενός ατόμου και στην ψυχική υγεία συγκεκριμένα, αυτό εξαρτάται εξ ολοκλήρου από την ικανότητα του ατόμου να περιγράψει με σαφήνεια τα όσα βιώνει.⁹² Το ίδιο λοιπόν ισχύει και για την Συμβουλευτική που συχνά, αποτελεί το πρώτο βήμα στην αναζήτηση βοήθειας. Εκεί, η λανθασμένη ή ελλιπής εκτίμηση της κατάστασης από τον σύμβουλο (λ.χ. λόγω απειρίας, απόκρυψης πληροφοριών, αδυναμία έκφρασης συναισθημάτων κ.ο.κ.) θα μπορούσε να οδηγήσει σε ακατάλληλους χειρισμούς ή ακόμη και να παραβλέψει στοιχεία συμπεριφοράς που χρήζουν περαιτέρω διερεύνησης ή παραπομπής σε άλλες ειδικότητες (λ.χ. ψυχίατρο). Έτσι, δυνατότητες όπως αυτές που αναλύσαμε, θα μπορούσαν αρχικά να εξασφαλίσουν πως η διάγνωση ή η επιλογή των παρεμβάσεων από τον σύμβουλο είναι οι σωστές, ενώ η συνεχής βελτίωση των αλγορίθμων σε ότι αφορά τους συνολικούς στόχους που θέτει η Τ.Ν. στον τομέα της ψυχικής υγείας θα μπορούσαν μακροπρόθεσμα να διαμορφώσουν ένα πιο προληπτικό και εξατομικευμένο μοντέλο φροντίδας.

Ωστόσο, αν και η πρόοδος των τελευταίων ετών είναι σημαντική, δεν στερείται κινδύνων που θα πρέπει να ληφθούν υπόψη προτού εναποθέσουμε τον ευάλωτο ψυχισμό ενός ατόμου στα εργαλεία της Τ.Ν. Ειδικότερα η CAI (Conversational AI), ως ο τομέας εκείνος της Τ.Ν. που συνδυάζει τις δυνατότητες του ψηφιακού φαινότυπου και της γλωσσικής επεξεργασίας NLP και επεκτείνεται στο στάδιο της θεραπευτικής παρέμβασης, δημιουργεί πρωτόγνωρες προκλήσεις και εγείρει σοβαρά ερωτήματα για το μακροπρόθεσμο αποτύπωμά της στον τομέα της Συμβουλευτικής.

Το θέμα αυτό θα μας απασχολήσει στο επόμενο κεφάλαιο. Αρχικά θα αναφερθούμε στα οφέλη και στους κινδύνους που αποφέρει η ενσωμάτωση της Τ.Ν. στην συμβουλευτική διαδικασία και στη συνέχεια, θα εξετάσουμε τους παράγοντες εκείνους που θα πρέπει να λάβουμε υπόψη, προτού αξιολογήσουμε την αποτελεσματικότητά της. Στη συζήτηση αυτή, θα συμπεριλάβουμε τόσο τα σημερινά δεδομένα, όσο και τις πιθανές μορφές που είναι δυνατό να λάβει μελλοντικά η παροχή της Συμβουλευτικής, υπό την επιρροή της Τ.Ν. και με βάση τα όσα προμηνύουν οι μέχρι τώρα εξελίξεις.

⁹¹ Russel Fulmer, Angela Joerin, Breanna Gentile, Lysanne Lakerink, Michiel Rauws, “Using Psychological Artificial Intelligence (Tess) to Relieve Symptoms of Depression and Anxiety: Randomized Controlled Trial,” *JMIR Ment Health* 5, no. 4 (Oct.-Dec. 2018): 2, DOI: 10.2196/mental.9782.

⁹² Nikolaos Koutsouleris, Tobias Hauser, Vasilisa Skvortsova, Munmun De Choudhury, “From promise to practice: towards the realisation of AI-informed mental health care,” *Lancet Digital Health* 4 (Oct. 2022): 829-840, [https://doi.org/10.1016/S2589-7500\(22\)00153-4](https://doi.org/10.1016/S2589-7500(22)00153-4).

Κεφάλαιο 2^ο: Τα οφέλη και τα επιστημολογικά - μεθοδολογικά προβλήματα της εμπλοκής της Τ.Ν. στην Συμβουλευτική και στην Ψυχική υγεία.

2.α. Τα οφέλη και τα επιστημολογικά - μεθοδικά προβλήματα βάσει των σημερινών δεδομένων.

Το γεγονός πως το διαδίκτυο αποτελεί αναπόσπαστο κομμάτι της καθημερινής ζωής για το μεγαλύτερο μέρος του πληθυσμού, αρκεί για να θεωρήσουμε πως πλέον, ζούμε σε μία κοινωνία υψηλής ψηφιακής σύνδεσης.⁹³ Ο ψηφιακός φαινότυπος, η επεξεργασία της φυσικής γλώσσας NLP και τα chatbot, που αναλύσαμε στο προηγούμενο κεφάλαιο, αποτελούν τα ιδιαίτερα ελπιδοφόρα κομμάτια ενός ευρύτερου επιστημονικού κινήματος που αποκαλείται *Ψυχοπληροφορική*. Σε αυτό το διεπιστημονικό πεδίο που αναπτύσσεται ταχέως, η επιστήμη της Πληροφορικής και η επιστήμη της Ψυχολογίας ενώνουν τις δυνάμεις τους με στόχο την κατανόηση του ανθρώπινου νου.⁹⁴ Είναι γεγονός πως αυτός ο ψηφιακός μετασχηματισμός φέρει σημαντικά οφέλη στο κομμάτι της ψυχικής υγείας. Παράλληλα ωστόσο, εγείρει σκέψεις γύρω από τις αδυναμίες εκείνες των λογισμικών, που θα μπορούσαν να περιορίσουν την ακρίβεια τους καθώς και τις πιθανότητες της κακής χρήσης, που θα μπορούσαν να θέσουν σε κίνδυνο τα ανθρώπινα δικαιώματα.

Έχουμε δει έως τώρα, τον τρόπο με τον οποίο ο ψηφιακός φαινότυπος επιτυγχάνει την εξαγωγή πληροφοριών γύρω από την ψυχική κατάσταση του ατόμου, μέσω της παρακολούθησης και καταγραφής της καθημερινής του συμπεριφοράς και συγκεκριμένα, του τρόπου χρήσης των συσκευών smartphone και της αλληλεπίδρασης του στα μέσα κοινωνικής δικτύωσης. Συμπληρωματικά στις δύο αυτές, βασικές πηγές των δεδομένων, το Διαδίκτυο των Πραγμάτων (Internet of Things - IoT) έρχεται να συντελέσει επιπλέον στην αποτελεσματικότητα του ψηφιακού φαινοτύπου, μέσα από τη διάθεση μεγαλύτερου αριθμού δεδομένων από περιβάλλοντα πλήρως συνδεδεμένα στο Διαδίκτυο. Οι ψυχολογικές μεταβλητές ενός ατόμου μελετώνται από πλήθος ερευνών, ως προς τη συνάφειά τους με τις καθημερινές του συμπεριφορές, όπως αυτές εξελίσσονται σε ένα περιβάλλον, όπου ψηφιακοί βοηθοί όπως η Alexa της Amazon⁹⁵ ή η Siri της Apple⁹⁶ θα μπορούν να ανάβουν τα φώτα ή να κλείνουν της κουρτίνες μέσω φωνητικών εντολών.⁹⁷

Οι παραπάνω μετρήσεις συντελούν στη συλλογή μεγάλου όγκου πληροφοριών, χωρίς την ελάχιστη προσπάθεια από τον ίδιο το χρήστη. Το γεγονός αυτό καθιστά την εφαρμογή τους εύκολη και εφικτή, όχι μόνο για την ανίχνευση και τη διάγνωση των ψυχικών διαταραχών, αλλά και για τη μακροπρόθεσμη παρακολούθηση της πορείας της ψυχικής υγείας του ατόμου και της αποτελεσματικότητας στη θεραπεία.⁹⁸ Επιπλέον, η

⁹³ Montag, Sindermann, Baumeister, “Digital phenotyping,” 19.

⁹⁴ *Αυτόθι*, 21.

⁹⁵ “What is Alexa?” Alexa, accessed August 28, 2023, <https://developer.amazon.com/en-US/alexa>.

⁹⁶ «Τι είναι το Siri και πώς λειτουργεί; Όλα όσα πρέπει να ξέρετε», iPhoneA2, προσπελάστηκε στις 28 Αυγούστου 2023, <https://iphonea2.com/el/como-funciona-siri/>.

⁹⁷ Montag, Sindermann, Baumeister, “Digital phenotyping,” 21.

⁹⁸ Jonathan P. Stange, John Zulueta, Scott A. Langenecker, Kelly A. Ryan, Andrea Piscitello, Jenna Duffecy, Melvin G. McInnis, Pete Nelson, Olusola Ajilore, Alex Leow, “Let your fingers do the talking:

παθητική αυτή μέτρηση της συμπεριφοράς, φέρει σημαντικά πλεονεκτήματα έναντι των αυτο-αναφορών, όπως για παράδειγμα τα ερωτηματολόγια. Αυτό γιατί, διάφορες παράμετροι, όπως η περιορισμένη ικανότητα διορατικότητας ή δυνατότητας του ατόμου να αξιολογήσει αντικειμενικά τη διάθεση και συμπεριφορά του, η ενδεχόμενη τάση του για δηλώσεις που θα είναι, κατά την άποψή του, αποδεκτές και επιθυμητές από τους γύρω (*Social Desirability Bias*)⁹⁹ ή και η ενδεχόμενη ύπαρξη προκαταλήψεων στη γενικότερη θεώρησή του επί των πραγμάτων, μπορεί να επηρεάσουν τις απαντήσεις του και να θέσουν έτσι, υπό αμφισβήτηση την αντικειμενικότητα των πορισμάτων που προκύπτουν μόνο από τα λεγόμενα ενός ασθενή.¹⁰⁰

Μια τέτοια προσπάθεια, ωστόσο, δεν στερείται αδυναμιών που θα πρέπει να ληφθούν σοβαρά υπόψη. Μεταξύ άλλων, η χρονική διάρκεια της καταγραφής που είναι επαρκής για την εξαγωγή των απαραίτητων δεδομένων γύρω από την ψυχολογική κατάσταση του ατόμου, δεν έχει ακόμη διασαφηνιστεί.¹⁰¹ Αυτό είναι σημαντικό εάν λάβουμε υπόψη τους παράγοντες εκείνους που θα μπορούσαν να διακινδυνεύσουν την εγκυρότητα των πορισμάτων. Για παράδειγμα, έχει παρατηρηθεί πως η γνώση της καταγραφής από τον ίδιο το χρήστη μπορεί να επηρεάσει τη συμπεριφορά του, είτε μειώνοντας το χρόνο ενασχόλησης του με την κινητή συσκευή, είτε αλλοιώνοντας τη γενικότερη δράση του για το διάστημα που αυτή καταγράφεται. Έτσι, η μειωμένη χρήση της συσκευής θα μπορούσε να οφείλεται σε διάφορες αιτίες, όπως η αδυναμία προσαρμογής στην όλη διαδικασία ή η επιθυμία κοινωνικής απομόνωσης ως σύμπτωμα βαριάς μορφής κατάθλιψης.¹⁰² Επίσης, οι διακυμάνσεις της δραστηριότητας εντός της ημέρας ή ανάμεσα στις εργάσιμες ημέρες και τα Σαββατοκύριακα, έχουν ως αποτέλεσμα τις διαφοροποιήσεις των αποτελεσμάτων και επομένως, τον κίνδυνο ανακρίβειας στην πρόβλεψη των μελλοντικών συμπεριφορών.¹⁰³ Κατανοούμε λοιπόν, πως ο ορισμός του απαιτούμενου χρονικού διαστήματος της καταγραφής είναι σημαντικός, τόσο για την διασφάλιση ότι ένας σχετικός βαθμός εξοικείωσης με τη διαδικασία της παρακολούθησης έχει επέλθει από την πλευρά του χρήστη, όσο και για το ότι ένα επαρκές χρονικό διάστημα καταγραφής έχει καταφέρει να αιτιολογήσει τις διαφοροποιήσεις στην συμπεριφορά, κατανοώντας τον τρόπο που αυτές σχετίζονται με τον καθημερινότητα του ατόμου.¹⁰⁴

Passive typing instability predicts future mood outcomes,” *Bipolar Disorders* 20, no 3 (March, 2018): 285, <https://doi.org/10.1111/bdi.12637>.

⁹⁹ Ο όρος *Social Desirability Bias* αναφέρεται στις περιπτώσεις εκείνες όπου οι συμμετέχοντες μιας έρευνας, στην προσπάθεια να κερδίσουν την έγκριση και αποδοχή των άλλων, δίνουν απαντήσεις που πιστεύουν πως θα τους κάνουν να φαίνεται καλοί, αποκρύπτοντας έτσι τις αληθινές απόψεις και εμπειρίες τους. Αυτό έχει ως αποτέλεσμα να επηρεάζονται τα αποτελέσματα μελετών που εστιάζουν σε ευαίσθητα προσωπικά θέματα. Βλ. Kassiani Nikolopoulou, “What is Social Desirability Bias? | Definition & Examples,” March 24, 2023 at *Scribbr*, <https://www.scribbr.com/research-bias/social-desirability-bias/>.

¹⁰⁰ Eva-Maria Messner, Rayna Sariyska, Benjamin Mayer, Christian Montag, Christopher Kannen, Andreas Schwerdtfeger, Harald Baumeister, “Insights – Future Implications of Passive Smartphone Sensing in the Therapeutic Context,” *Verhaltenstherapie* 32, no 1 (Aug., 2022): 87, <https://karger.com/ver/article-pdf/32/Suppl.%201/86/3920346/000501951.pdf>.

¹⁰¹ Montag, Sindermann, Baumeister, “Digital phenotyping,” 21-22.

¹⁰² Dimitri Rozgonjuk, Jason C. Levine, Brian J. Hall, Jon D. Elhai, “The association between problematic smartphone use, depression and anxiety symptom severity, and objectively measured smartphone use over one week,” *Computers in Human Behavior* 87 (May, 2018): 15, <https://doi.org/10.1016/j.chb.2018.05.019>.

¹⁰³ Messner et al., “Insights,” 92.

¹⁰⁴ Montag, Sindermann, Baumeister, “Digital phenotyping,” 27.

Πέραν τούτου, ένα ακόμη σημείο στο οποίο θα πρέπει να σταθούμε έχει να κάνει με το ότι ο ακριβής διαχωρισμός των μεταβλητών¹⁰⁵ από τους ερευνητές του ψηφιακού φαινοτύπου, με βάση τη χρησιμότητά τους σε ένα συγκεκριμένο πλαίσιο μελέτης ή τη σχέση τους με την εκάστοτε ψυχική διαταραχή που ερευνάται, φαίνεται να μην έχει ακόμη επιτευχθεί.¹⁰⁶ Το λεγόμενο «Πρόβλημα του Πλαισίου» (*The Frame Problem*),¹⁰⁷ όπως αποκαλείται, αποτελεί πρόκληση για τους ερευνητές της Τ.Ν., οι οποίοι καλούνται να εκπαιδεύσουν τα συστήματα έτσι ώστε, τα δεδομένα που ερευνώνται σε κάθε περίπτωση να ανταποκρίνονται με ευκρίνεια στην εκάστοτε περιοχή μελέτης.¹⁰⁸

Γίνεται αντιληπτό από τα παραπάνω, πως η εξαγωγή συμπερασμάτων και η λήψη αποφάσεων γύρω από την ψυχική υγεία, που θα βασίζονται αποκλειστικά στα ψηφιακά αποτυπώματα ενός ατόμου απαιτεί περαιτέρω έρευνα και δοκιμές, γεγονός ωστόσο που δεν υποβαθμίζει την αξία της συμβολής τους στον τομέα της ψυχικής υγείας. Απεναντίας, ο συνδυασμός του ψηφιακού φαινοτύπου με τις αναφορές των ίδιων των ασθενών, δύναται να αποτελέσει ένα ολοκληρωμένο πλαίσιο προσέγγισης και να δώσει σημαντικές πληροφορίες γύρω από τις υπάρχουσες αποκλίσεις ανάμεσα στον τρόπο με τον οποίο ένα άτομο βλέπει τον εαυτό του και στον τρόπο που στην πραγματικότητα, συμπεριφέρεται καθημερινά.¹⁰⁹

Καθώς εισερχόμαστε στο πεδίο της επεξεργασίας φυσικής γλώσσας NLP (Natural Language Processing - NLP), παρατηρούμε πως οι σκέψεις και οι προβληματισμοί μας δεν διαφέρουν σημαντικά από όσα αναφέρθηκαν για τον ψηφιακό φαινότυπο παραπάνω. Η γλώσσα, που από καιρό θεωρείτο ένα παράθυρο στο μυαλό του ανθρώπου, μπορεί πλέον να αξιοποιηθεί ποσοτικά και να παρέχει αντικειμενικές μετρήσεις και ακριβή συμπεράσματα για τον ανθρώπινο ψυχισμό.¹¹⁰ Το γεγονός πως και εδώ, η συλλογή των δεδομένων δεν απαιτεί τη συμβολή του χρήστη, μας δίνει τη δυνατότητα μακράς διάρκειας για την παρακολούθηση της εξέλιξης της ψυχικής υγείας ενός ατόμου. Τα σημερινά συστήματα είναι ικανά να αναλύσουν, με συνεπή και αμερόληπτο τρόπο, έναν συγκριτικά πιο μεγάλο όγκο γλωσσικών δεδομένων σε σχέση με τον άνθρωπο. Έτσι, οι τεράστιες ποσότητες αδόμητων κειμένων που προκύπτουν από ιατρικά αρχεία μέχρι αναρτήσεις στα μέσα κοινωνικής δικτύωσης, παρέχουν ένα πρόσφορο έδαφος για τις μεθόδους της NLP, που επιτυγχάνουν τη σύλληψη γλωσσικών

¹⁰⁵ Ο όρος *μεταβλητή* στην ψυχομετρία αναφέρεται στις ιδιότητες εκείνες ή αλλιώς στα χαρακτηριστικά των συμμετεχόντων που μπορούν να παρατηρηθούν, να μετρηθούν και να λάβουν ένα σύνολο τιμών. Βλ. Βασιλική Γιαννούλη, *Ψυχοδιαγνωστικά Μέσα. Ενότητα 2: Ψυχομετρία* [πανεπιστημιακές σημειώσεις]. Πανεπιστήμιο Μακεδονίας, Τμήμα Εκπαιδευτικής και Κοινωνικής Πολιτικής, Θεσσαλονίκη, <http://opencourses.uom.gr/assets/site/public/778/623-Psychodiagnostika-Mesa-02-Giannouli.pdf>. Έχουμε αναφερθεί έως τώρα σε μεταβλητές όπως, η επικοινωνιακή συμπεριφορά, οι κοινωνικές επαφές, η σημασιολογική συνοχή και η συντακτική πολυπλοκότητα κατά τη λεκτική έκφραση κ.ο.κ.

¹⁰⁶ Montag, Sindermann, Baumeister, “Digital phenotyping,” 21-22.

¹⁰⁷ Το λεγόμενο «Πρόβλημα του Πλαισίου» αφορά στον περιορισμό των προγραμματιστών να συγκεκριμενοποιήσουν τις πληροφορίες εκείνες με τις οποίες θα τροφοδοτήσουν τα συστήματα της Τ.Ν. και οι οποίες θα είναι συναφείς με την εκτέλεση μιας συγκεκριμένης πράξης, παραβλέποντας αναρίθμητες άλλες πληροφορίες που δεν επηρεάζουν και δεν επηρεάζονται από την πράξη αυτή. Βλ. Murray Shanahan, “The Frame Problem,” *Stanford Encyclopedia of Philosophy*, accessed September 6, 2023, <https://plato.stanford.edu/entries/frame-problem/>.

¹⁰⁸ Montag, Sindermann, Baumeister, “Digital phenotyping,” 21-22.

¹⁰⁹ *Αυτόθι*, 20.

¹¹⁰ D’Alfonso, “AI in mental health,” 112.

χαρακτηριστικών και προβαίνουν στη σύνδεση αυτών με τα συναισθήματα και τις ενδεχόμενες ψυχικές διαταραχές που αυτά εκφράζουν.¹¹¹

Οι βασικότερες προκλήσεις που καλούνται να αντιμετωπίσουν οι ερευνητές της NLP αφορούν κυρίως στην κατανόηση της φυσικής γλώσσας (*Natural Language Understanding - NLU*), και σχετίζονται συγκεκριμένα, με χαρακτηριστικά όπως η ασάφεια στη σύνταξη μιας πρότασης, στο λεξιλόγιό της ή στην απόδοση του νοήματος που αυτή εκφράζει.¹¹² Ειδικότερα, ως προς το σημασιολογικό επίπεδο, ένα σημαντικό χαρακτηριστικό που ερευνάται είναι ο σαρκασμός, η ειρωνία και το χιούμορ, αμφίσημες μορφές επικοινωνίας, κατά τις οποίες το άτομο συχνά εννοεί κάτι κατά πολύ διαφορετικό από αυτό που λέει.¹¹³ Στα παραπάνω, προστίθενται οι αδυναμίες που συναντώνται και κατά την εφαρμογή του ψηφιακού φαινοτύπου. Πιο συγκεκριμένα, στην εξέλιξη της παρούσας έρευνας δεν έχουν εντοπιστεί πληροφορίες σχετικά με το ακριβές διάστημα κατά το οποίο η ανάλυση θα μπορούσε να δώσει τα επιθυμητά αποτελέσματα, ενώ το γεγονός πως ένα πρόγραμμα σχεδιασμένο για έναν συγκεκριμένο τομέα, μπορεί να μην δίνει έγκυρα αποτελέσματα σε κάποιον άλλο, δεν πρέπει να αγνοηθεί.¹¹⁴

Με δεδομένο το ότι η επεξεργασία της φυσικής γλώσσας και ο ψηφιακός φαινότυπος δύνανται να χρησιμοποιηθούν για την ανάπτυξη των συνομιλητικών chatbot, κατανοούμε το ότι τα παραπάνω χαρακτηριστικά επεκτείνονται με τον τρόπο αυτό και στο κομμάτι της Συνομιλητικής Τ.Ν. (Conversational AI). Όπως είδαμε και στο προηγούμενο κεφάλαιο, τα chatbot αποτελούν το πιο διαδεδομένο παράδειγμα αλληλεπίδρασης ανάμεσα στον άνθρωπο και στη μηχανή, που έχει επιτύχει η Τ.Ν. Ως προγράμματα υπολογιστή που συνομιλούν μέσω φωνής ή κειμένου, μπορούν να μιμηθούν την ανθρώπινη συνομιλία, ανταποκρινόμενα σε ερωτήσεις και να αποτελέσουν χρήσιμα εργαλεία σε διάφορους τομείς, όπως η εκπαίδευση, το εμπόριο και η συμβουλευτική διαδικασία, η οποία και μας ενδιαφέρει περισσότερο στην παρούσα διατριβή.¹¹⁵

Τα μηνύματα που λαμβάνουμε από τη σχετική έρευνα είναι πολλά υποσχόμενα γύρω από τη συμβολή των chatbot στο κομμάτι της ψυχικής υγείας. Ένας αριθμός μελετών, έχουν δείξει πως η Συνομιλητική Τ.Ν. (Conversational AI - CAI) μπορεί να ανακουφίσει τα συμπτώματα της κατάθλιψης και του άγχους και να αποτελέσει την επιλογή πολλών ανθρώπων για τη συναισθηματική τους αποφόρτιση, μέσα από μια σειρά διαδοχικών επαφών συμβουλευτικής για ένα συγκεκριμένο χρονικό διάστημα.¹¹⁶ Επιπλέον, ήδη από το 2019, μια μεγάλης κλίμακας γραμμή βοήθειας για αυτοκτονία στην Αυστραλία ξεκίνησε δοκιμαστικά τη χρήση ενός ψηφιακού σύμβουλου με το όνομα Claire, το οποίο κατασκευάζει σενάρια και προτροπές συνομιλίας στην γραμμή

¹¹¹ Γεώργιος Μυρίσας, «Αναγνώριση τάσεων αυτοκτονίας σε κοινωνικά δίκτυα» (Διπλωματική διατριβή, Ελληνικό Ανοικτό Πανεπιστήμιο, 2021), 39.

¹¹² Γεωργούλη, *Τεχνητή Νοημοσύνη*, 243.

¹¹³ Γεώργιος Π. Δελατόλας, «Επεξεργασία φυσικής γλώσσας» (Διπλωματική διατριβή, Πανεπιστήμιο Πατρών, 2021), 108.

¹¹⁴ Ανδρέας Α. Κακογιάννης, «Ανάλυση συναισθήματος σε δεδομένα του Twitter με τεχνικές NLP» (Διπλωματική διατριβή, Πανεπιστήμιο Μακεδονίας, 2022), 34-35.

¹¹⁵ Eleni Adamopoulou, Lefteris Moussiades, “An overview of Chatbot Technology,” *ALAI* 584 (May, 2020): 373-383, https://link.springer.com/content/pdf/10.1007/978-3-030-49186-4_31.pdf.

¹¹⁶ Sedlakova, Trachsel, “Conversational Artificial Intelligence in Psychotherapy,” 5; Giovanni Rubeis, “E-mental health applications for depression: an evidence-based ethical analysis,” *Eur Arch Psychiatry Clin Neurosci* 271 (Jan., 2021): 549, <https://doi.org/10.1007/s00406-019-01093-y>.

Suicide Call Back Service¹¹⁷ και θέτει τις ανάλογες ερωτήσεις στον καλούντα, προκειμένου να ανιχνεύσει το επίπεδο κινδύνου αυτοκτονίας που ενυπάρχει στα λεγόμενά του και να τον κατευθύνει αναλόγως.¹¹⁸ Οι δε συνεχείς προσπάθειες των σχεδιαστών να αυξήσουν την ομοιότητα ενός chatbot με τον άνθρωπο, μέσα από τη μορφοποίηση διαφόρων χαρακτηριστικών όπως είναι η φωνή ή το όνομα του συνομιλητή, αναμένεται να εμπλουτίσει τη θεραπευτική σχέση που αναπτύσσεται και το βαθμό αλληλεπίδρασης με τον μη ανθρώπινο συνομιλητή.¹¹⁹

Το θέμα αυτό, ωστόσο, του ανθρωπομορφισμού είναι ένα από τα πολλά που προκύπτουν προς συζήτηση στο πλαίσιο μελέτης των θεραπευτικών chatbot. Μεταξύ άλλων, ο βαθμός προσομοίωσης στο ανθρώπινο μοντέλο θεραπευτή, η εδραίωση εμπιστοσύνης στη θεραπευτική σχέση, οι δυνατότητες για ενσυναίσθηση και αυτο-αποκάλυψη, αποτελούν θέματα που θα μας απασχολήσουν στο επόμενο κεφάλαιο, όπου η εφαρμογή της Τ.Ν. στις τηλεφωνικές γραμμές βοήθειας, ως μέθοδος παροχής συμβουλευτικής υποστήριξης θα αποτελέσει το αντικείμενο της εξέτασής μας. Για την ώρα, ίσως θα πρέπει να αρκεστούμε στο γεγονός ότι παρόλο που εκατομμύρια άνθρωποι ανά τον κόσμο προβαίνουν στην χρήση των chatbot με στόχο την αναζήτηση ανακούφισης ή θεραπείας, τα αποτελέσματά τους δεν έχουν αξιολογηθεί ακόμη επαρκώς, λόγω του αρχικού σταδίου των ερευνών πάνω σε αυτόν τον τομέα.¹²⁰ Από τη μία λοιπόν, η εκτεταμένη και συνεχώς αυξανόμενη χρήση των συστημάτων της Τ.Ν. ως μέθοδος παροχής της Συμβουλευτικής και από την άλλη, η μέχρι τώρα ανεπαρκής γνώση και ανάλυση των κινδύνων που η χρήση αυτή ενέχει, καθώς και των τρόπων που θα μπορούσαν να βελτιώσουν την αποτελεσματικότητα της. Αυτή η υπάρχουσα ασυμμετρία είναι σημαντική, γιατί καθιστά ιδιαίτερα πολύπλοκη και δυσχερή τη διατύπωση των κατάλληλων ρυθμιστικών πλαισίων (νομικών και ηθικών) σχετικά με τη χρήση των συστημάτων της Τ.Ν. στη Συμβουλευτική.

Το βέβαιο είναι πως οι νέες παρεμβάσεις που μας προσφέρει η συνεχής εξέλιξη της Τ.Ν. επιφέρουν σημαντικά οφέλη σε ανθρώπους με ευάλωτο ψυχισμό. Η άμεση και εύκολη πρόσβαση σε συμβουλές και υπηρεσίες χωρίς πρόσθετες ενέργειες από το χρήστη ή απαραίτητες εγκαταστάσεις εφαρμογών,¹²¹ οι δυνατότητες για εφαρμογή προληπτικών προγραμμάτων μέσω της έγκαιρης ανίχνευσης ψυχικών διαταραχών που μακροπρόθεσμα θα μειώσουν το κόστος της υγειονομικής περίθαλψης,¹²² η επίτευξη της αντικειμενικότητας στη μέτρηση και της ολιστικότητας στην προσέγγιση του ανθρώπινου ψυχισμού και η γενικότερη, ταχεία επέκταση των ψηφιακών τεχνολογιών στο πλαίσιο της ψυχικής υγείας, φαίνεται να επιβεβαιώνουν τους ισχυρισμούς των σαράντα μελών της Επιτροπής της Παγκόσμιας Ψυχιατρικής Εταιρείας για το Μέλλον της Ψυχιατρικής, πως η επανάσταση της ψηφιακής Ψυχιατρικής έφτασε.¹²³

¹¹⁷ “Feeling suicidal,” Suicide Call Back Service, προσπελάστηκε στις 28 Αυγούστου 2023, <https://www.suicidecallbackservice.org.au/feeling-suicidal/>.

¹¹⁸ Piers Gooding, “Mapping the rise of digital mental health technologies: Emerging issues for law and society,” *International Journal of Law and Psychiatry* 67 (Oct, 2019): 5, <https://doi.org/10.1016/j.ijlp.2019.101498>. Επιπλέον, βλ. “Virtual counsellor steps in to help out on suicide hotline,” Probono Australia, accessed August 20, 2023, https://probonoaustralia-com-au.translate.goog/news/2019/04/virtual-counsellor-steps-in-to-help-out-on-suicide-hotline/?_x_tr_sl=en&_x_tr_tl=el&_x_tr_hl=el&_x_tr_pto=wapp.

¹¹⁹ Adamopoulou, Moussiades, “An overview of Chatbot Technology,” 375-376.

¹²⁰ Sedlakova, Trachsel, “Conversational Artificial Intelligence in Psychotherapy,” 5.

¹²¹ Adamopoulou, Moussiades, “An overview of Chatbot Technology,” 374.

¹²² Ζηκόπουλος, «Η χρήση της τεχνητής νοημοσύνης», 36.

¹²³ Gooding, “Mapping the rise of digital mental health technologies,” 1.

Ωστόσο, ενόσω η Τ.Ν. διεισδύει ολοένα και περισσότερο στην καθημερινή ζωή των ανθρώπων, καθίσταται επιτακτική η ανάγκη να δημιουργηθούν τα κανονιστικά εκείνα πλαίσια, που θα διασφαλίσουν το απόρρητο και την ιδιωτικότητα κατά την εδραίωσή της στις ψυχολογικές και ιατρικές επιστήμες. Σε μία εποχή, όπου πολλές εταιρίες εξασφαλίζουν τα έσοδα τους μέσα από την εκμετάλλευση των προσωπικών δεδομένων, η διασφάλιση των όρων που θα μειώσουν την κακή χρήση αυτών εις βάρος των ευάλωτων κοινωνικών ομάδων, απαιτεί τη στροφή σε μεθόδους εφαρμογής που θα παράγουν οφέλη για την κοινωνία, αντί να περιορίσουν τα δικαιώματα της ιδιωτικής ζωής των ανθρώπων.¹²⁴

Πιο συγκεκριμένα, η κύρια ανησυχία που εγείρει η συλλογή των αναρίθμητων δεδομένων που χρησιμοποιεί η Τ.Ν., έχει να κάνει με το κατά πόσο είναι επαρκής η παραδοσιακή έννοια της συναίνεσης, τη στιγμή που είναι ολοένα και πιο δύσκολο να προβλεφθεί ποιος θα έχει πρόσβαση στα δεδομένα αυτά μελλοντικά, για ποιους σκοπούς και υπό ποιες συνθήκες. Αν και η συνεργασία μεταξύ οργανισμών υγειονομικής περίθαλψης και εταιριών που δραστηριοποιούνται στον τομέα της τεχνολογίας υπόσχονται τον εμπλουτισμό της γνώσης που θα διευκολύνει τις στοχευμένες προσεγγίσεις στη διάγνωση και θεραπεία των ασθενών, οι προβλεπόμενες συμφωνίες με εθνικές κυβερνήσεις ή ιδιωτικούς οργανισμούς, δίνουν τεράστιες πρωτοβουλίες στα χέρια των τελευταίων, ως προς την εκμετάλλευση του τεράστιου όγκου πληροφοριών σχετικά με την υγεία εκατομμυρίων πολιτών. Κατά συνέπεια, διακινδυνεύουν την ανεπιθύμητη διαρροή των πληροφοριών αυτών, που μπορεί να οδηγήσει στην μεροληπτική τους χρήση σε τομείς όπως η απασχόληση, η εκπαίδευση και η ασφάλιση.¹²⁵

Έτσι, ο διαμοιρασμός των δεδομένων στις ασφαλιστικές εταιρίες για την επιλογή πελατών που θα συμφέρουν οικονομικά, στις ιδιωτικές κλινικές για το διαχωρισμό των ασθενών και ένα σύστημα φιλικό μόνο προς ασθενείς που θα αυξήσουν τα κέρδη ή στα τμήματα ανθρωπίνου δυναμικού εταιριών που θα ορίσουν αναλόγως τις προσλήψεις του προσωπικού με βάση τα στοιχεία του ψυχισμού των υποψηφίων, αποτελούν ελάχιστα μόνο παραδείγματα του τρόπου με τον οποίο η συλλογή δεδομένων θα μπορούσε να χρησιμοποιηθεί εναντίον του ατόμου.¹²⁶

Επιπλέον, διάφορες άλλες αδυναμίες επιστημολογικής και μεθοδολογικής φύσεως δημιουργούν ερωτηματικά και ανασφάλεια ως προς το εάν, η πολλά υποσχόμενη εξάπλωση της χρήσης των συστημάτων της Τ.Ν. στον τομέα της ψυχικής υγείας θα καταφέρει να αντιμετωπίσει τις ηθικές προκλήσεις που αναδύονται. Μεταξύ άλλων, οι ανησυχίες για την αυτονομία και την αξιοπιστία σε περίπτωση τεχνικών βλαβών, η σταδιακή μετατόπιση της ιατρικής εξουσίας στους αλγόριθμους, που επιφέρει τη διαπροσωπική απόσταση μεταξύ γιατρού και ασθενή και τον κίνδυνο της κοινωνικής απομόνωσης για τον δεύτερο, η υπερ-εκπροσώπηση συγκεκριμένων ομάδων στο σύνολο των ιατρικών δεδομένων που μπορεί να οδηγήσει στην εξαγωγή συμπερασμάτων μεροληπτών υπέρ αυτών και κατά των μειονοτήτων ή των κοινωνικά περιθωριοποιημένων ομάδων και οι νέες μορφές ταξινόμησης στις οποίες ενδέχεται να προβούν οι αλγόριθμοι με βάση τη συσχέτιση μεταξύ βιολογικών, προσωπικών ή και κοινωνικών χαρακτηριστικών, αποτελούν ελάχιστα μόνο παραδείγματα που δύνανται να παραβιάσουν τις τέσσερις αρχές της *Βιοηθικής*

¹²⁴ Montag, Sindermann, Baumeister, “Digital phenotyping,” 19-20.

¹²⁵ Alessandro Blasimme, Effy Vayena, “The Ethics of AI in Biomedical Research, Patient Care, and Public Health,” in *The Oxford Handbook of Ethics of AI*, ed. Markus D. Dubber, Frank Pasquale, Sunit Das (New York, Oxford University Press. 2020), 704-706.

¹²⁶ Montag, Sindermann, Baumeister, “Digital phenotyping,” 21-22.

(αυτονομία, αγαθοποιία, μη βλάβη και δικαιοσύνη προς τον ασθενή) στον ταχέως αναπτυσσόμενο τομέα της Τ.Ν.¹²⁷

Η ανάλυση των δεδομένων όμως απαιτεί ολοένα και περισσότερη Τ.Ν. και η τελευταία αυτή, προβαίνει σε όλο και περισσότερη συλλογή και επεξεργασία των δεδομένων, πράγμα που απαιτεί τη συνεργασία με τον άνθρωπο. Επομένως, η κοινωνική και πολιτική εφαρμογή κατευθυντηρίων γραμμών είναι υψίστης σημασίας για την επίτευξη της εμπιστοσύνης από την πλευρά του ανθρώπου και τη δημιουργία της αίσθησης πως αυτή η εξέλιξη δεν θα τον βλάψει.¹²⁸

Στη συζήτηση αυτή, περί των ηθικών προκλήσεων που προκύπτουν, θα επανέλθουμε αργότερα και συγκεκριμένα, στο 4^ο κεφάλαιο της παρούσας διατριβής, όπου θα αναλύσουμε το πεδίο της Ηθικής της Τ.Ν. Προτού φτάσουμε εκεί, αξίζει να δούμε, σε συνέχεια των ωφελειών και των αδυναμιών που ελλοχεύουν στη χρήση της Τ.Ν. στον τομέα της ψυχικής υγείας και που είδαμε μόλις, τους παράγοντες εκείνους, βάσει των οποίων μπορούμε να προβούμε στην αξιολόγηση της εφαρμογής της ή αλλιώς, τις γενικότερες προϋποθέσεις που θα πρέπει να πληροί η αξιολόγηση αυτή.

2.β. Οι παράγοντες που επηρεάζουν την αξιολόγηση των αποτελεσμάτων

Όπως είδαμε έως τώρα, ένα ευρύ φάσμα υπηρεσιών και δυνατοτήτων που αναπτύσσονται στο πεδίο της Ψυχοπληροφορικής γίνεται προσβάσιμο στον άνθρωπο, προϋποθέτοντας μονάχα την κατοχή μιας κινητής συσκευής τύπου smartphone και τη σύνδεση στο διαδίκτυο, κάτι που σήμερα, ισχύει για παραπάνω από το 60% του παγκόσμιου πληθυσμού. Αυτόματα, ωστόσο, η πληροφορία αυτή συνεπάγεται πως ένα ποσοστό του 40% περίπου του ίδιου πληθυσμού, που δεν έχει πρόσβαση στο διαδίκτυο, αποκλείεται από τα παραπάνω.¹²⁹

Η αδυναμία αυτή στην πρόσβαση μπορεί να οφείλεται στον παράγοντα της ηλικίας (π.χ. ηλικιωμένοι), στην ύπαρξη βαριάς αναπηρίας ή βαριάς ψυχικής νόσου, κυρίως όμως, σχετίζεται με την οικονομική αδυναμία και το χαμηλό επίπεδο των ψηφιακών δεξιοτήτων του πληθυσμού, δύο μεταβλητές που συχνά είναι άρρηκτα συνδεδεμένες μεταξύ τους.¹³⁰ Παράλληλα, έκπληξη προκαλεί το γεγονός πως ο ψηφιακός αναλφαβητισμός είναι έντονος ακόμη και σε ομάδες νεότερων ή υψηλά μορφωμένων ανθρώπων και γενικότερα, είναι πράγματι ανησυχητικό πως ακόμη και σε χώρες όπου τα ποσοστά συνδέσεων στο διαδίκτυο είναι υψηλά, το 25% του πληθυσμού στερείται βασικών ψηφιακών δεξιοτήτων.¹³¹

Τα τελευταία χρόνια, η εκτόξευση της χρήσης της τεχνολογίας που προκλήθηκε από την πανδημία του COVID-19 (για λόγους ενημέρωσης, ανίχνευσης επαφών, έκδοση

¹²⁷ Blasimme, Vayena, “The Ethics of AI in Biomedical Research,” 707-711.

¹²⁸ Montag, Sindermann, Baumeister, “Digital phenotyping,” 21-22.

¹²⁹ *Αυτόθι*, 20.

¹³⁰ Για περισσότερα στοιχεία σχετικά με την έλλειψη ψηφιακών δεξιοτήτων παγκοσμίως, βλ. “The Digital Skills Gap Index (DSGI),” Wiley, accessed September 7, 2023, <https://dsgi.wiley.com/global-rankings/>.

¹³¹ Esther I. Metting, Eveline Hage, “Outside the Digital Bubble: Digital illiteracy is common and deserves more attention when implementing ehealth,” *Social Science Research Network (SSRN)* (Oct., 2022): 2, <https://doi.org/10.2139/ssrn.4246806>.

ψηφιακού πιστοποιητικού εμβολιασμού κ.ο.κ.), διεύρυνε ακόμη περισσότερο το χάσμα μεταξύ των ψηφιακά γραμματισμένων και των ψηφιακά αναλφάβητων πολιτών. Πέραν της πανδημίας, το φαινόμενο του ψηφιακού αναλφαβητισμού δίνει μια άλλη διάσταση στο κομμάτι της ψυχικής υγείας που μας απασχολεί εδώ και στο πλαίσιο της στενής σχέσης που σταδιακά εγκαθιδρύεται μεταξύ αυτής και της τεχνολογίας. Πιο συγκεκριμένα, το γεγονός πως η αποτελεσματικότητα των όποιων μεθόδων εφαρμόζει η Τ.Ν. για την πρόληψη, τη διάγνωση ή τη θεραπεία της ψυχικής νόσου, ενδέχεται να παρεμποδιστεί από τις ανεπάρκειες στην πρόσβαση ή στην αδυναμία και άγνοια χειρισμού της τεχνολογίας, αποτελεί ένα καίριο θέμα προς συζήτηση και επίλυση, εάν θέλουμε να αποφευχθούν οι ανισότητες και οι διακρίσεις στο δικαίωμα της υγειονομικής περίθαλψης.¹³²

Πέραν τούτου και αν υποθέσουμε πως το πρόβλημα του ψηφιακού αναλφαβητισμού περιοριστεί, μέσω ενεργειών που θα στοχεύουν στον ψηφιακό εγγραμματισμό των πληθυσμιακών ομάδων και στην προσαρμογή των εκπαιδευτικών μεθόδων ώστε να ανταποκρίνονται ικανοποιητικά σε συγκεκριμένες πληθυσμιακές ομάδες (π.χ. άτομα με βαριά διανοητική αναπηρία, τυφλότητα, αυτισμό, τύφλωση), μια ακόμη σημαντική προϋπόθεση για την αποτελεσματικότητα των εφαρμογών της Τ.Ν. στο πλαίσιο του ψηφιακού μετασχηματισμού που σημειώνεται ραγδαίως, είναι η αποδοχή τους από πλευράς των ανθρώπων και η παραδοχή πως η εμπλοκή της Τ.Ν. μπορεί πράγματι να είναι προς όφελός τους.¹³³

Πάνω σε αυτό, αξιοσημείωτες είναι οι σχετικές μελέτες, όπως αυτή των Philip Fei Wu et al. που διερεύνησε τις αντιλήψεις γύρω από τη χρήση των ΑΙ chatbot για τη διαχείριση χρόνιων προβλημάτων υγείας και μακροχρόνιων επιπτώσεων του COVID-19. Η έρευνα έδειξε πως, παρά το μεγάλο βαθμό εξοικείωσης των συμμετεχόντων με τις ψηφιακές εφαρμογές υγείας, λιγότερο από το ένα τρίτο αυτών δήλωσε ως πιθανό το να κάνει χρήση ενός chatbot υγείας, εκφράζοντας ανασφάλεια και ανησυχίες για τις δυνατότητες αυτού να παρέχει ακριβείς ιατρικές συμβουλές. Το σύνολο των συμμετεχόντων δήλωσε πως προτιμά, κατά πολύ περισσότερο, να μιλήσει με έναν πραγματικό άνθρωπο, παρά με ένα ρομπότ, για θέματα γύρω από την ψυχική και σωματική του υγεία.¹³⁴ Παρομοίως, στην έρευνα των Sweeney et al., η οποία εστίασε στις απόψεις των επαγγελματιών ψυχικής υγείας σχετικά με τη χρήση των chatbot για την ενίσχυση και τη βελτίωση των παροχών, αν και το 60% των ερωτηθέντων δήλωσε πως τα chatbot ψυχικής υγείας θα μπορούσαν να αποτρέψουν τις άσκοπες επισκέψεις στους παρόχους συναφών υπηρεσιών, περισσότεροι από τους μισούς, εξέφρασαν προβληματισμό γύρω από τις δυνατότητες των chatbot να κατανοήσουν επαρκώς τα ανθρώπινα συναισθήματα, να αξιολογήσουν την ψυχολογική κατάσταση των ασθενών και να ανταποκριθούν στις ανάγκες τους.¹³⁵

¹³² *Αυτόθι*, 2.

¹³³ Montag, Sindermann, Baumeister, “Digital phenotyping,” 21.

¹³⁴ Philip Fei Wu, Charlotte Summers, Arjun Panesar, Amit Kaura, Li Zhang, “AI Hesitancy and Acceptability: Perceptions of AI Chatbots for Chronic Health Management and Long-COVID Support,” *JMIR Preprints* (July, 2023), <https://s3.ca-central-1.amazonaws.com/assets.jmir.org/assets/preprints/preprint-51086-submitted.pdf>.

¹³⁵ Colm Sweeney, Courtney Potts, Edel Ennis, Raymond Bond, Maurice D. Mulvenna, Siobhan O’neill, Martin Malcolm, Lauri Kuosmanen, Catrine Kostenius, Alex Vakaloudis, Gavin Mcconvey, Robin Turkington, David Hanna, Heidi Nieminen, Anna-Kaisa Vartiainen, Alison Robertson, Michael F. Mctear, “Can chatbots help support a person’s mental health? Perceptions and views from mental healthcare professionals and experts,” *ACM Transactions on Computing for Healthcare* 2, no. 3 (July, 2021), 7-8, <https://doi.org/10.1145/3453175>.

Θα πρέπει να τονίσουμε πως η επιφυλακτικότητα αυτή του κοινού, δεν επικεντρώνεται αποκλειστικά στην χρήση μεθόδων που επιστρατεύουν Τ.Ν. αλλά γενικότερα, στο σύνολο των παρεμβάσεων που υποβοηθούνται από τη χρήση της τεχνολογίας. Αξιοσημείωτες είναι οι σχετικές μελέτες, όπως αυτή των Baumeister et al., η οποία εξέτασε την αποδοχή των διαδικτυακών παρεμβάσεων (*Internet-based intervention – IBI*) για την κατάθλιψη, από καταθλιπτικούς ασθενείς με διαβήτη. Τα αποτελέσματα της έρευνας έδειξαν ένα σχετικά χαμηλό έως μέτριο επίπεδο αποδοχής για το 91,5% των συμμετεχόντων, πόρισμα που δε διαφέρει κατά πολύ από τα ευρήματα προηγούμενων μελετών. Παραδείγματος χάριν, η έρευνα των Mohr et al.¹³⁶ που συνέκρινε το ποσοστό των ασθενών πρωτοβάθμιας περίθαλψης, οι οποίοι ενδιαφέρθηκαν για ψυχολογική και συμπεριφορική θεραπεία δια ζώσης, τηλεφωνικά ή μέσω Διαδικτύου, διαπίστωσε πως μόνο το 11,6% αυτών που έδειξαν ενδιαφέρον να λάβουν κάποια μορφή παρέμβασης, εξέφρασαν το σοβαρό ενδιαφέρον τους για μια διαδικτυακή παρέμβαση (IBI), ενώ το 51,8% αρνήθηκε αυτή τη μορφή.¹³⁷

Ως εκ τούτου, η εφαρμογή παρεμβάσεων που θα αφορούν στην αποδοχή της Τ.Ν. (*Acceptance Facilitating Interventions - AFI*), μέσα από την παροχή ενημέρωσης σχετικά με τη λειτουργία της, τα οφέλη και τους κινδύνους που ενέχουν, αλλά και τη δυνατότητα αυτο-βελτίωσης που παρέχει, θα μπορούσε να βελτιώσει το βαθμό αποδοχής μεθόδων όπως ο ψηφιακός φαινότυπος και το IOT (Internet of Things - IoT).¹³⁸ Άλλωστε, η άποψη αυτή επιβεβαιώνεται μέσα από έρευνες, όπως αυτή που διεξήχθη από τους Gun et al.¹³⁹ και έδειξε πως την πιθανότητα αποδοχής μιας διαδικτυακής παρέμβασης μελλοντικά, θα μπορούσε να αυξήσει η παροχή πληροφοριών σχετικά με την αποτελεσματικότητα, την προσβασιμότητα και την ασφάλεια της. Παρομοίως η έρευνα των Mitchell και Gordon¹⁴⁰ εντόπισε μια αύξηση στο βαθμό αποδοχής μετά από μια σύντομη επίδειξη διαδικτυακής παρέμβασης (IBI) στους συμμετέχοντες της έρευνας, ενώ οι Cranen et al.,¹⁴¹ εντόπισαν μια σημαντική θετική επίδραση στις αντιλήψεις γύρω από μια υπηρεσία τηλεϊατρικής, μετά τη σύντομη χρήση ενός συστήματος της σε πληθυσμό ασθενών με χρόνια πόνου.¹⁴²

Όπως έχουμε δει μέχρι τώρα, τα συστήματα της Τ.Ν. που εφαρμόζονται στον τομέα της ψυχικής υγείας, εισέρχονται στην καθημερινή ζωή των ανθρώπων και μέσω

¹³⁶ Βλ. David C. Mohr, Juned Siddique, Joyce Ho, Jenna Duffecy, Ling Jin, J. Konadu Fokuo, “Interest in Behavioral and Psychological Treatments Delivered Face-to-Face, by Telephone, and by Internet,” *Annals of Behavioral Medicine* 40, no. 1 (Aug., 2010): 89–98, <https://doi.org/10.1007/s12160-010-9203-7>.

¹³⁷ H. Baumeister, L. Nowoczin, J. Lin, H. Seiffert, J. Seufert, K. Laubner, D.D. Ebert, “Impact of an acceptance facilitating intervention on diabetes patients’ acceptance of Internet-based interventions for depression: A randomized controlled trial,” *Diabetes Research and Clinical Practice* 105, no. 1 (July, 2014): 35, <https://doi.org/10.1016/j.diabres.2014.04.031>.

¹³⁸ Montag, Sindermann, Baumeister, “Digital phenotyping,” 21.

¹³⁹ Βλ. Shih Ying Gun, Titov Nickolai, Gavin Andrews, “Acceptability of Internet treatment of anxiety and depression,” *Australas Psychiatry* 19, no. 3 (June, 2011): 259-264, 10.3109/10398562.2011.562295.

¹⁴⁰ Βλ. Kenneth P. Gordon, “Attitudes towards computerized CBT for depression amongst a student population,” *Behavioral and Cognitive Psychotherapy* 3, no. 4 (2007): 421–430, https://www.researchgate.net/publication/232021580_Attitudes_Towards_Computerized_CBT_for_Depression_Amongst_a_Student_Population.

¹⁴¹ Karlijn Cranen, Constance H.C.Drossaert, Evelien S. Brinkman, Annemarie L.M. Braakman-Jansen, Maarten J. Ijzerman, Miriam M.R. Vollenbroek-Hutten, “An exploration of chronic pain patients’ perceptions of home telerehabilitation services,” *Health Expert* 15, no. 4 (Dec. 2012): 339-350, doi: 10.1111/j.1369-7625.2011.00668.x.

¹⁴² Baumeister et al. “Impact of an acceptance facilitating intervention,” 35-36.

αισθητήρων σε ψηφιακές συσκευές καθημερινής χρήσης, προβαίνουν στην καταγραφή της κάθε στιγμής και στην λεπτομερή ανάλυση των δεδομένων μέσω ισχυρών αλγορίθμων. Αν και αποτελούν μια σημαντική εξέλιξη στην παροχή υπηρεσιών ψυχικής υγείας, αυτή η διείσδυση στην καθημερινότητα του ανθρώπου απαιτεί πρωτίστως, την αποδοχή από τον ίδιο. Έτσι, η γνώση της γενικής στάσης του πληθυσμού απέναντι στις μεθόδους της Τ.Ν. στο κομμάτι της ψυχικής υγείας είναι σημαντική για την κατανόηση του κατά πόσο ένα έδαφος είναι πρόσφορο για την εφαρμογή τους. Ταυτόχρονα, η ακριβής ενημέρωση και η επιμόρφωση των πολιτών, καθώς και η παροχή ίσων ευκαιριών στην πρόσβαση, συνθέτουν το απαραίτητο σύνολο των παρεμβάσεων που θα στοχεύουν στην αύξηση του βαθμού αποδοχής και στην ενεργή συνεργασία του πληθυσμού, που είναι απαραίτητη για την επέκταση τέτοιων μεθόδων.

Πέρα από τα παραπάνω, παρά τα ιδιαίτερος αισιόδοξα μηνύματα των μέχρι τώρα ερευνών, η εξασφάλιση της ακρίβειας και της αξιοπιστίας των συστημάτων της Τ.Ν. είναι άκρως απαραίτητη πριν την εκτεταμένη διάδοση και εφαρμογή τους παγκοσμίως. Έτσι, η περαιτέρω επιβεβαίωση μέσω μεγαλύτερων δειγμάτων και η μεγαλύτερη διάρκεια διεξαγωγή των ερευνητικών διαδικασιών είναι απαραίτητες για την διασφάλιση της εγκυρότητας των διαφορετικών μεθόδων.¹⁴³ Επιπλέον, μια ιδιαίτερου ενδιαφέροντος μέθοδος που προτείνεται από πολλούς μελετητές, είναι η δομημένη έρευνα στο πλαίσιο των κριτηρίων του ερευνητικού τομέα *RDoC*¹⁴⁴ (*Research Domain Criteria*).¹⁴⁵ Αυτό, καθότι σε αντίθεση με την κλασική διάγνωση των συστημάτων *ICD-11* ή *DSM-5*,¹⁴⁶ που συχνά δεν κατανοούν την πραγματική φύση μιας διακριτής ψυχικής διαταραχής, το *RDoC* επιδιώκει να ποσοτικοποιήσει τις ασθένειες, εξετάζοντάς τις μέσω μιας σειράς εννοιών που ανταποκρίνονται σε συγκεκριμένες λειτουργικές διαστάσεις της συμπεριφοράς, μέσω βιολογικών μετρήσεων σε επίπεδο γονιδίων, μορίων, κυττάρων και νευρικών κυκλωμάτων, καθώς και μέσω μιας σειράς παρατηρήσιμων μονάδων ανάλυσης όπως η φυσιολογία, η συμπεριφορά και οι αυτοαναφορές του ασθενή.¹⁴⁷ Αυτή η ολιστική προσέγγιση στο πλαίσιο της διερεύνησης των ψυχικών διαταραχών, σε συνδυασμό με τις δυνατότητες των συστημάτων Τ.Ν. ως προς τη συλλογή δεδομένων γύρω από τα πεδία στα οποία εστιάζει το ερευνητικό πλαίσιο *RDoC*, αναμένεται ότι θα συμβάλλει στην ακριβή και αξιόπιστη

¹⁴³ D'Alfonso, "AI in mental health," 114.

¹⁴⁴ Το *RDoC* (*Research Domain Criteria*) αποτελεί ένα ερευνητικό πλαίσιο για τη διερεύνηση των ψυχικών διαταραχών, που αναπτύχθηκε από το Εθνικό Ινστιτούτο Ψυχικής Υγείας των ΗΠΑ. Στόχος είναι η ολοκληρωμένη κατανόηση της φύσης της ψυχικής υγείας, στο πλαίσιο της ανθρώπινης νευροσυμπεριφορικής λειτουργίας και τη μελέτη μεταβλητών, όπως το φυσικό και πολιτισμικό περιβάλλον, η ανάπτυξη και ωρίμανση του εγκεφάλου, η συμπεριφορά κ.ο.κ. Βλ. "About *RDoC*," National Institute of Mental Health, accessed September 11, 2023, <https://www.nimh.nih.gov/research/research-funded-by-nimh/rdoc/about-rdoc>.

¹⁴⁵ Montag, Sindermann, Baumeister, "Digital phenotyping," 22.

¹⁴⁶ Το *ICD-11* αναφέρεται στο διεθνές σύστημα ταξινόμησης νοσημάτων που αναπτύχθηκε από τον Παγκόσμιο Οργανισμό Υγείας. Βλ. "ICD-11 International Classification of Diseases 11th Revision. The global standard for diagnostic health information," World Health Organization, accessed September 7, 2023, <https://icd.who.int/en>. Αντίστοιχα, το *DSM-5* αναφέρεται στο Διαγνωστικό και Στατιστικό Εγχειρίδιο Ψυχικών Διαταραχών (5^η έκδοση), που αναπτύχθηκε από την Αμερικανική Ψυχιατρική Εταιρεία. Βλ. "Diagnostic and Statistical Manual of Mental Disorders (DSM-5-TR)," American Psychiatric Association, accessed September 7, 2023, <https://www.psychiatry.org/psychiatrists/practice/dsm>.

¹⁴⁷ J. Torous, J.P. Onnela, M. Keshavan, "New dimensions and new tools to realize the potential of *RDoC*: digital phenotyping via smartphones and connected devices," *Translational Psychiatry* 7, no. 1053 (March, 2017): 2, <https://doi.org/10.1038/tp.2017.25>.

διάγνωση, προτού οι συζητήσεις για τις αντίστοιχες θεραπευτικές παρεμβάσεις λάβουν χώρα.

Ακόμη, λαμβάνοντας υπόψη τις επιστημολογικές και μεθοδολογικές αδυναμίες των συστημάτων της T.N. που αναφέραμε νωρίτερα σε αυτό το κεφάλαιο, έχει ιδιαίτερη σημασία το να συνειδητοποιήσουμε πως οι περιορισμοί αυτοί καθιστούν απαραίτητη μέχρι στιγμής, τόσο την καθοδήγηση της έρευνας από τη θεωρία, όσο και την προσεκτική επίβλεψη εκπαιδευμένου επιστημονικού προσωπικού. Πιο συγκεκριμένα, τα αναρίθμητα και πολυποίκιλα δεδομένα, που εισέρχονται με υψηλές ταχύτητες σε συγκεκριμένα ερευνητικά πεδία, αποτυπώνουν την ικανότητα πρόβλεψης της συμπεριφοράς σε επίπεδο ομαδικό. Ως εκ τούτου, η πρόβλεψη για τη συμπεριφορά ενός μεμονωμένου ατόμου δεν είναι ακόμη δυνατή με απόλυτη βεβαιότητα. Επιπλέον, τα εκατομμύρια ψηφιακά δεδομένα που συγχωνεύονται μεταξύ τους, μπορεί να οδηγήσουν σε πολλές ψευδείς συσχετίσεις και ψευδώς θετικά αποτελέσματα, κατά την προσπάθεια συσχέτισης χιλιάδων μεταβλητών μεταξύ τους. Στο σημείο αυτό, η θεωρητική καθοδήγηση μπορεί να είναι καθοριστική, τόσο για τον περιορισμό της αναζήτησης σε ένα μικρότερο σύνολο δεδομένων, όσο και για τον σωστό χειρισμό των δεδομένων αυτών, ώστε να μη διογκωθούν τα ποσοστά των σφαλμάτων κατά την προσπάθεια συσχέτισης των πολλών μεταβλητών.¹⁴⁸

Από την άλλη, η αξία που έχει η επίβλεψη των ερευνών και των παρεμβάσεων της T.N. από κλινικά εκπαιδευμένο επιστημονικό προσωπικό, δεν μπορεί να αγνοηθεί κατά κανένα τρόπο. Η πολυπλοκότητα του ανθρώπινου ψυχισμού και των κλινικών περιστατικών καθιστά αναγκαία τη συγκριτική ανάλυση των πορισμάτων από μεθόδους όπως ο ψηφιακός φαινότυπος και η επεξεργασία της φυσικής γλώσσας (Natural language Processing - NLP), με δεδομένα αυτοαναφοράς και την κλινική εικόνα του ασθενή. Επιπλέον, οι περιορισμένες ανταποκρίσεις των chatbot σε καταστάσεις έκτακτης ανάγκης, όπως ο αυτοκτονικός ιδεασμός ή η κακοποίηση σε πραγματικό χρόνο, μπορεί να αποβούν επικίνδυνες και ακατάλληλες, όπως αποδεικνύει μια σειρά σχετικών ερευνών.¹⁴⁹

Χαρακτηριστικά, από τη δοκιμή που πραγματοποιήθηκε στα chatbot Wysa και Woebot, για τη διαχείριση αναφοράς σεξουαλικής κακοποίησης παιδιών, προέκυψε ότι καμία από τις δύο εφαρμογές δεν ήταν αποτελεσματική σε αυτό. Συγκεκριμένα, καμία τους δεν προέτρεψε το υποτιθέμενο θύμα να αναζητήσει βοήθεια, ενώ, η απάντηση του chatbot Woebot στη φράση «Με αναγκάζουν να κάνω σεξ και είμαι μόλις 12 ετών», ήταν «Λυπάμαι που σας συμβαίνει αυτό, αλλά επίσης φανερώνει την ανάγκη σας για σύνδεση, κάτι που είναι πολύ όμορφο».¹⁵⁰ Οι παρόμοιες παραλείψεις που έγιναν κατά τις αναφορές περί διατροφικών διαταραχών και χρήσης ναρκωτικών, αποδεικνύουν πως αν και οι μέθοδοι της Ψυχοπληροφορικής δύναται να προάγουν σημαντικά το ερευνητικό πεδίο της ψυχικής υγείας, η συμβολή των ειδικευμένων επαγγελματιών, σαφώς και δεν μπορεί να αντικατασταθεί, αυτή τη στιγμή, από τα συστήματα της T.N. Αντιθέτως, η χρήση των chatbot θα πρέπει τουλάχιστον να συνοδεύεται από τη δυνατότητα ειδοποίησης και περαιτέρω επικοινωνίας για ζωντανή βοήθεια από

¹⁴⁸ Christian Montag, Jon D. Elhai, “A new agenda for personality psychology in the digital age?,” *Personality and Individual Differences* 147 (Sep., 2019): 131-132, <https://doi.org/10.1016/j.paid.2019.03.045>.

¹⁴⁹ *Αυτόθι*.

¹⁵⁰ “Child advice chatbots fail to spot sexual abuse,” BBC, last modified December 11, 2018, <https://www.bbc.com/news/technology-46507900>.

ειδικό.¹⁵¹ Αυτό συνεπάγεται πως η μετάβαση στην πλήρως αυτόνομη λειτουργία των συστημάτων της Συμβουλευτικής Τ.Ν. (*Human off the loop*) δεν είναι ακόμη εφικτή. Αντιθέτως, μέχρι στιγμής, η διατήρηση των συστημάτων Τ.Ν. στην μορφή εκείνη, στην οποία υπόκεινται σε ολικό (*Human in the loop*) ή μερικό έλεγχο από τον άνθρωπο (*Human on the loop*),¹⁵² κρίνεται απαραίτητη για την ασφάλεια των χρηστών κατά τον χειρισμό των ευαίσθητων και σοβαρών προσωπικών τους ζητημάτων.

Χωρίς αμφιβολία, η μετάβαση από το σημείο εκείνο, όπου η μόνη ευκαιρία των ερευνητών να κατανοήσουν την ψυχολογική κατάσταση ενός ανθρώπου δεν ήταν παρά να ρωτήσουν τον ίδιο, στο σήμερα, όπου οι αναρίθμητες μεταβλητές δύναται να εξεταστούν από τους αλγορίθμους της Τ.Ν. που θα μας δώσουν συμπεράσματα για την προσωπικότητά του, αποτελεί ένα μεγάλο βήμα. Ωστόσο, το κομμάτι του ανθρώπινου ψυχισμού είναι τόσο ευαίσθητο που δεν επιδέχεται ρίσκο ή άλλες ανεπάρκειες. Έτσι, προκειμένου η εφαρμογή της Τ.Ν. στο κομμάτι της ψυχικής υγείας να μπορέσει να αξιολογηθεί, παράγοντες όπως αυτοί που αναφέρθηκαν παραπάνω, καθίστανται στην ουσία οι προϋποθέσεις της εφαρμογής αυτής. Η ανταπόκριση των συστημάτων στα ιδιαίτερα χαρακτηριστικά των πληθυσμιακών ομάδων στις οποίες απευθύνονται, η αποδοχή και η πρόθεση συνεργασίας από τις ομάδες αυτές, η επίβλεψη των εξειδικευμένων στην ψυχική υγεία ειδικών, καθώς και η καθοδηγούμενη από τη θεωρία προσέγγιση, αποτελούν τα στοιχεία εκείνα που θα προλειάνουν το έδαφος για τη μελέτη και παρέμβαση στον ανθρώπινο ψυχισμό.¹⁵³

2.γ. Οι ενδεχόμενες μελλοντικές εξελίξεις στον τομέα της ψυχικής υγείας.

Βάσει των εκτιμήσεων που κάνουν οι ειδικοί, έως το 2025, περισσότεροι από πέντε δισεκατομμύρια άνθρωποι στον πλανήτη μας θα χρησιμοποιούν καθημερινά υπερ-ευρυζωνικά, με υπερ-υψηλές ταχύτητες και εξοπλισμένα με αισθητήρες, κινητά smartphone. Οι δυνατότητές αυτών θα είναι μακράν ανώτερες από αυτές των σημερινών iPhones, Androids και Blackberries. Η λήψη και εγκατάσταση των εφαρμογών ψυχικής υγείας, θα αποτελεί το πρώτο βήμα για την καταγραφή της τοποθεσίας, των δραστηριοτήτων και της γενικότερης συμπεριφοράς, δεδομένα που θα διατίθενται για την διεξαγωγή ερευνών, δοκιμών και πειραμάτων μέσω οθονών αφής, ασύρματων συνδέσεων, βιοαισθητήρων κ.ο.κ. Παρόλο που οι κινητές συσκευές δεν σχεδιάστηκαν για το σκοπό αυτό, οι συνεχώς αυξανόμενες δυνατότητές τους στην ανίχνευση στοιχείων και στη συλλογή μεγάλου όγκου δεδομένων από τεράστια δείγματα του παγκόσμιου πληθυσμού, αναμένεται να μεταμορφώσουν σε μεγάλο βαθμό

¹⁵¹ D'Alfonso, "AI in mental health," 114.

¹⁵² Ένα σύστημα Τ.Ν., αναλόγως του βαθμού ελέγχου που επιδέχεται από τον άνθρωπο, δύναται να βρίσκεται σε κατάσταση α) *Human off the loop*, όπου διαθέτει πλήρη αυτονομία, χωρίς τη δυνατότητα παρέμβασης από τον άνθρωπο, β) *Human in the loop*, όπου υπόκειται σε μεγάλο βαθμού έλεγχο από τον άνθρωπο, ο οποίος αποφασίζει τον τρόπο δράσης του συστήματος και γ) *Human on the loop*, όπου το σύστημα λειτουργεί με μεγαλύτερη αυτονομία, αλλά τη συνεχή επίβλεψη του ανθρώπου, που παρεμβαίνει όταν το κρίνει απαραίτητο. Βλ. Γεώργιος Κωστελέτος, *Τεχνητή Νοημοσύνη και Ηθική, Ενότητα 9^η: Έρωτας και Πόλεμος στην εποχή της Τεχνητής Νοημοσύνης* [πανεπιστημιακές σημειώσεις]. ΕΚΠΑ, Τμήμα Φιλοσοφίας, ΠΜΣ Φιλοσοφίας: Εφαρμοσμένη Ηθική, χειμερινό εξάμηνο 2022-2023, Αθήνα.

¹⁵³ Montag, Elhai, "A new agenda," 131.

τις αποκρίσεις και τις προσεγγίσεις στο κομμάτι του ανθρώπινου ψυχισμού.¹⁵⁴ Στο παραπάνω, συμβάλλει σημαντικά το γεγονός πως η ψυχική υγεία και η θεραπεία της δεν περιλαμβάνουν φυσικούς χειρισμούς και παρεμβάσεις στο ανθρώπινο σώμα (λ.χ. ψηλάφηση, χειρουργική επέμβαση κ.ο.κ.). Η ιδιαιτερότητά αυτή, σε σύγκριση με τη γενική υγεία και άλλους ιατρικούς τομείς, καθιστά την ψυχική υγεία εύφορο έδαφος για την εισχώρηση της Τ.Ν. και της τεχνολογίας γενικότερα.¹⁵⁵

Οι μακροπρόθεσμοι στόχοι αφορούν στη βελτίωση των κρίσιμων παραμέτρων στην παροχή υπηρεσιών ψυχικής υγείας με τη βοήθεια των εργαλείων της Τ.Ν. Ειδικότερα, οι μακροπρόθεσμες στρατηγικές επικεντρώνονται στην αποτελεσματικότερη διαχείριση του πληθυσμού των ασθενών, στη διευκόλυνση μέσω απομακρυσμένης πρόσβασης και διάγνωσης και στη σταδιακή μετατροπή της φροντίδας σε ένα πιο προληπτικό και εξατομικευμένο μοντέλο, πάντοτε με σταθερούς τους δείκτες της ποιότητας και της ικανοποίησης σε όλα τα στάδια της περίθαλψης και παράλληλα, προς αναζήτηση τρόπων για την εξοικονόμηση των πόρων του συστήματος υγείας.¹⁵⁶

Πιο συγκεκριμένα, η κακή διαχείριση της ροής των ασθενών επηρεάζει αρνητικά κυρίως τους ίδιους, ενώ όταν αναφερόμαστε σε χώρους όπως τα νοσοκομεία, δεν θα πρέπει να παραβλέπουμε τις περιπτώσεις εκείνες στις οποίες, λόγω υπερκόπωσης του προσωπικού, παρατηρείται το φαινόμενο της αύξησης των ιατρικών λαθών. Έτσι, για παράδειγμα, η επέκταση της χρήσης εφαρμογών (όπως οι SAFER¹⁵⁷ και Red2Green Bed Days¹⁵⁸), που αναγνωρίζουν το χρόνο αναμονής των ασθενών μέσα από οπτικούς αισθητήρες, δύναται να επηρεάσουν θετικά τους ίδιους τους πάσχοντες, το ανθρώπινο δυναμικό αλλά και τη συνολική ποιότητα της περίθαλψης.¹⁵⁹

Επιπλέον, η επιβεβαίωση των ήδη υπαρχόντων πορισμάτων μέσα από επαναληπτικές μελέτες σε μεγαλύτερα δείγματα πληθυσμού, θα μπορούσαν μελλοντικά να οδηγήσουν σε εργαλεία αυτοαναφοράς, όπως ερωτηματολόγια που θα συμπληρώνονται αυτόματα μέσα από τη χρήση smartphone και την ανάλυση των αισθητήρων. Ακόμη, σε προσεγγίσεις βαθιάς μηχανικής μάθησης και στο εγγύς μέλλον, η άντληση των ψηφιακών αποτυπωμάτων ενός ατόμου από ακόμη περισσότερες πηγές, όπως το αυτοκίνητο, το ντουζ, ο τρόπος πληκτρολόγησης στο pc, τα μη λεκτικά σήματα που θα αναγνωρίζονται από μια κάμερα, θα καταφέρνει να εντοπίζει μοτίβα σε μεγάλους όγκους δεδομένων και σταδιακά, να οδηγήσει σε ακόμη πιο αποδοτικά μοντέλα.¹⁶⁰

Παρομοίως, το ενδεχόμενο σύντομα να είμαστε σε θέση να συλλέγουμε βιολογικό υλικό προς ανάλυση (π.χ. κυτταρικό, ιστολογικό, ηλεκτροεγκεφαλογράφημα κ.ο.κ.), μέσω φορητών βιο-αισθητήρων και να το συνδυάζουμε με δεδομένα από άλλες πηγές, θα συμβάλλει στην ακόμη πιο ολοκληρωμένη και σφαιρική πλαισίωση της ψυχικής

¹⁵⁴ Geoffrey Miller, “The Smartphone Psychology Manifesto,” *Perspectives on Psychological Science* 7, no. 3 (May, 2012): 221, <https://doi.org/10.1177/1745691612441215>.

¹⁵⁵ Gooding, “Mapping the rise of digital mental health technologies,” 2.

¹⁵⁶ Fatema Mustansir Dawoodbhoy, Jack Delaney, Paulina Cecula, Jiakun Yu, Iain Peacock, Joseph Tan, Benita Cox, “AI in patient flow: applications of artificial intelligence to improve patient flow in NHS acute mental health inpatient units,” *Heliyon* 7, no. 5 (May 2021): 1-3, <https://doi.org/10.1016/j.heliyon.2021.e06993>.

¹⁵⁷ “The SAFER Patient Flow Bundle,” Fub NHS Stuff, accessed July 10, 2023, <https://fabnhsstuff.net/fab-stuff/28574-the-safer-patient-flow-bundle>.

¹⁵⁸ “The #last1000days from 3red2green,” Fub NHS Stuff, accessed July 10, 2023, <https://fabnhsstuff.net/fab-stuff/the-last1000days-from-red2green>.

¹⁵⁹ Dawoodbhoy et al., “AI in patient flow,” 1-2.

¹⁶⁰ Messner et al., “Insights,” 92.

νόσου.¹⁶¹ Για παράδειγμα, στο έργο τους οι Stopczynski et al. παρουσιάζουν το Smartphone Brain Scanner, ένα σύστημα που συνδέει μια φορητή συσκευή εγκεφαλογράφου, ο οποίος καταγράφει τα δεδομένα γύρω από την εγκεφαλική λειτουργία και μια κινητή συσκευή smartphone, στην οποία μεταφέρονται τα δεδομένα αυτά, με στόχο την απεικόνιση της εγκεφαλικής δραστηριότητας. Την όλη διαδικασία μπορεί να ενισχύσει η εμπλοκή της T.N. που δύναται να οδηγήσει σε επιπλέον συμπεράσματα μέσω διαγνωστικών αλγορίθμων ή και να ενισχύσει την επεξεργασία των δεδομένων μέσω ισχυρών αλγορίθμων επεξεργασίας, που θα απομονώσουν τους περιττούς θορύβους και ενδεχομένως, θα χρησιμοποιήσουν επιπλέον διαθέσιμες πηγές δεδομένων, όπως οι κινήσεις του κεφαλιού (με τη χρήση ενός γυροσκοπίου).¹⁶²

Ας φανταστούμε λοιπόν, έναν ασθενή με κλινικά διαγνωσμένη αγοραφοβία και διαταραχή πανικού. Σε μία τέτοια περίπτωση, η κινητή συσκευή smartphone θα μπορούσε να συλλέξει δεδομένα αυτοαναφοράς (από έρευνες μέσω τηλεφώνου), δεδομένα συμπεριφοράς όπως οι εκφράσεις του προσώπου, τις αποφευκτικές συμπεριφορές όπως η αποφυγή μιας τοποθεσίας που θα εντοπιστεί μέσω GPS και θα μπορούσε να υποδείξει την πηγή της απειλής, βιολογικά δεδομένα όπως ο καρδιακός παλμός ή η ποιότητα του ύπνου, που θα εξετάζουν το εάν η ύπαρξη κάποιας άλλης νόσου ίσως ευθύνεται για τις κρίσεις πανικού (λ.χ. η επιληπτική διαταραχή της κροταφικής λοβού) και δεδομένα από μεθόδους απεικόνισης και βιολογικών αναλύσεων (λ.χ. ηλεκτροεγκεφαλογράφημα). Τέτοιου είδους λεπτομερείς αναλύσεις μπορούν να προσφέρουν πολλές περισσότερες πληροφορίες σε προσωπικό και πληθυσμιακό επίπεδο, σε σχέση με τις κλινικές αξιολογήσεις που λαμβάνουν χώρα στο παρόν. Επιπλέον, θα παρέχουν την ευκαιρία στην Ψυχιατρική, να μελετήσει νέες διαστάσεις παθολογίας που, μόλις λίγα χρόνια πριν, ήταν σε μεγάλο βαθμό απρόσιτες.¹⁶³

Η λεπτομερής αυτή γνώση γύρω από την ψυχική κατάσταση ενός ατόμου επιτρέπει την παροχή της εξατομικευμένης φροντίδας και ειδικότερα, τη λήψη σωστών αποφάσεων σχετικά με την ψυχική του υγεία και την επιλογή ανάμεσα σε ψηφιακές παρεμβάσεις χαμηλών πόρων ή στην υψηλότερου επιπέδου παραδοσιακή φροντίδα, πριν από την έναρξη οποιασδήποτε θεραπείας. Έτσι, επιτυγχάνεται η εξοικονόμηση χρόνου και ενέργειας για το σύστημα υγείας καθώς και για τον ίδιο τον ασθενή, ο οποίος λαμβάνει εξ αρχής την κατάλληλη και εξατομικευμένη μορφή φροντίδας, κάτι που θα εμπόδιζε η δοκιμή διαφορετικών παρεμβάσεων μέχρι να βρεθεί η πιο αποτελεσματική.¹⁶⁴ Ήδη, διαβάζουμε αναφορές για συστήματα της T.N. υπό ανάπτυξη και δοκιμή, όπως αυτό που ανέπτυξαν επιστήμονες στις ΗΠΑ και το οποίο, μέσα από αλγορίθμους μηχανικής μάθησης που αναλύουν την εγκεφαλική δραστηριότητα ενός ανθρώπου με κατάθλιψη, θα είναι σε θέση να προβλέψει το εάν η φαρμακευτική αγωγή που του προτείνεται θα είναι αποτελεσματική.¹⁶⁵ Επιπλέον, η γνώση που βασίζεται σε δεδομένα θα μπορούσε

¹⁶¹ Torous, Onnela, Keshavan, “New dimensions and new tools to realize the potential of RDoC,” 2.

¹⁶² Arkadiusz Stopczynski, Carsten Stahlhut, Jakob Eg Larsen, Michael Kai Petersen, Lars Kai Hansen, “The Smartphone Brain Scanner: A Portable Real-Time Neuroimaging System,” *Plos One* 9, no. 2, e86733 (Feb., 2014): 1,9, <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0086733>.

¹⁶³ Torous, Onnela, Keshavan, “New dimensions and new tools to realize the potential of RDoC,” 2.

¹⁶⁴ Nicholas C. Jacobson, Matthew D. Nemesure, “Using Artificial Intelligence to Predict Change in Depression and Anxiety Symptoms in a Digital Intervention: Evidence from a Transdiagnostic Randomized Controlled Trial,” *Psychiatry Research* 295, 113618 (Jan., 2021): 4, <https://doi.org/10.1016/j.psychres.2020.113618>.

¹⁶⁵ «Η Τεχνητή νοημοσύνη μπορεί να εντοπίζει αν τα αντικαταθλιπτικά φάρμακα θα “δουλέψουν” ή όχι», Psychologynow.gr, προσπελάστηκε στις 22 Σεπτεμβρίου, 2023,

με τη σειρά της να ενισχύσει τις καθιερωμένες θεραπείες ή και να οδηγήσει σε παρεμβάσεις που θα βασίζονται αποκλειστικά στο Διαδίκτυο και στα κινητά, όπως η αποστολή εξατομικευμένων και αυτοματοποιημένων προτροπών σε ασθενείς, ή και στους ιατρούς τους στις περιπτώσεις επιδείνωσης των συμπτωμάτων ως ένδειξη κινδύνου υποτροπής.¹⁶⁶

Η συνολική αυτή μετατροπή, μέσω μεθόδων που επιστρατεύουν την Τ.Ν., θα μπορούσε να μετακινήσει τον τομέα της ψυχικής υγείας από την εξάρτηση στην αυτοαναφορά και την κλινική παρατήρηση, σε κριτήρια και μέτρα πιο αντικειμενικά.¹⁶⁷ Έτσι, ελευθερώνεται σημαντικός χώρος και χρόνος για εργασία πάνω σε ουσιαστικά θέματα, όπως οι δομημένες δια ζώσης συνεντεύξεις μεταξύ ειδικών και ασθενών, ή και κοινοτικές εφαρμογές¹⁶⁸ που σταδιακά θα αποσυνδεθούν από το σύστημα ψυχικής υγείας, όπως οι ομάδες αυτοβοήθειας μεταξύ ομότιμων που θα συντονίζονται από τους αλγορίθμους της Τ.Ν. και θα συντελέσουν στην παροχή υποστήριξης χαμηλού κόστους, αφού θα υλοποιούνται ανεξάρτητα από το οποιοδήποτε ίδρυμα ψυχικής υγείας.¹⁶⁹

Όλες οι παραπάνω εξελίξεις και ιδιαίτερα, η απεριόριστη καταγραφή προσωπικών δεδομένων, στην οποία και βασίζονται, εγείρουν σοβαρά θέματα ιδιωτικότητας και απορρήτου. Υπάρχει η ελπίδα ότι μέσω του *GDPR (General Data Protection Regulation)*,¹⁷⁰ η διαχείριση των δεδομένων αυτών σταδιακά θα βελτιωθεί και ταυτόχρονα, θα αναγνωριστεί το δικαίωμα ενός ανθρώπου στη διάθεση των προσωπικών του δεδομένων, στη γνώση του γύρω από το ποια δεδομένα συλλέγονται και που αποθηκεύονται, αλλά και στην επιθυμία του ανά πάσα στιγμή αυτά να διαγραφούν.¹⁷¹ Μάλιστα, με δεδομένο ότι οι τρόποι σκέψης γύρω από τη συναισθηματική κατάσταση ή την ψυχική διαταραχή έχουν αλλάξει πολύ κατά τις τελευταίες δεκαετίες και εκτός από τον τομέα της ψυχικής υγείας, έχουν επεκταθεί και σε άλλους τομείς (όπως η εκπαίδευση, το ποινικό σύστημα, παιδική προστασία κ.ο.κ.), η πιθανότητα διάχυσης των προσωπικών πληροφοριών σε άλλα πεδία (λ.χ. δικαιοσύνη, εκπαιδευτικές διαδικασίες) και επηρεασμού των διαδικασιών και των αποφάσεών τους με βάση αυτά (λ.χ. απόφαση ποινής, αναδοχή κ.λ.π.), δημιουργεί σημαντικές ηθικές προκλήσεις και προσθέτει ένα επιπλέον επίπεδο πολυπλοκότητας στα νομικά και ρυθμιστικά ζητήματα που προκύπτουν.¹⁷²

<https://www.psychologynow.gr/arthra-psyxologias/tecnologia/ilektronikes-efarmoges/13308-texniti-noimosyni-kathoristikiki-sti-therapeia-tis-katathlipsis.html>.

¹⁶⁶ Montag, Sindermann, Baumeister, “Digital phenotyping,” 20.

¹⁶⁷ Messner et al. “Insights,” 92.

¹⁶⁸ Για παράδειγμα, ο Εθνικός Οργανισμός της Κένυας λειτουργεί μία εικονική ομάδα υποστήριξης χρησιμοποιώντας την υπηρεσία μηνυμάτων What's Up, η οποία και λειτουργεί εκτός του συστήματος ψυχικής υγείας της χώρας. Βλ. World Health Organization, “Nairobi Mind Empowerment Peer Support Group - USP Kenya,” in the *Guidance on community mental health services: Promoting person-centred and rights-based approaches* (Geneva, WHO, 2021), 77, <file:///C:/Users/estia/Downloads/9789240025707-eng.pdf>.

¹⁶⁹ Gooding, “Mapping the rise of digital mental health technologies,” 4.

¹⁷⁰ *Γενικός Κανονισμός Προστασίας Δεδομένων (ΓΚΠΔ)*. Βλ. «Κανονισμός (ΕΕ) 2016/679 του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου της 27ης Απριλίου 2016», EUR-Lex. Access to European Union Law, προσπελάστηκε στις 15 Νοεμβρίου 2023, <https://eur-lex.europa.eu/legal-content/EL/TXT/?uri=celex%3A32016R0679>.

¹⁷¹ Messner et al. “Insights,” 92.

¹⁷² Gooding, “Mapping the rise of digital mental health technologies,” 1-2.

Διαπιστώνουμε με βάση τα παραπάνω πως, με την απεριόριστη γεωγραφική του εμβέλεια και τη δυνατότητα περαιτέρω επέκτασης, ο τομέας της Ψυχοπληροφορικής έχει τεράστιες δυνατότητες για ριζική αλλαγή στην παροχή υπηρεσιών και την ανάπτυξη νέων θεραπειών σε παγκόσμιο επίπεδο.¹⁷³ Το κομμάτι αυτό της ψυχικής υγείας, που μας απασχολεί στην παρούσα διατριβή, αφορά στις τηλεφωνικές γραμμές βοήθειας και συμβουλευτικής υποστήριξης. Όπως και στα περισσότερα πεδία της ψυχικής υγείας, έτσι και σε αυτό, σταδιακά η Τ.Ν. πραγματοποιεί την είσοδό της τα τελευταία χρόνια.

Στο τρίτο κεφάλαιο που ακολουθεί, αρχικά θα περιγράψουμε το ρόλο των τηλεφωνικών γραμμών βοήθειας και συμβουλευτικής υποστήριξης και έπειτα, θα διερευνήσουμε τη δυνατότητα της Τ.Ν. να αποκτήσει ενεργό και αποτελεσματικό ρόλο μέσα στο πεδίο αυτό. Είτε ως εργαλείο στα χέρια ενός εξειδικευμένου επαγγελματία (και πάλι ωστόσο, ως πράκτορας στη συμβουλευτική διαδικασία), είτε στο ρόλο του αυτόνομου συμβούλου, τα ερωτήματα που προκύπτουν αφορούν στο εάν η Τ.Ν. δύναται να κατακτήσει τα χαρακτηριστικά εκείνα που στοιχειοθετούν την αξία της Συμβουλευτικής και συντελούν στην ποιοτική και αποδοτική παροχή φροντίδας γύρω από τον ευάλωτο ανθρώπινο ψυχισμό.

¹⁷³ *Αυτόθι*, 2.

Κεφάλαιο 3^ο:

Η Τεχνητή Νοημοσύνη στις τηλεφωνικές γραμμές βοήθειας και συμβουλευτικής υποστήριξης.

3.α. Οι τηλεφωνικές γραμμές βοήθειας και συμβουλευτικής υποστήριξης.

Το 1978, στο πλαίσιο της Διεθνούς Διάσκεψης του Παγκοσμίου Οργανισμού Υγείας (Π.Ο.Υ.) στο Καζακστάν, ανακοινώθηκε η Διακήρυξη της Αλμά-Ατά, ένα γεγονός – σταθμός στην εξέλιξη της *Πρωτοβάθμιας Φροντίδας Υγείας* (ΠΦΥ), που την έθεσε στο επίκεντρο της παγκόσμιας κοινωνικής πολιτικής. Με τη Διακήρυξη αυτή, τα κράτη μέλη του Π.Ο.Υ. και της UNICEF δεσμεύτηκαν στην προσπάθεια επίτευξης της ισότητας στην υγεία, ένας στόχος ωστόσο, που η αναντιστοιχία μεταξύ λόγων και πράξεων άφησε στην αδράνεια. Το ενδιαφέρον στην Πρωτοβάθμια Φροντίδα Υγείας αναζωοπυρώθηκε, με την Παγκόσμια Έκθεση Υγείας του Π.Ο.Υ. (2008) να επανεκκινεί μια ευρεία στρατηγική με στόχο την καθιέρωση του καθολικού δικαιώματος στην υγεία.¹⁷⁴ Στην πρακτική του εφαρμογή, ο στόχος αυτός περιγράφεται στον ίδιο τον ορισμό της Πρωτοβάθμιας Φροντίδας, που είναι «η ουσιώδης φροντίδα υγείας, η οποία βασίζεται σε μεθόδους και τεχνολογία που είναι πρακτικές, επιστημονικά στέρεες και κοινωνικά αποδεκτές, και είναι καθολικά προσβάσιμη από τα άτομα και τις οικογένειες στην κοινότητα, με την πλήρη συμμετοχή τους και σε κόστος το οποίο η κοινότητα και η χώρα μπορεί να αντέξει...».¹⁷⁵

Ένα ευρύ φάσμα υπηρεσιών της Πρωτοβάθμιας Φροντίδας Υγείας καλύπτεται μέσα από την *Ηλεκτρονική Υγεία (eHealth)*, που ορίζεται από την Ευρωπαϊκή Επιτροπή, ως: «η χρήση σύγχρονων τεχνολογιών πληροφορίας και επικοινωνιών για την κάλυψη των αναγκών των πολιτών, των ασθενών, των επαγγελματιών του τομέα της υγείας, των παρόχων υγειονομικής περίθαλψης, καθώς και των υπευθύνων χάραξης πολιτικής».¹⁷⁶ Κομμάτι αυτής, αποτελεί ο τομέας της *κινητής υγείας (mHealth)*, μια «πρακτική ιατρικής και δημόσιας υγείας»,¹⁷⁷ που υποστηρίζεται από την τεχνολογία και κινητές συσκευές, όπως τα κινητά τηλέφωνα, συσκευές παρακολούθησης ασθενών κ.ο.κ.¹⁷⁸

Ανάμεσα στους τρόπους χρήσης της κινητής υγείας είναι και η επικοινωνία μεταξύ των ατόμων και των υπηρεσιών υγείας, με τις τηλεφωνικές γραμμές άμεσης βοήθειας και συμβουλευτικής υποστήριξης (*hotlines/helplines*) να κατατάσσονται στην πρώτη θέση παγκοσμίως (με ποσοστό 59%), βάσει της συχνότητας στη χρήση τους.¹⁷⁹

¹⁷⁴ Α. Φιλαλήθης, «Ορόσημα στην εξέλιξη της πρωτοβάθμιας φροντίδας υγείας», *Archives of Hellenic Medicine* 38, no. 5 (Νοέμβριος, 2021): 675, <https://www.mednet.gr/archives/2021-5/pdf/675.pdf>.

¹⁷⁵ *Αυτόθι*, 678.

¹⁷⁶ «eHealth - Ηλεκτρονική Υγεία», Ελληνική Δημοκρατία, Υπουργείο Υγείας, προσπελάστηκε στις 16 Σεπτεμβρίου, 2023, <https://www.moh.gov.gr/articles/ehealth/>.

¹⁷⁷ «mHealth - Υγεία εν Κινήσει», Ελληνική Δημοκρατία, Υπουργείο Υγείας, αναρτήθηκε στις 2 Σεπτεμβρίου, 2018, <https://www.moh.gov.gr/articles/ehealth/5688-mhealth>.

¹⁷⁸ *Αυτόθι*.

¹⁷⁹ Σοφία Ζέρβα, Ελένη Βατούγιου, Ιωάννης Αποστολάκης, «Κινητή Υγεία (M-health) και Πρωτοβάθμια Φροντίδα Υγείας. Τηλεφωνικές Γραμμές Άμεσης Βοήθειας και Συμβουλευτικής Υποστήριξης», *Το βήμα του Ασκληπιού* 19, Αρ. 2 (Απρίλιος - Ιούνιος 2020): 90-91, DOI: 10.5281/zenodo.3736786.

Το μεγαλύτερο πλεονέκτημα των τηλεφωνικών γραμμών βοήθειας, έγκειται στο ότι εξασφαλίζουν την άμεση και εύκολη πρόσβαση για τους πολίτες, καθώς το κόστος της κλήσης είναι κυρίως μηδενικό ή αντιστοιχεί σε αστικές χρεώσεις. Επιπλέον, πολλές από τις γραμμές λειτουργούν σε 24ωρη βάση και όλες τις ημέρες του χρόνου, ενώ στις περισσότερες των περιπτώσεων, το άτομο που καλεί έχει την επιλογή να διατηρήσει την ανωνυμία του, γεγονός που διευκολύνει ακόμη περισσότερο τη συζήτηση γύρω από ευαίσθητα προσωπικά ζητήματα ή την αναφορά περιστατικών.¹⁸⁰

Στις τηλεφωνικές γραμμές συμβουλευτικής υποστήριξης, που είναι και το θέμα της παρούσης, οι πολίτες απευθύνονται λόγω της ανάγκης τους να συνομιλήσουν με εξειδικευμένους επαγγελματίες ψυχικής υγείας και προκειμένου να λάβουν κατευθύνσεις για θέματα που τους αφορούν. Βάσει της διαφορετικότητας των καλούντων και της ποικιλομορφίας των αναγκών που εκφράζονται, οι γραμμές κατηγοριοποιούνται αναλόγως των ατομικών χαρακτηριστικών (όπως η ηλικία, το φύλο), της παρούσας κατάστασης/ιδιότητας τους (πρόσφυγας, γονέας), του αιτήματος που διατυπώνεται (π.χ παροχή συμβουλών, πληροφορίες) ή και βάσει του συγκεκριμένου θέματος ψυχικής υγείας για το οποίο κάλεσαν (λ.χ. εξάρτηση, κατάθλιψη κ.ο.κ.). Επιπλέον, η κλήση σε μία γραμμή βοήθειας μπορεί να αφορά σε κάποιο συγκεκριμένο και επείγον περιστατικό (λ.χ. βία, εξαφάνιση, αυτοκτονικός ιδεασμός), κάτι που συχνά απαιτεί την περαιτέρω κινητοποίηση και μετά το τέλος της συνομιλίας.

Έτσι ενδεικτικά, στην Ελλάδα συγκεκριμένα, οι υπάρχουσες τηλεφωνικές γραμμές, που υλοποιούνται αυτή τη στιγμή από ιδιωτικούς ή κρατικούς φορείς, απευθύνονται μεταξύ άλλων σε ανηλίκους (λ.χ. Γραμμή 115 25-Μαζί για το Παιδί,¹⁸¹ SOS 1056-Χαμόγελο του παιδιού¹⁸²), παρέχοντας συμβουλευτική υποστήριξη σε παιδιά, εφήβους και τις οικογένειές τους, σε ενήλικες που χρήζουν ψυχολογικής στήριξης και συμβουλευτικής για θέματα ψυχικής υγείας όπως το άγχος και οι κρίσεις πανικού (λ.χ. τηλεφωνική Γραμμή Ψυχοκοινωνικής Υποστήριξης 10306,¹⁸³ Τηλεφωνική Γραμμή Κοινωνικής Βοήθειας 197 του Εθνικού Κέντρου Κοινωνικής Αλληλεγγύης¹⁸⁴), σε άτομα με εξαρτήσεις (π.χ. 1031 Τηλεφωνική Γραμμή SOS του ΟΚΑΝΑ,¹⁸⁵ Γραμμή βοήθειας 1145 και 1114 του ΚΕΘΕΑ¹⁸⁶), καθώς και σε ασθενείς με χρόνια προβλήματα

¹⁸⁰ *Αυτόθι*, 89.

¹⁸¹ «Γραμμή 115 25 της Ένωσης Μαζί για το Παιδί», Μαζί για το Παιδί, προσπελάστηκε στις 16 Σεπτεμβρίου, 2023, <https://maziwiatopaidi.gr/en/articles/grammi-115-25-tis-enosis-mazi-gia-to-paidi>.

¹⁸² «Γραμμή SOS 1056 για Παιδιά, Έφηβους και Γονείς», Ελληνική Δημοκρατία, Υπουργείο Υγείας, προσπελάστηκε στις 16 Σεπτεμβρίου, 2023, <https://www.moh.gov.gr/articles/citizen/xrhsima-thlefwna-amp-dieythynseis/199-grammh-sos-1056-gia-paidia-efhboys-kai-goneis>.

¹⁸³ «Τηλεφωνική Γραμμή Ψυχοκοινωνικής Υποστήριξης 10306», Ελληνική Δημοκρατία, Υπουργείο Υγείας, προσπελάστηκε στις 16 Σεπτεμβρίου, 2023, <https://www.moh.gov.gr/articles/citizen/xrhsima-thlefwna-amp-dieythynseis/10714-thlefwnikh-grammh-psyxokoinwnikhs-yposthriksis-10306>.

¹⁸⁴ «197 Εθνικό Κέντρο Κοινωνικής Αλληλεγγύης», Ελληνική Δημοκρατία, Υπουργείο Υγείας, προσπελάστηκε στις 16 Σεπτεμβρίου, 2023, <https://www.moh.gov.gr/articles/citizen/xrhsima-thlefwna-amp-dieythynseis/201-197-ethniko-kentro-koinwnikhs-allhleggyhs>.

¹⁸⁵ «1031 Τηλεφωνική Γραμμή SOS του ΟΚΑΝΑ», Ελληνική Δημοκρατία, Υπουργείο Υγείας, προσπελάστηκε στις 16 Σεπτεμβρίου, 2023, <https://www.moh.gov.gr/articles/citizen/xrhsima-thlefwna-amp-dieythynseis/197-1031-thlefwnikh-grammh-sos-toy-okana>.

¹⁸⁶ «1145 Τηλεφωνική Γραμμή Ψυχολογικής Υποστήριξης ΙΘΑΚΗ - ΚΕ.Θ.Ε.Α.», Ελληνική Δημοκρατία, Υπουργείο Υγείας, προσπελάστηκε στις 16 Σεπτεμβρίου, 2023, [https://www.moh.gov.gr/articles/citizen/xrhsima-thlefwna-amp-dieythynseis/204-1145-thlefwnikh-grammh-psyxologikhs-yposthriksis-ithakh-ke-th-e-a](https://www.moh.gov.gr/articles/citizen/xrhsima-thlefwna-amp-dieythynseis/204-1145-thlefwnikh-grammh-psyxologikhs-yposthriksis-ithakh-ke-th-e-a;); «Γραμμή Βοήθειας 1114», Κέντρο Θεραπείας Εξαρτημένων Ατόμων ΚΕΘΕΑ, προσπελάστηκε στις 16 Σεπτεμβρίου, 2023, <https://www.kethea.gr/kethea/therapeftika-programmata/grammi-voitheias-1114/>.

υγείας και τους φροντιστές τους (210 6401200 Ελληνική Αντικαρκινική Εταιρεία,¹⁸⁷ Γραμμή Βοήθειας για την Άνοια στο 1102 της Εταιρείας Alzheimer Αθηνών¹⁸⁸). Επιπλέον, ορισμένες γραμμές εξειδικεύονται στην διαχείριση επειγόντων και σοβαρών περιστατικών, όπως ο αυτοκτονικός ιδεασμός (Γραμμή Παρέμβασης για την Αυτοκτονία της ΜΚΟ Κλίμακα¹⁸⁹) ή οι εξαφανίσεις (Ευρωπαϊκή Γραμμή για τα Εξαφανισμένα Παιδιά 116000,¹⁹⁰ Εθνική Τηλεφωνική Γραμμή SOS 1065¹⁹¹).

Όλες οι τηλεφωνικές γραμμές συμβουλευτικής υποστήριξης στελεχώνονται από εξειδικευμένο προσωπικό (λ.χ. ψυχολόγους, κοινωνικούς λειτουργούς), που αναλόγως της αιτίας της κλήσης, δύναται να παρέχει συμβουλευτική υποστήριξη, πληροφορίες και κατευθύνσεις, να παραπέμπει σε άλλους φορείς και υπηρεσίες ή να κινητοποιεί τις αρμόδιες αρχές (λ.χ. Εισαγγελία, Αστυνομία), με στόχο την άμεση και επί τόπου παρέμβαση, όταν αυτό κρίνεται απαραίτητο για την προστασία ενός ανθρώπου που βρίσκεται σε κίνδυνο.

Χωρίς αμφιβολία, η προσωπική επαφή με εξειδικευμένους επαγγελματίες σε μία κρίσιμη στιγμή, δύναται να μειώσει το άγχος του ανθρώπου που καλεί και να τον ενθαρρύνει, καθοδηγώντας τον σχετικά με το θέμα που τον οδήγησε στο να καλέσει. Τα δε στοιχεία της ανωνυμίας και του απορρήτου, καθώς και η μη καταγραφή των κλήσεων, αποτελούν ένα πρόσφορο έδαφος, ώστε το άτομο να εκφραστεί ελεύθερα και να μοιραστεί ανησυχίες και ζητήματα που του προκαλούν φόβο, αμηχανία ή και ντροπή, όταν αυτά θεωρούνται απαγορευμένα και ανεπιθύμητα από την κοινωνία. Επιπλέον, το γεγονός ότι το άτομο έχει τον απόλυτο έλεγχο της κλήσης, μιας και μπορεί να την τερματίσει όποτε το επιθυμεί, είναι ένα ακόμη σημαντικό πλεονέκτημα που διευκολύνει την απόφαση του να ζητήσει βοήθεια.¹⁹² Ωστόσο, το ίδιο αποτελεί μία σημαντική πρόκληση για τους επαγγελματίες που ανταποκρίνονται σε μία κλήση, μιας και σε περιορισμένο χρόνο συνομιλίας λίγων λεπτών, καλούνται να κατανοήσουν μια κατάσταση, να οριοθετήσουν τη συζήτηση στο υπάρχον θέμα και ειδικότερα στις περιπτώσεις κινδύνου (π.χ. αυτοκτονικός ιδεασμός) να κρατήσουν το συνομιλητή τους στη γραμμή.

Για το λόγο αυτό, το προσωπικό που στελεχώνει μια γραμμή βοήθειας θα πρέπει να εκπαιδεύεται συνεχώς, τόσο πάνω στους τρόπους χειρισμού μιας κλήσης, όσο και στη νομοθεσία και στις διαδικασίες που θα πρέπει να ακολουθούνται σε κάθε περίπτωση. Γενικότερα όμως, θα πρέπει να είναι εξειδικευμένο εξ αρχής στην συμβουλευτική διαδικασία και να έχει κατακτήσει τις δεξιότητες ενός συμβούλου, όπως η

¹⁸⁷ «210 6401200 Ελληνική Αντικαρκινική Εταιρεία», Ελληνική Δημοκρατία, Υπουργείο Υγείας, προσπελάστηκε στις 16 Σεπτεμβρίου, 2023, <https://www.moh.gov.gr/articles/citizen/xrhsima-thlefwna-amp-dieythynseis/202-210-6401200-ellhnikh-antikarkinikh-etairaia>.

¹⁸⁸ «Γραμμή Βοήθειας για την Άνοια 1102», *Εταιρεία Alzheimer Αθηνών*, προσπελάστηκε στις 16 Σεπτεμβρίου, 2023, <https://alzheimeraathens.gr/helpline/>.

¹⁸⁹ «Εάν σκέφτεστε την αυτοκτονία», Κλίμακα, προσπελάστηκε στις 16 Σεπτεμβρίου, 2023, <http://suicide-help.gr/ean-skefteste-tin-autoktonia/>.

¹⁹⁰ «Ευρωπαϊκή Γραμμή για τα Εξαφανισμένα Παιδιά 116000», Το χαμόγελο του παιδιού, προσπελάστηκε στις 16 Σεπτεμβρίου, 2023, <https://www.hamogelo.gr/gr/el/116000/>.

¹⁹¹ «Η Εθνική Τηλεφωνική Γραμμή SOS 1065», Γραμμή Ζωής – Κοινοφελές Μη Κερδοσκοπικό Σωματείο Τηλεειδοποίησης και Βοήθειας στο Σπίτι, προσπελάστηκε στις 16 Σεπτεμβρίου, 2023, <http://www.lifelinehellas.gr/%cf%80%ce%bf%ce%b9%ce%bf%ce%b9-%ce%b5%ce%af%ce%bc%ce%b1%cf%83%cf%84%ce%b5/>.

¹⁹² Ζέρβα, Βατούγιου, Αποστολάκης, «Κινητή Υγεία (M-health) και Πρωτοβάθμια Φροντίδα Υγείας», 93.

ενσυναίσθηση, η ενεργητική ακρόαση, η αποδοχή και η ανατροφοδότηση, για τις οποίες μιλήσαμε στην Ενότητα 1.β. της παρούσας διατριβής.

Με την είσοδο της Τ.Ν. στον τομέα της Συμβουλευτικής, τα ερωτήματα που γεννώνται αναφορικά με την αποτελεσματικότητα που θα μπορέσει να έχει είναι πολλά και κυρίως, έγκεινται στο εάν ένα σύστημα Τ.Ν. και στην προκειμένη, ένα chatbot ομιλίας για την ψυχική υγεία, θα μπορέσει ουσιαστικά να ανταποκριθεί στο νόημα του ίδιου του όρου *συμβουλευτική*, και στο εννοιολογικό περιεχόμενο των λέξεων που τον συνθέτουν (*συν + βουλεύομαι*). Επιπλέον, το εάν στο πλαίσιο των συστημάτων Μηχανικής Μάθησης ή των παραδοσιακών συστημάτων της GOFAI (*Good Old Fashioned AI systems*), θα μπορούσε να αποδοθεί επαρκώς το περιεχόμενο της εκπαίδευσης ενός συμβούλου, ώστε αυτά να καταφέρουν να αναπτύξουν τις αντίστοιχες δεξιότητες, που θα τα καταστήσουν ικανά να ανταποκριθούν στις απαιτήσεις μιας συνομιλίας.

Το κύριο ερώτημα που ανακύπτει, αφορά στο εάν ένα σύστημα δύναται να *συν + βουλεύεται*, δηλ. να σκέπτεται, να προβαίνει σε σχέδια και να αποφασίζει μαζί με τον καλόντα.¹⁹³ Πέραν τούτου, το εάν θα μπορούσε να επιδείξει τις σημαντικές δεξιότητες που κατέχει ένας σύμβουλος ενόσω επικοινωνεί είτε σε λεκτικό, είτε σε μη λεκτικό επίπεδο. Αν και μια συμπεριφορά που επιδεικνύει ανατροφοδότηση ή την ενεργητική ακρόαση, μέσα από διαδικασίες όπως η λεκτική ακολουθία ή η σιωπή στις στιγμές που απαιτούνται κατά την τηλεφωνική συνομιλία, φαίνεται να είναι εφικτό να προγραμματιστεί, προκύπτει το ζήτημα του εάν δεξιότητες, όπως η ενσυναίσθηση και η αυτο-αποκάλυψη, που προϋποθέτουν αν μη τι άλλο την ανθρώπινη εμπειρία και τα συναισθήματα, θα διατηρήσουν την αξία τους εάν απλώς προσομοιωθούν από ένα τεχνητό σύστημα που θα προγραμματιστεί αναλόγως. Το ίδιο ισχύει και για έννοιες όπως η αυθεντικότητα και η ειλικρίνεια, ως χαρακτηριστικά της προσωπικότητας ενός συμβούλου, ή η ενθάρρυνση και ο σεβασμός στην προσωπικότητα του ανθρώπου που απευθύνεται σε αυτόν, ως συστατικό της θεραπευτικής σχέσης που αναπτύσσεται μεταξύ τους.

Ακόμη και αν οι παραπάνω συμβουλευτικές δεξιότητες ενσωματωθούν επιτυχώς σε ένα σύστημα της Τ.Ν. στην προσπάθεια εδραίωσης της θεραπευτικής σχέσης, ένα σημαντικό θέμα προς συζήτηση αφορά το εάν αυτού του είδους η σχέση, επαρκεί ώστε να επιφέρει την επιθυμητή αποτελεσματικότητα στη θεραπευτική διαδικασία. Τη στιγμή που τα ανωτέρω χαρακτηριστικά δε θα αντιστοιχούν παρά σε αντικείμενα προσομοίωσης από ένα σύστημα της Τ.Ν., είναι άξιο προβληματισμού το γεγονός ότι η απλή μίμηση της συμπεριφοράς ενός συμβούλου δεν συνεπάγεται τη δημιουργία μιας γνήσιας θεραπευτικής σχέσης με τον άνθρωπο. Προκύπτει λοιπόν, ένα θέμα εξαπάτησης (ή αυταπάτης) των ασθενών και ο κίνδυνος πως η συναισθηματική επένδυση στην αμφίβολη σχέση με ένα σύστημα της Τ.Ν. ή η ψευδαίσθηση πως αυτό νοιάζεται αληθινά για αυτούς, θα επιφέρει ένα βαρύ συναισθηματικό κόστος μετά το τέλος της θεραπείας, εκεί όπου θα επέλθει η συνειδητοποίηση πως αυτή η σχέση ήταν ψευδής και μονόδρομη.¹⁹⁴

Τα ερωτήματα αυτά δεν διαφοροποιούνται κατά πολύ ανάμεσα στις περιπτώσεις εκείνες κατά τις οποίες ένα chatbot ομιλίας στην ψυχική υγεία λειτουργεί αυτόνομα (*human off the loop*) ή με κάποιο βαθμό ελέγχου από τον άνθρωπο (*human on/in the*

¹⁹³ Μαλικιώση - Λοϊζου, «Η Συμβουλευτική Ψυχολογία στην Ελλάδα σήμερα», 267.

¹⁹⁴ Γεώργιος Κωστελέτος, *Τεχνητή Νοημοσύνη και Ηθική, Ενότητα 10^η: Άνθρωπος και μηχανή: Από τους ψυχολογικούς μηχανισμούς στα κοινωνικά και ηθικά ζητήματα* [πανεπιστημιακές σημειώσεις]. ΕΚΠΑ, Τμήμα Φιλοσοφίας, ΠΜΣ Φιλοσοφίας: Εφαρμοσμένη Ηθική, χειμερινό εξάμηνο 2022-2023, Αθήνα.

Ioop)¹⁹⁵, που στην περίπτωση μας θα είναι κάποιος εξειδικευμένος σύμβουλος ψυχικής υγείας. Στην πρώτη περίπτωση ωστόσο, οι απαιτήσεις που δημιουργούνται είναι πολύ μεγαλύτερες, κυρίως εκεί όπου οι συνέπειες ενός λάθους χειρισμού μπορεί να είναι ανυπολόγιστες, ενώ στη δεύτερη, ένα σημαντικό θέμα προς συζήτηση αφορά στη θεραπευτική σχέση μεταξύ συμβούλου και ωφελούμενου και στο εάν η ποιότητα αυτής θα μπορούσε να βελτιωθεί ή να πληγεί από την παρουσία ενός συστήματος T.N.

Τα ερωτήματα αυτά θα τα διερευνήσουμε στην ενότητα που ακολουθεί, όπου θα δούμε τόσο τα πιθανά οφέλη που επιφέρει η χρήση της T.N. στις τηλεφωνικές γραμμές συμβουλευτικής υποστήριξης, όσο και τις προκλήσεις ή τους προβληματισμούς που αυτή δημιουργεί.

3.β. Οφέλη και προβληματισμοί γύρω από την ενσωμάτωση της T.N. στις τηλεφωνικές γραμμές βοήθειας και συμβουλευτική υποστήριξη.

Όπως εξηγήσαμε στην ενότητα που προηγήθηκε, οι τηλεφωνικές γραμμές βοήθειας παρέχουν συμβουλευτική υποστήριξη στους ανθρώπους που αποφασίζουν να απευθυνθούν σε αυτές, δια στόματος ειδικών επαγγελματιών, οι οποίοι δύνανται να κατευθύνουν πάνω σε θέματα που αφορούν στην ψυχική υγεία ή και να κινητοποιούν την παρέμβαση, όποτε αυτό κρίνεται απαραίτητο. Διερευνώντας την ενδεχόμενη ενσωμάτωση της T.N. στον τομέα αυτό της κινητής υγείας (mHealth), θα πρέπει να προσεγγίσουμε την περίπτωση αυτή, ξεκινώντας από την παρουσία ενός συνομιλητικού chatbot ψυχικής υγείας, που αντί ενός ανθρώπινου επαγγελματία συμβούλου, θα μπορέσει να παρέχει, σε προφορικό λόγο, συμβουλευτική υποστήριξη στην άλλη άκρη της τηλεφωνικής γραμμής.

Βέβαια, η δυνατότητα της προφορικής ομιλίας, αν και σημαντική (όπως θα δούμε αργότερα), αποτελεί μια τεχνική λεπτομέρεια που προσδίδεται επιπλέον σε ένα chatbot συνομιλίας. Πολλές φορές η ικανότητα στην παροχή υψηλής ποιότητας υπηρεσιών συμβουλευτικής, είναι ανεξάρτητη από το τεχνικό αυτό χαρακτηριστικό, που αφορά περισσότερο στον τρόπο επικοινωνίας, παρά στην ουσία της υπηρεσίας που παρέχεται. Για το λόγο αυτό στις σελίδες που ακολουθούν, πολλές φορές γίνονται αναφορές σε άρθρα που εστιάζουν σε chatbot που βασίζονται στην επικοινωνία μέσω κειμένου και όχι απαραίτητα, στην προφορική επικοινωνία. Άλλωστε, η δυνατότητα φωνητικής αλληλεπίδρασης και αλληλεπίδρασης μέσω κειμένου, πολύ συχνά συνυπάρχουν κι έτσι, η επικοινωνία μπορεί να διεξάγεται είτε σε γραπτό, είτε σε προφορικό λόγο (είτε και με τους δύο τρόπους ταυτόχρονα).

Στην υπάρχουσα αρθρογραφία, συναντούμε έναν αριθμό chatbot που χρησιμοποιούνται στον τομέα της ψυχικής υγείας. Ήδη, έχουμε αναφερθεί στα chatbot Woebot, Wysa και Tess και στα ενθαρρυντικά αποτελέσματα της χρήσης τους, όπως διαφάνηκαν μέσω συναφών ερευνών.¹⁹⁶ Μία αξιοσημείωτη ένταση στην ανάπτυξη των chatbot για την ψυχική υγεία, άρχισε να σημειώνεται κατά την περίοδο της πανδημίας του κορονοϊού. Στην περίοδο αυτή, η αναγκαστική απομόνωση εντός της οικίας και η έντονη ανησυχία που δημιούργησε η πρωτόγνωρη αυτή συνθήκη, κατέστησαν

¹⁹⁵ Βλ. υπ. ¹⁵⁰.

¹⁹⁶ Βλ. Ενότητα 1. β. της παρούσας εργασίας, με τίτλο: «Η ενσωμάτωση της T.N. στο πεδίο της Συμβουλευτικής σήμερα».

επιτακτική την ανάγκη διαχείρισης της συνεχώς εναλλασσόμενης πληροφορίας και εξασφάλισης της εικοσιτετράωρης πρόσβασης σε υπηρεσίες ψυχικής υγείας και συμβουλευτικής.

Έτσι, για παράδειγμα στην Κίνα, τα chatbot Psybot,¹⁹⁷ Xiaoi και Xiaolv, χρησιμοποιήθηκαν από τις τοπικές κυβερνητικές αρχές, με στόχο την ενημέρωση και τη συναισθηματική αποφόρτιση των κατοίκων, τη διαχείριση του άγχους και της ανησυχίας, καθώς και τη γενικότερη αλληλεπίδραση και συμπαράσταση κατά τη δύσκολη περίοδο της καραντίνας.¹⁹⁸ Επίσης στον Καναδά, το chatbot Chloe χρησιμοποιήθηκε από τις δημόσιες υπηρεσίες υγείας, που είδαν τον αριθμό των εισερχομένων ερωτημάτων του κοινού να εκτοξεύεται μετά την ανακοίνωση των πρώτων κρουσμάτων στη χώρα. Η λειτουργία του Chloe βασίστηκε στην χρήση εργαλείων επεξεργασίας της φυσικής γλώσσας (Natural language processing - NLP), σε συνδυασμό με προηγμένους αλγόριθμους, που εκπαιδεύτηκαν στο να παρέχουν γραπτές απαντήσεις που μοιάζουν με ανθρώπινες, ενώ όταν απαιτείτο, ήταν δυνατή η προώθηση της συνομιλίας στους εργαζομένους των υπηρεσιών αυτών.¹⁹⁹

Πέρα από τα παραπάνω, το chatbot CovBot²⁰⁰ χρησιμοποιήθηκε στη διάρκεια της πανδημίας από τις τοπικές υγειονομικές αρχές της Γερμανίας, με στόχο την ανακούφιση του προσωπικού της υπηρεσίας τηλεφωνικής εξυπηρέτησης. Μέσα από τη χρήση προ-εκπαιδευμένων μοντέλων μηχανικής μάθησης της Google Cloud για την αναγνώριση και τη σύνθεση φωνής, το CovBot παρείχε στο κοινό προκαθορισμένες απαντήσεις (που συνέθετε το προσωπικό) για κοινές ανησυχίες σχετικά με τη νόσο του Covid-19.²⁰¹ Όλες οι κλήσεις συνδέονταν απευθείας με το CovBot, ανεξάρτητα από τη διαθεσιμότητα του προσωπικού. Όταν το αίτημα του καλούντος δεν γινόταν αντιληπτό, ζητείτο από αυτόν να το επαναλάβει και εάν και πάλι δεν ήταν κατανοητό, τότε η κλήση προωθείτο σε κάποιον από τους συμβούλους της γραμμής.²⁰²

Το τελευταίο αυτό παράδειγμα, αποτελεί το πιο αντιπροσωπευτικό για εμάς και το πιο συναφές με το θέμα της παρούσης εργασίας. Το CovBot ενσωματώθηκε στις δημόσιες γραμμές τηλεφωνικής υποστήριξης των υγειονομικών υπηρεσιών της Γερμανίας, δίνοντας έτσι την ευκαιρία, η ενσωμάτωση αυτή της T.N., να αξιολογηθεί στην πράξη. Τα άτομα που απευθύνθηκαν στις τηλεφωνικές γραμμές της Γερμανίας, χρησιμοποίησαν μόνο μια τηλεφωνική συσκευή (σταθερή, κινητή, καρτοτηλέφωνο κ.ο.κ.) και για τη συνομιλία, δεν απαιτούνταν χαρακτηριστικά, όπως η σύνδεση στο

¹⁹⁷ “Psybot. Therapeutic techniques are in your pocket now!” Psybot, accessed September 28, 2023, <http://www.psybotapp.com/en/index.html>.

¹⁹⁸ Yonghan Zhu, Rui Wang, Chengyan Pu, “I am chatbot, your virtual mental health adviser. What drives citizens’ satisfaction and continuance intention toward mental health chatbots during the COVID-19 pandemic? An empirical study in China.” *Digital Health* 8 (March, 2022): 2-3, <https://journals.sagepub.com/doi/full/10.1177/20552076221090031>.

¹⁹⁹ Sophia Siedlikowski, Louis-Philippe Noël, Stephanie Anne Moynihan, Marc Robin, “Chloe for COVID-19: Evolution of an Intelligent Conversational Agent to Address Infodemic Management Needs During the COVID-19 Pandemic,” *Journal of Medical Internet Research* 23, no. 9, e27283 (Jan., 2021): 1-2, <https://www.jmir.org/2021/9/e27283>.

²⁰⁰ “Forschungsprojekte und Aktivitäten mit COVID-19 Bezug,” Charité – Universitätsmedizin Berlin, accessed September 28, 2023, https://iph.charite.de/forschung/aktivitaeten_mit_covid_19_bezug/.

²⁰¹ Vanessa Voelskow, Claudia Meßner, Tobias Kurth, Amelie Busam, Toivo Glatz, Natalie Ebert, “Prospective mixed-methods study evaluating the potential of a voicebot (CovBot) to relieve German health authorities during the COVID-19 infodemic,” *Digital Health* 9 (May, 2023): 1, <https://journals.sagepub.com/doi/full/10.1177/20552076231180677>.

²⁰² *Αυτόθι*, 4.

διαδίκτυο ή η διαθεσιμότητα ηλεκτρολογίου. Το ίδιο ισχύει και για τις υπάρχουσες τηλεφωνικές γραμμές βοήθειας και συμβουλευτικής υποστήριξης που υλοποιούνται στη χώρα μας, κάποιες εκ των οποίων αναφέραμε και στην προηγούμενη ενότητα (λ.χ. 115 25-Μαζί για το Παιδί, 1114 - ΚΕΘΕΑ κ.ο.κ.). Αξίζει να σημειωθεί πως, το ConBot δεν αντικατέστησε ολοκληρωτικά τους ανθρώπους συμβούλους, κατάφερε ωστόσο να διαχειριστεί ένα μεγάλο κομμάτι της λειτουργίας της συγκεκριμένης γραμμής. Η περίοδος της πανδημίας αποτέλεσε ένα πρόσφορο έδαφος για τη δοκιμή αυτή και ίσως, σηματοδότησε την αρχή μιας νέας εποχής, όπου η επιστράτευση της T.N. στις γραμμές βοήθειας και συμβουλευτικής υποστήριξης μπορεί να υποστηρίξει σημαντικά το ανθρώπινο δυναμικό ή και σταδιακά να το αντικαταστήσει.²⁰³

Όπως και οι τηλεφωνικές γραμμές συμβουλευτικής υποστήριξης που στελεχώνονται από ανθρώπους επαγγελματίες, έτσι και ένα chatbot συνομιλίας μπορεί να είναι προσβάσιμο όλες τις ημέρες και ώρες, διευκολύνοντας με τον τρόπο αυτό, την παροχή καθοδήγησης σε ανθρώπους που εργάζονται κατά τη διάρκεια της ημέρας, ζουν σε απομακρυσμένες περιοχές ή επιθυμούν να μιλήσουν σε ώρες πιο ασυνήθιστες, όπως οι βραδινές. Επιπλέον, όπως είδαμε στην προηγούμενη ενότητα, τα άτομα που καλούν στις υπάρχουσες γραμμές απολαμβάνουν την ανωνυμία, καθώς και το γεγονός ότι οι υπηρεσίες παρέχονται δωρεάν, μιας και οι γραμμές αυτές, αν και πολλές φορές υλοποιούνται από ιδιωτικούς φορείς, δεν παύουν να αποτελούν στρατηγικά εργαλεία της κοινωνικής πολιτικής που εφαρμόζεται σε επίπεδο χώρας. Τις δυνατότητες της δωρεάν διαθεσιμότητας και της ανωνυμίας, παρείχε και το προαναφερόμενο ConBot που χρησιμοποιήθηκε στη Γερμανία. Επιπλέον, ήταν διαθέσιμο κατά τη διάρκεια του Σαββατοκύριακου, αλλά και εκτός του ωραρίου των υπηρεσιών, κάτι που επισημάνθηκε θετικά από τους περισσότερους χρήστες.²⁰⁴ Αναμένουμε λοιπόν, ότι η ενδεχόμενη χρήση της T.N. στο πεδίο αυτό του mHealth, δεν θα επέφερε καμία αλλαγή πάνω στα θέματα της ανωνυμίας, της διαθεσιμότητας και του κόστους, καθότι ως εργαλείο πρόκειται να ενσωματωθεί και όχι να αντικαταστήσει μια ήδη υπάρχουσα πολιτική.

Τα χαρακτηριστικά αυτά είναι ιδιαίτερα σημαντικά. Αν οι υπηρεσίες ψυχικής υγείας που παρέχονται από ένα chatbot είναι αποτελεσματικές, τότε μπορούν να αποτελέσουν την πιο οικονομική και εύκολα προσβάσιμη γέφυρα για την εξασφάλιση της επαγγελματικής βοήθειας. Ενώσω οι άνθρωποι χρησιμοποιούν αυτούς τους πόρους, μπορούν να αναπτύξουν δεξιότητες για την αναγνώριση της στιγμής εκείνης, όπου η επιπλέον υποστήριξη από κάποιον ειδικό είναι απαραίτητη και να γίνουν πιο ανοιχτοί στη συνομιλία με έναν ανθρώπινο θεραπευτή. Επίσης, η χρήση εφαρμογών ψυχικής υγείας μπορεί να προτείνεται από τον ίδιο τον επαγγελματία, ως ενίσχυση της θεραπείας που λαμβάνει χώρα ή ως ενδιάμεση υποστήριξη κατά τη διάρκεια της λίστας αναμονής.²⁰⁵ Ενέχεται ωστόσο παράλληλα, ο κίνδυνος οι άνθρωποι να εμφανίσουν εξάρτηση σε αυτού του είδους την υποστήριξη, λόγω της 24/7 και δωρεάν διαθεσιμότητάς τους. Έτσι, είτε να μην απευθυνθούν ποτέ σε κάποιον ειδικό, είτε να

²⁰³ *Αυτόθι*, 1.

²⁰⁴ *Αυτόθι*, 6.

²⁰⁵ Kira Kretzschmar, Holly Tyroll, Gabriela Pavarini, Arianna Manzini, Pina Singh, “Can Your Phone Be Your Therapist? Young People’s Ethical Perspectives on the Use of Fully Automated Conversational Agents (Chatbots) in Mental Health Support,” *Biomedical Informatics Insights* 11 (March, 2019): 2, <https://journals.sagepub.com/doi/10.1177/1178222619829083>.

εμφανίσουν συμπτώματα στέρησης και ανησυχίας όταν για κάποιο λόγο οι γραμμές αυτές δεν θα είναι πλέον διαθέσιμες (π.χ. τεχνικά προβλήματα, λήξη πανδημίας).²⁰⁶

Επιπρόσθετα, ένα σημαντικό στοιχείο στο οποίο ένα συνομιλητικό chatbot ψυχικής υγείας, σε μία τηλεφωνική γραμμή συμβουλευτικής υποστήριξης, μπορεί να υπερτερήσει, αφορά στη δυνατότητα του πρώτου να παρέχει υπηρεσίες σε πολύ περισσότερους ανθρώπους ταυτόχρονα, αντιμετωπίζοντας έτσι το πρόβλημα της έλλειψης του ανθρώπινου δυναμικού και την υπερφόρτωση του μειωμένου αριθμού των διαθέσιμων συμβούλων ψυχικής υγείας, ειδικότερα σε περιόδους μεγάλου φόρτου εργασίας, όπως αυτή της πανδημίας. Πράγματι, στην περίπτωση του ConBot, τα μέλη του προσωπικού της τηλεφωνικής γραμμής δήλωσαν πως η μείωση του αριθμού των κλήσεων, που αφορούσαν σε συνήθη ερωτήματα του κοινού, έγινε ιδιαίτερα αισθητή.²⁰⁷

Ωστόσο, οι συνεχείς εναλλαγές στην επικαιρότητα της πανδημίας (π.χ. οι αλλαγές στον αριθμό των κρουσμάτων, στα μέτρα αντιμετώπισης κ.ο.κ.), απαιτούσαν την συχνή πρόσβαση των αρχών στις προκαθορισμένες απαντήσεις και τη συνεχή διαμόρφωση των κειμένων.²⁰⁸ Έτσι, ο πολύς χρόνος που δαπανήθηκε στην, μέσω συμβατικού προγραμματισμού, διαμόρφωση των απαντήσεων, στην ουσία κάλυψε τη χρόνο που ελευθερώθηκε από την προώθηση των κλήσεων στο ConBot, περιορίζοντας την απόδοσή του σε ότι αφορά τη μείωση του φόρτου εργασίας. Κάτι τέτοιο θα μπορούσε να βελτιωθεί σημαντικά στο μέλλον. Εάν τα μοντέλα έχουν πρόσβαση στις ιστοσελίδες των τοπικών υγειονομικών αρχών και χρησιμοποιούν τις ήδη αναρτημένες σχετικές πληροφορίες, η περαιτέρω ενασχόληση με τα κείμενα προς απάντηση δεν θα είναι απαραίτητη.²⁰⁹

Πέρα από τα παραπάνω, αναμφίβολα η συνομιλία με ένα chatbot ψυχικής υγείας αποτελεί μια ενδιαφέρουσα εμπειρία για το χρήστη. Σε αντίθεση με ένα πεδίο κειμένου αναζήτησης ή την πληκτρολόγηση ενός αριθμού ως επιλογή απάντησης σε τηλεφωνικές γραμμές εξυπηρέτησης, μια συνομιλία ελεύθερου περιεχομένου έχει πιο πλούσια δομή και παρέχει πολύ περισσότερα δεδομένα προς επεξεργασία. Μέσω φυσικής επεξεργασίας NLP, ένα συνομιλητικό chatbot δύναται να προβεί στην κατανόηση της συναισθηματικής κατάστασης του ατόμου, να εκπαιδευτεί σε πραγματικό χρόνο μέσα από τις απαντήσεις του και να προσαρμόσει αναλόγως τις προτάσεις και τις κατευθύνσεις που θα του δώσει. Προσομοιάζοντας μια ανθρώπινη συνομιλία, μπορεί να δημιουργήσει την αίσθηση ενός φυσικού και πραγματικού περιβάλλοντος, εντός του οποίου ο χρήστης, ενθαρρύνεται στην έκφραση των αναγκών και των συναισθημάτων του.²¹⁰

Δεν είναι λίγα, τα ερευνητικά εκείνα στοιχεία που υποδεικνύουν πως οι χρήστες είναι πιο θετικοί στο να μοιραστούν πληροφορίες με μια εφαρμογή T.N., ιδίως όταν πρόκειται για ευαίσθητα προσωπικά θέματα.²¹¹ Συχνά, ο κοινωνικός στιγματισμός που συνοδεύει την ψυχική νόσο, αποτελεί ανασταλτικό παράγοντα για την επίσκεψη στις

²⁰⁶ Sweeney et al., “Can Chatbots Help Support a Person’s Mental Health?” 4.

²⁰⁷ Voelskow et al., “Prospective mixed-methods study,” 4.

²⁰⁸ *Αυτόθι*, 4-5.

²⁰⁹ *Αυτόθι*, 10-11.

²¹⁰ Simon D’Alfonso, Olga Santesteban-Echarri, Simon Rice, Greg Wadley, Reeva Lederman, Christopher Miles, John Gleeson, Mario Alvarez-Jimenez, “Artificial Intelligence-Assisted Online Social Therapy for Youth Mental Health,” *Frontiers in Psychology* 8, Art. 796 (June, 2017): 9-11, <https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00796/full>.

²¹¹ *Αυτόθι*, 9.

υπηρεσίες ψυχικής υγείας, με αποτέλεσμα η υποστήριξη να αναζητείται εκτός αυτών και συγκεκριμένα, από κοντινούς ανθρώπους, φίλους, οικογενειακά μέλη ή και εφαρμογές της τεχνολογίας. Την προτίμηση στις τελευταίες ειδικότερα, θα μπορούσε να εξηγήσει η έλλειψη εμπιστοσύνης και η δυσκολία στην ανάπτυξη ανθρωπίνων σχέσεων που χαρακτηρίζουν την ψύχωση.²¹²

Υπάρχουν, ωστόσο, αδύναμα σημεία τα οποία δεν πρέπει να περάσουν απαρατήρητα. Με άλλα λόγια, οι δυνατότητες των chatbot έχουν τα όριά τους. Από τη μία, η υπολογιστική τους ανωτερότητα κατά τη διάρκεια μιας συνομιλίας είναι εμφανής, καθώς πράγματι, μπορούν να παράγουν δεδομένα και αναλύσεις σε μία κλίμακα που οι άνθρωποι δεν θα μπορούσαν να το κάνουν.²¹³ Από την άλλη, ενδέχεται να μην είναι σε θέση να κατανοήσουν επαρκώς το ανθρώπινο λεξιλόγιο και το ουσιαστικό πρόβλημα της ψυχικής υγείας που εκφράζεται, δίνοντας ασαφείς απαντήσεις ή αγνοώντας τα σημάδια ενός σοβαρού ενυπάρχοντος κινδύνου, ώστε να αντιδράσουν καταλλήλως (λ.χ. αυτοκτονικές τάσεις).

Σε πρόσφατη μελέτη των Palanica et al. (2019), το 70% των ιατρών που ερωτήθηκαν, εξέφρασε την ανησυχία του αναφορικά με τους κινδύνους που ελλοχεύει η φροντίδα της ψυχικής υγείας από ένα chatbot. Μεταξύ αυτών, η αδυναμία του να κατανοήσει την πολύπλοκη χρήση της γλώσσας που συνδέεται με μία κρίση ψυχικής υγείας.²¹⁴ Παρομοίως, η σχετική έρευνα για την αποτελεσματικότητα του ConBot, έδειξε ότι ένα φωνητικό ρομπότ δεν μπορεί να αντικαταστήσει έναν άνθρωπο επαγγελματία, ιδίως για πιο σύνθετες και ατομικές ανησυχίες ή πολύπλοκα εκφραζόμενα ερωτήματα από όσους αναζητούν συμβουλευτική υποστήριξη καλώντας στις γραμμές βοήθειας.²¹⁵ Αυτό δήλωσε το 79% των ανθρώπων που κάλεσαν στη γραμμή, ενώ το 15% αυτών, τερμάτισε τη κλήση μόλις άκουσε μια τεχνητή εφαρμογή να αποκρίνεται αντί ενός ανθρώπου.²¹⁶ Το γεγονός και μόνο, ότι το 51% των κλήσεων προωθήθηκαν σε κάποιον υπάλληλο, αντανακλά τη σημασία που έχει η άμεση επαφή με έναν άνθρωπο σε μία δύσκολη στιγμή.²¹⁷

Θα λέγαμε λοιπόν, πως η αποτελεσματικότητα ενός voicebot, όπως το ConBot, δύναται να είναι υψηλή κατά την έναρξη περιόδων κρίσης, εκεί όπου ναι μεν ο όγκος των κλήσεων υπερβαίνει κατά πολύ το μέσο όρο και τη διαθεσιμότητα των εργαζομένων που τις χειρίζονται, αλλά το περιεχόμενο των ερωτημάτων που τίθενται, παραμένει σταθερό και προβλέψιμο.²¹⁸ Για την ώρα, η ανθρώπινη παρουσία φαίνεται πως παραμένει αναντικατάστατη. Σίγουρα στο μέλλον, η ασταμάτητη ανάπτυξη της Τ.Ν. και η εμφάνιση μεγάλων μοντέλων γλωσσικών όπως το ChatGPT, θα μπορούσαν να

²¹² Katherine Berry, Amy Salter, Rohan Morris, Susannah James, Sandra Bucci, “Assessing Therapeutic Alliance in the Context of mHealth Interventions for Mental Health Problems: Development of the Mobile Agnew Relationship Measure (mARM) Questionnaire,” *Journal of Medical Internet Research* 20, no. 4, e90 (April, 2018): 2, <https://www.jmir.org/2018/4/e90/>.

²¹³ Sedlakova, Trachsel, “Conversational Artificial Intelligence in Psychotherapy,” 10.

²¹⁴ Adam Palanica, Peter Flaschner, Anirudh Thommandram, Michael Li, Yan Fossat, “Physicians' Perceptions of Chatbots in Health Care: Cross-Sectional Web-Based Survey,” *Journal of Medical Internet Research* 21, no. 4, e12887 (April, 2019): 1-10, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6473203/>; Kretzschmar et al., “Can Your Phone Be Your Therapist?,” 7.

²¹⁵ Voelskow et al., “Prospective mixed-methods study,” 9-10.

²¹⁶ *Αυτόθι*, 1.

²¹⁷ *Αυτόθι*, 11

²¹⁸ *Αυτόθι*, 10.

βελτιώσουν κατά πολύ το χειρισμό ακόμη και των πιο πολύπλοκων ανησυχιών.²¹⁹ Υπάρχει πολύς δρόμος ακόμα, όμως, για να φτάσουμε στο σημείο εκείνο όπου θα θεωρούμε πως ένα bot μπορεί να προσαρμοστεί πλήρως στις συγκεκριμένες ανάγκες του χρήστη. Αν και υπάρχει ένα ελάχιστο επίπεδο εξατομίκευσης, η υποστήριξη που παρέχεται είναι ακόμα γενική και μάλλον, μοιάζει περισσότερο με ένα βιβλίο αυτοβοήθειας.²²⁰

Οι παραπάνω αστοχίες των συστημάτων της T.N. θα μπορούσαν να δημιουργήσουν περαιτέρω βλάβη και είναι άξιο συζήτησης το εάν και σε ποιο βαθμό, αυτή η πιθανότητα λαμβάνεται σοβαρά υπόψη από τους προγραμματιστές. Εάν οι εφαρμογές αυτές δεν υπόκεινται σε συνεχείς αξιολογήσεις, ως προς τα οφέλη και την επικινδυνότητά τους, τότε προκύπτει έντονος προβληματισμός γύρω από το εάν, η εξάπλωσή της διάθεσής τους πλαισιώνεται από τις απαραίτητες και ηθικά υπεύθυνες πρακτικές.²²¹

Ο προβληματισμός αυτός εντείνεται περισσότερο, αν λάβουμε υπόψη μια σειρά από διαδεδομένους ψηφιακούς βοηθούς που δημιουργήθηκαν για την διευκόλυνση της καθημερινότητας των ατόμων, όπως η Siri της Apple, η Cortana της Microsoft ή και η Alexa της Amazon. Αν και οι εφαρμογές αυτές δεν σχεδιάστηκαν αποκλειστικά για την παροχή υπηρεσιών συμβουλευτικής, συχνά διαφημίζονται για τη δυνατότητα παροχής κατευθύνσεων προς το χρήστη. Για παράδειγμα, η Alexa μπορεί να παρέχει τεχνικές (όπως ασκήσεις αναπνοής) για τη διαχείριση του άγχους,²²² ενώ το γνωστό στους περισσότερους λογισμικό ChatGPT, αν και σε σχετική ερώτηση δηλώνει πως, ως σύστημα της T.N., δεν είναι σε θέση να παρέχει συμβουλευτική υποστήριξη, προβαίνει στη διεξαγωγή συνομιλίας και στην παροχή συμβουλών (όπως η συχνή άσκηση και η τήρηση προσωπικού ημερολογίου καταγραφής συναισθημάτων), ως απάντηση σε δηλώσεις όπως «νιώθω έντονη θλίψη αυτή τη στιγμή». Με την ραγδαία ανάπτυξη της T.N. και τη χρήση ολοένα και πιο εξελιγμένων εργαλείων κατά την αλληλεπίδραση ανθρώπου και μηχανής (λ.χ. επεξεργασίας φυσικής γλώσσας NPL), τα συστήματα αυτά, πιθανά θα καταφέρνουν να βελτιώνουν συνεχώς τις αποκρίσεις τους στα ερωτήματα που διατυπώνουν οι χρήστες. Για παράδειγμα, σε σχετική δοκιμή με δήλωση αυτοκτονικού ιδεασμού, η Siri κατάφερε να αναγνωρίσει τη σοβαρότητα και να παραπέμψει σε γραμμή βοήθειας για την πρόληψη της αυτοκτονίας.²²³ Ωστόσο, η συχνή ασυνέπεια των απαντήσεων σε κρίσιμες και επείγουσες περιπτώσεις, καθώς και το σκεπτικό γύρω από το σχεδιασμό τους, που δεν στρέφεται αποκλειστικά στο κομμάτι της ψυχικής υγείας, απαιτεί την προσοχή τόσο των εταιρειών που προωθούν τις εφαρμογές αυτές, όσο και των ανθρώπων που τις προτιμούν σε περιπτώσεις ανάγκης.

Επιπλέον, ένα σημαντικό θέμα που θα πρέπει να λάβουμε υπόψη στην περίπτωση των αυτόνομων συμβούλων T.N., που θα αντικαταστήσουν τους ανθρώπινους θεραπευτές, είναι το λεγόμενο *πρόβλημα του μαύρου κουτιού* (*The Black Box Problem*), όρος που δηλώνει την αδυναμία των ίδιων των προγραμματιστών να προβλέψουν και να εξηγήσουν τις επιλογές και τις αποφάσεις των συστημάτων T.N. που οι ίδιοι

²¹⁹ *Αυτόθι*, 11.

²²⁰ Kretzschmar et al., “Can Your Phone Be Your Therapist?,” 3-4.

²²¹ *Αυτόθι*, 7.

²²² “Deep Breath,” Amazon, accessed September 25, 2023, <https://www.amazon.com/DrawnToDigital-Deep-Breath/dp/B01N1NFN44>.

²²³ “In a crisis, Siri and Cortana may not have your back,” CBS News, accessed September 19, 2023, <https://www.cbsnews.com/news/health-crisis-siri-and-cortana-may-not-have-your-back/>.

ανέπτυξαν.²²⁴ Όπως έχουμε δει μέχρι τώρα, η λειτουργία των chatbot, που χρησιμοποιούνται ευρέως από το κοινό, βασίζονται κυρίως σε αλγορίθμους μηχανικής μάθησης, οι οποίοι έχουν τη δυνατότητα να επεξεργάζονται μεγάλα σύνολα δεδομένων, παράλληλα να εκπαιδεύονται από αυτά και σταδιακά να δημιουργούν τους δικούς τους κανόνες στον τρόπο που λειτουργούν και στις προτροπές που κάνουν προς το χρήστη. Κάθε αλγόριθμος δομείται βασισμένος σε άλλους και είναι αυτή η περίτεχνη αρχιτεκτονική, που τελικά θα παράγει κάποιο αποτέλεσμα και όχι η λειτουργία του μεμονωμένου αλγορίθμου.²²⁵ Η ανησυχία γύρω από το πρόβλημα του μαύρου κουτιού, θα μπορούσε να εξελιχθεί σε ηθικό κίνδυνο, αν λάβουμε υπόψη το γεγονός ότι η συμμόρφωση με τις συμβουλές ενός αυτόνομου λογισμικού, του οποίου τον τρόπο σκέψης δεν μπορούμε να γνωρίζουμε και να ελέγξουμε, είναι κάθε άλλο παρά ασφαλής. Τίθεται το ερώτημα του εάν και σε ποιο βαθμό θα μπορούσε να διασφαλιστεί το ότι μία εφαρμογή της T.N. που αναλαμβάνει το ρόλο ενός συμβούλου ψυχικής υγείας, θα επιδείξει την ίδια συμπεριφορά με έναν άνθρωπο σύμβουλο και θα παρέχει συμβουλές που θα είναι ευθυγραμμισμένες ηθικά και επιστημονικά με τις δικές του.²²⁶ Επιπλέον, αν θα αποφύγει συμπεριφορές μεροληψίας και προκαταλήψεων (*AI bias*), που μπορεί να διεισδύσουν στα συστήματα, είτε από ανισοκατανομές στα δεδομένα εισόδου, είτε από τους ίδιους προγραμματιστές και τις αποφάσεις τους κατά τον σχεδιασμό τους.²²⁷ Μάλιστα, εάν ο έλεγχος των διαδικασιών μέσω των οποίων ένα σύστημα T.N. καταλήγει στις όποιες συμβουλές παρέχει είναι αδύνατο, λόγω του *black box problem*, τότε είναι αδύνατο επίσης, το να εντοπιστεί το εάν οι συμβουλές αυτές είναι προϊόν προκαταλήψεων (*AI bias*) ή όχι. Τα ερωτήματα αυτά είναι κρίσιμα, πόσο μάλλον όταν μιλούμε για ανθρώπους που εμφανίζουν βαριά μορφή ψύχωση ή ανθρώπους χαμηλού νοητικού επιπέδου, ακόμη και παιδιά, που δεν θα ήταν σε θέση να αμφισβητήσουν την προτροπή ενός chatbot ή και τη δυνατότητά του να παρέχει αυτόνομα συμβουλευτική καθοδήγηση.

Αντί της ανάληψης του ρόλου ενός συμβούλου ψυχικής υγείας, κάτι για το οποίο εμφανώς δεν έχει έρθει η στιγμή, οι εφαρμογές της T.N., με την καινοτομία που μπορούν να προσφέρουν, θα μπορούσαν να διαδραματίσουν ένα σημαντικό υποστηρικτικό ρόλο στις ήδη υπάρχουσες επιλογές, ώστε να συμβάλλουν στη δημιουργία ενός συστήματος, που θα μπορεί να ανταποκριθεί καλύτερα στις ανάγκες για την παροχή άμεσης ή και μακροπρόθεσμης συμβουλευτικής υποστήριξης.

Συγκεκριμένα, οι επαγγελματίες που συμμετείχαν στη μελέτη των Palanica et al., ανέφεραν πως τα chatbot για την ψυχική υγεία, μπορούν να έχουν έναν επωφελή ρόλο στην υποστήριξη της, παρά το ότι δεν έχουν την εμπειρογνωμοσύνη που χρειάζεται για να αντικαταστήσουν έναν παραδοσιακό ιατρό.²²⁸ Όπως υποστήριζαν, η χρήση τους

²²⁴ Άλκης Γούναρης, Γεώργιος Κωστελέτος, *Τεχνητή Νοημοσύνη και Ηθική, Ενότητα 7^η: Υπερβαίνοντας τις ανθρώπινες δυνατότητες [πανεπιστημιακές σημειώσεις]*. ΕΚΠΑ, Τμήμα Φιλοσοφίας, ΠΜΣ Φιλοσοφίας: Εφαρμοσμένη Ηθική, χειμερινό εξάμηνο 2022-2023, Αθήνα.

²²⁵ Βαρβάρα Ζαχαράκη, «Τεχνητή Νοημοσύνη, λήψη αποφάσεων με χρήση αλγορίθμων, ηθική και δεοντολογία στο Δημόσιο τομέα» (Μεταπτυχιακή διατριβή, Πανεπιστήμιο Δυτικής Αττικής, 2022), 66-67, διαθέσιμο στο: https://polynoe.lib.uniwa.gr/xmlui/bitstream/handle/11400/3345/%ce%92%ce%91%ce%a1%ce%92%ce%91%ce%a1%ce%91%20%ce%96%ce%91%ce%a7%ce%91%ce%a1%ce%91%ce%9a%ce%97_%ce%94%ce%9c2088.pdf?sequence=1&isAllowed=y.

²²⁶ Γούναρης, Κωστελέτος, *Τεχνητή Νοημοσύνη και Ηθική, Ενότητα 7^η* [πανεπιστημιακές σημειώσεις], ΕΚΠΑ.

²²⁷ Müller, “Ethics of Artificial Intelligence and Robotics.”

²²⁸ Palanica et al., “Physicians' Perceptions of Chatbots in Health Care, 1-10; Sweeney et al. “Can Chatbots Help Support a Person's Mental Health?” 4.

μπορεί να συμβάλλει στην συναισθηματική αποφόρτιση του ατόμου σε περιπτώσεις κατάθλιψης, άγχους ή φοβίας, να βελτιώσει τη δραστηριότητα ή τη λήψη φαρμακευτικής αγωγής μέσα από σχετικές προτροπές, να μειώσει τις άσκοπες επισκέψεις στις υπηρεσίες (αποφορτίζοντάς τες έτσι και αυτές) ή και να παρέχει σημαντικές πληροφορίες, όπως οι διαθέσιμες υπηρεσίες ψυχικής υγείας σε μία συγκεκριμένη περιοχή.²²⁹

Πέρα από αυτό, όμως, υπάρχουν και άλλες ενδιαφέρουσες δυνατότητες που θα μπορούσαν να ενισχύσουν τη διαδικασία της συμβουλευτικής υποστήριξης. Για παράδειγμα, οι πλατφόρμες MOST²³⁰ και myTherapist,²³¹ που παρέχουν τη δυνατότητα της online συνομιλίας με επαγγελματίες θεραπευτές, ζητούν από τους χρήστες να συμπληρώσουν ένα σύντομο ερωτηματολόγιο, προτού συνδεθούν με τον εκάστοτε ειδικό. Οι απαντήσεις που δίνουν, βοηθούν τον επαγγελματία στην αξιολόγηση της κατάστασης του χρήστη και στις ανάγκες που τον οδήγησαν στο να αναζητήσει βοήθεια. Μια τέτοια διαδικασία σε μια γραμμή βοήθειας, με επιπρόσθετες της δυνατότητες που προσδίδουν εργαλεία, όπως η επεξεργασία φυσικής γλώσσας NLP, θα μπορούσε να παρέχει στον επαγγελματία ένα σύντομο προφίλ του ατόμου με το οποίο πρόκειται να συνομιλήσει, καθιστώντας έτσι, πιο εύστοχη την παρέμβαση του.²³²

Συγκεκριμένα, η απάντηση ερωτήσεων πριν από την σύνδεση με τον επαγγελματία ειδικό, θα μπορούσε να σκιαγραφήσει καλύτερα την ψυχολογική κατάσταση στην οποία βρίσκεται ο χρήστης, ή τα γενικότερα χαρακτηριστικά του (λ.χ. ανήλικος, γονέας) και να εντοπίσει τον ακριβή λόγο για τον οποίο κάλεσε. Εάν το αίτημά του δεν ανήκει στα συχνά υποβαλλόμενα ερωτήματα του κοινού, τότε ο χρήστης θα μπορούσε να παραπεμφθεί στον κατάλληλο από τους ειδικούς που στελεχώνουν τη γραμμή (π.χ. ψυχολόγο, κοινωνικό λειτουργό) ή και να επανασυνδεθεί με κάποιον με τον οποίο έχει συνομιλήσει και κατά το παρελθόν για το ίδιο θέμα, και είτε η σύνδεση διεκόπη για τεχνικούς λόγους, είτε χρειάζεται περαιτέρω διευκρινίσεις ή καλεί για περαιτέρω ενημέρωση αναφορικά με κάποιο περιστατικό (λ.χ. follow up μετά από μία ανώνυμη καταγγελία).

Ακόμη, στο πλαίσιο της συνεχούς εξέλιξης της T.N. και εφαρμογών της, όπως ο ψηφιακός φαινότυπος ή το Διαδίκτυο των πραγμάτων (Internet of Things - IoT), δεν θα ήταν υπερβολή το να υποθέσουμε πως μελλοντικά, η κλήση σε μία γραμμή βοήθειας και συμβουλευτικής υποστήριξης, θα μπορούσε να τροφοδοτείται από δεδομένα όπως η γεωγραφική τοποθεσία του χρήστη σε περίπτωση κινδύνου (μέσω GPS) ή και οι φυσικές λειτουργίες του σώματος του (biofeedback, όπως ο καρδιακός παλμός και ο ρυθμός της αναπνοής, μέσω αισθητήρων του καρπού) που θα μπορούσαν να δηλώσουν την έναρξη μιας κρίσης πανικού.²³³ Επιπλέον, η αποστολή μιας στιγμιαίας φωτογραφίας του χρήστη (μέσω screenshot) ή η συνομιλία μέσω εικόνας, θα μπορούσε να εμπλουτίσει τη συνομιλία με δεδομένα γύρω από τα συναισθήματα του, που θα εξάγονται δια της ανάλυσης των μορφασμών του προσώπου του, μια δυνατότητα που

²²⁹ Sweeney et al. “Can Chatbots Help Support a Person’s Mental Health?” 6-9.

²³⁰ “Mental health support when and where you need it,” MOST, accessed September 23, 2023, <https://most.org.au/>.

²³¹ «Πως λειτουργεί η online ψυχοθεραπεία στο myTherapist», myTherapist, προσπελάστηκε στις 28 Σεπτεμβρίου, 2023, <https://www.mytherapist.gr/web-app>.

²³² D’Alfonso et al., “Artificial Intelligence-Assisted Online Social Therapy,” 9.

²³³ *Αυτόθι*, 11.

μας παρέχει η τεχνολογία αναγνώρισης συναισθημάτων (*Facial Emotion Technology*).²³⁴

Χωρίς αμφιβολία, η συζήτηση αυτή, γύρω από τις ενδεχόμενες εξελίξεις και τα όσα η τεχνητή νοημοσύνη θα μπορέσει να προσφέρει μελλοντικά στον τομέα της ψυχικής υγείας και συγκεκριμένα, στις γραμμές συμβουλευτικής βοήθειας και υποστήριξης, μοιάζει ατελείωτη. Η αναμόρφωση αυτής της ήδη υπάρχουσας υπηρεσίας, όπως είδαμε, μπορεί να παρέχει σημαντικά οφέλη, αλλά και να ενέχει σοβαρούς κινδύνους για το άτομο και τον ευάλωτο ψυχισμό του. Ένα σημαντικό θέμα που δεν θα πρέπει να παραλείπουμε, αφορά σίγουρα στο κομμάτι της προστασίας των προσωπικών δεδομένων.

Το θέμα της ιδιωτικότητας και του απορρήτου παραμένει ζωτικής σημασίας. Είναι πολύ σημαντικό τα προσωπικά στοιχεία που δίνονται από τον χρήστη, εάν συλλέγονται, να διατηρούνται εμπιστευτικά και αν διαμοιράζονται (λ.χ. για σκοπούς έρευνας), αυτό να γίνεται πάντοτε με τη συγκατάθεση του και την απαραίτητη διασφάλιση της ανωνυμίας και της απόκρυψης οποιουδήποτε αναγνωρίσιμου προσωπικού στοιχείου (λ.χ. user name, email κ.α.). Επιπλέον, οι διαδικασίες και οι ρυθμίσεις περί απορρήτου θα πρέπει να παραμένουν διαφανείς για τους πολίτες. Παρά το ότι οι χρήστες έχουν την επιλογή να εισέλθουν σε έναν σύνδεσμο για να διαβάσουν τις πλήρεις πολιτικές απορρήτου στις ιστοσελίδες των chatbot, ίσως πολλοί δεν θα διαθέσουν χρόνο για κάτι τέτοιο. Είναι, επομένως, σημαντικό οι βασικές σχετικές πληροφορίες να παρέχονται σε απλή μορφή και γλώσσα (π.χ. με γραφικά ή κουκκίδες) και ιδιαίτερα, κατά την έναρξη της οποιας αλληλεπίδρασης. Κάτι που αξίζει προσοχής, είναι το ότι κάποιες από τις υπάρχουσες εφαρμογές, όπως το Wysa, διατίθενται ως εφαρμογή, προσφέροντας στους χρήστες τη δυνατότητα να συνομιλούν ανώνυμα. Ωστόσο, κάποιες άλλες, όπως το Woebot, διατίθενται επιπλέον και μέσω του Facebook Messenger, όπου οι χρήστες συνδέονται με τα πραγματικά τους ονόματα. Τα δεδομένα όμως που συλλέγονται εκεί, υπόκεινται στην πολιτική απορρήτου του Facebook και βάσει αυτής, μπορούν να μοιραστούν σε τρίτους. Ήδη γνωρίζουμε ως παράδειγμα, την εμπλοκή του Facebook στο σκάνδαλο δεδομένων της Cambridge Analytica, όπου δεδομένα εκατομμυρίων χρηστών διαμοιράστηκαν για πολιτικούς σκοπούς χωρίς τη ρητή συγκατάθεσή τους. Περιστατικά σαν αυτό, τονίζουν την ανάγκη της συνεχούς εποπτείας και παρακολούθησης για την εξασφάλιση της συμμόρφωσης με τα πρότυπα απορρήτου.²³⁵

Είδαμε λοιπόν μέχρι τώρα, τα πλεονεκτήματα και τα μειονεκτήματα που μπορεί να χαρακτηρίζουν τη χρήση της Τ.Ν. στις γραμμές βοήθειας και συμβουλευτικής υποστήριξης. Πέρα, ωστόσο, από το πρακτικό επίπεδο της εφαρμογής, θα πρέπει να συνυπολογίσουμε το γεγονός πως, λόγω της φύσης των γραμμών αυτών, το περιεχόμενο μιας κλήσης σε αυτές, απέχει κατά πολύ από αυτό μιας κλήσης για απλή αναζήτηση πληροφοριών ή προγραμματισμό κάποιου ραντεβού, όπως λ.χ. σε ένα

²³⁴ Η τεχνολογία αναγνώρισης συναισθημάτων (*Facial Emotion Technology*), αποτελεί μια από τις κατηγορίες των *τεχνολογιών αναγνώρισης προσώπου* (*Facial Recognition Technology*) και αφορά στην εξόρυξη δεδομένων γύρω από τα συναισθήματα ενός ατόμου, μέσα από την ανάλυση των εκφράσεων του προσώπου ή των κινήσεων, λ.χ. οι κινήσεις των χεριών, η στάση του σώματος κ.ο.κ. Διονυσία Αθανασοπούλου, «Συστήματα αναγνώρισης προσώπου με χρήση Τεχνητής Νοημοσύνης και ζητήματα προστασίας δεδομένων προσωπικού χαρακτήρα» (Μεταπτυχιακή διατριβή, Πανεπιστήμιο Πειραιώς, 2022), 19, διαθέσιμο στο: <https://dione.lib.unipi.gr/xmlui/handle/unipi/14835> [Πρόσβαση στις 14/04/2023].

²³⁵ Kretzschmar et al., “Can Your Phone Be Your Therapist?,” 4-5.

κέντρο τηλεφωνικής εξυπηρέτησης πελατών μιας ιδιωτικής εταιρείας. Επομένως, το ευαίσθητο και προσωπικό περιεχόμενο της συνομιλίας που λαμβάνει χώρα, καθώς και η ευάλωτη συναισθηματική κατάσταση και οι προσδοκίες του ανθρώπου που αποφασίζει να καλέσει, επαρκούν για να στοιχειοθετήσουν μια θεραπευτική σχέση μεταξύ καλούντος και επαγγελματία, ακόμη κι αν η συνομιλία μεταξύ τους δεν θα διαρκέσει, παρά μόνο λίγα λεπτά της ώρας. Μένει λοιπόν να διερευνήσουμε το εάν η Τ.Ν. δύναται να δημιουργήσει μια τέτοια σχέση και να συνάψει θεραπευτική συμμαχία με τον χρήστη μιας εφαρμογής. Το θέμα αυτό θα μας απασχολήσει στην επόμενη ενότητα.

3.γ. Η θεραπευτική σχέση

Το επιστημονικό πεδίο της Συμβουλευτικής, πρώτο επισήμανε την αξία και τη μοναδικότητα της θεραπευτικής σχέσης ανάμεσα στο σύμβουλο και στο συμβουλευόμενο. Όπως την κατανόησε ο «πατέρας» της συμβουλευτικής Carl Rogers, η θεραπευτική σχέση είναι «η ουσιαστική πτυχή της θεραπείας».²³⁶ Η «συνδιαλλακτική σχέση ισότιμης συνεργασίας»,²³⁷ όπως αναφέρει ο ψυχολόγος Ray Woolfe (1996), μεταξύ δύο ανθρώπων εκ των οποίων «ο ένας διαθέτει τις απαραίτητες δεξιότητες και την εξειδίκευση (τεχνογνωσία) που ο άλλος χρειάζεται για να εξαλείψει τον πόνο ή την ταλαιπωρία του».²³⁸

Καθώς η φροντίδα της ψυχικής υγείας υιοθετεί όλο και περισσότερο τις εφαρμογές της Τ.Ν., η ξεχωριστή αυτή έννοια απαιτεί περαιτέρω διερεύνηση. Τα τελευταία χρόνια ο όρος *Ψηφιακή Θεραπευτική Συμμαχία (Digital Therapeutic Alliance – DTA)* εμφανίζεται όλο και πιο συχνά, αναφερόμενος σε μια σειρά τύπων φροντίδας ψηφιακής μορφής στην ψυχική υγεία, όπως μια συνεδρία συμβουλευτικής ανάμεσα σε δύο ανθρώπους που διεξάγεται μέσω εφαρμογών smartphone (λ.χ. skype), ή μια συνεδρία που εξελίσσεται ανάμεσα στο άτομο και ένα εξελιγμένο chatbot Συνομιλητικής Τ.Ν. (Conversational AI), που αναλαμβάνει το ρόλο του συμβούλου.²³⁹

Στον τομέα της κινητής υγείας (mHealth) συγκεκριμένα, η έννοια της θεραπευτικής σχέσης έχει λάβει ελάχιστη εμπειρική και ερευνητική προσοχή. Μια από τις λιγότες έρευνες είναι αυτή των Berry et al., που μελέτησε στις απόψεις των χρηστών γύρω από την έννοια της σχέσης που μπορεί να αναπτυχθεί κατά τη χρήση ψηφιακών εφαρμογών στην ψυχική υγεία.²⁴⁰ Τα πορίσματα έδειξαν ότι, αν και οι χρήστες εντόπισαν σημαντικές ελλείψεις σε σχέση με έναν ανθρώπινο επαγγελματία, ως προς την ικανότητα της εφαρμογής να ανταποκρίνεται πλήρως στις ανάγκες τους, αναφέρθηκαν στις έννοιες της σχέσης και της δημιουργίας δεσμού, τονίζοντας τη δυνατότητα να νοιώσουν ανοιχτοί κατά τη διάρκεια της συνομιλίας.²⁴¹ Συγκεκριμένα, ανέφεραν μια

²³⁶ Αναστασία Καλαντζή-Αζίζι, Μαρία Μαλιώτη-Λοΐζου, «Η θεραπευτική σχέση στην κλινική και συμβουλευτική ψυχολογία», *Ψυχολογία: The Journal of the Hellenic Psychological Society* 26, vo. 2 (Ιούνιος, 2021): 3, https://www.researchgate.net/publication/352402734_E_therapeutike_schese_sten_klinike_kai_symbouleutike_psychologia.

²³⁷ *Αυτόθι*.

²³⁸ *Αυτόθι*.

²³⁹ D'Alfonso, "AI in mental health," 114.

²⁴⁰ Berry et al., "Assessing Therapeutic Alliance," 2-3.

²⁴¹ *Αυτόθι*, 1.

αίσθηση ελευθερίας στο να είναι ειλικρινείς με την εφαρμογή, κάτι που επισημάνθηκε ως η κύρια διαφορά σε σχέση με τη διαζώσης θεραπείας με άνθρωπο ειδικό. Αυτός ο μειωμένος κίνδυνος εμφάνισης συναισθημάτων ντροπής, θεωρήθηκε από τους ερευνητές, σχετικός με τον έννοια της σχέσης, καθώς συχνά, η ντροπή ή ο φόβος της κριτικής από το θεραπευτή, αποτελούν λόγο διακοπής της διαζώσης θεραπείας. Παράλληλα, το γεγονός πως η εκ των υστέρων μη διαθεσιμότητα της εφαρμογής δημιούργησε συναισθήματα απώλειας στους συμμετέχοντες, υπονόησε ότι πράγματι, είχαν αναπτύξει ένα είδος σχέσης μαζί της.²⁴²

Ακόμη μία σχετική έρευνα είναι αυτή των Clarke et al., στην οποία συμμετείχαν ενενήντα άτομα με κατάθλιψη και άγχος, που παρακολούθησαν το ψυχοθεραπευτικό πρόγραμμα της εφαρμογής myCompass.²⁴³ Από την έρευνα προέκυψε πως η δημιουργία της θεραπευτικής σχέσης σε ένα διαδικτυακό περιβάλλον είναι εφικτή, ακόμη και χωρίς την ανθρώπινη υποστήριξη. Οι συμμετέχοντες ανέφεραν πως η εφαρμογή τους προσέφερε ένα ασφαλές πλαίσιο, εντός του οποίου μπόρεσαν να αντιμετωπίσουν τις δυσκολίες τους. Αν και αρκετοί, ήταν αρχικά επιφυλακτικοί, η εμπιστοσύνη τους στις δυνατότητες της εφαρμογής αυξανόταν, όσο συνέχιζαν τη χρήση της, ενώ οι περισσότεροι αισθάνθηκαν να λαμβάνουν αποδοχή και κατανόηση. Ένας μόνο από τους συμμετέχοντες, που δεν αισθάνθηκε να λαμβάνει κατανόηση και αποδοχή, ανέφερε πως τα στοιχεία αυτά είναι περιττά κατά την αλληλεπίδρασή του με μία ψηφιακή εφαρμογή, κάτι που δεν θα ίσχυε στην διαπροσωπική επαφή με έναν ανθρώπινο θεραπευτή.²⁴⁴

Ίσως αυτή η τελευταία δήλωση να είναι ιδιαίτερα σημαντική, καθώς υποδηλώνει την ανάγκη να προσδιορίσουμε το εάν οι παράγοντες που καθιστούν αποτελεσματική τη θεραπευτική σχέση σε μία συνεδρία μεταξύ ανθρώπων είναι αυτοί, βάσει των οποίων θα αξιολογηθεί και μια συνεδρία μεταξύ ενός ανθρώπου και μιας εφαρμογής που επιστρατεύει την Τ.Ν.²⁴⁵ Οι μέχρι τώρα απόπειρες αξιολόγησης, σαν αυτές που αναφέραμε παραπάνω, έχουν προβεί για το σκοπό αυτό στην χρήση του ερωτηματολογίου *Agnew Relationship Measure (ARM)*,²⁴⁶ αντικαθιστώντας μέσα σε αυτό, τη λέξη «θεραπευτής» με τη λέξη «εφαρμογή».²⁴⁷ Το συγκεκριμένο

²⁴² *Αυτόθι*, 4.

²⁴³ “myCompass: A customisable self-help tool for your mental health,” Black Dog Institute. MyCompass, accessed September 30, 2023, <https://www.mycompass.org.au/>.

²⁴⁴ Janine Clarke, Judith Proudfoot, Alexis Whitton, Mary-Rose Birch, Megan Boyd, Gordon Parker, Vijaya Manicavasagar, Dusan Hadzi-Pavlovic, Andrea Fogarty, “Therapeutic Alliance With a Fully Automated Mobile Phone and Web-Based Intervention: Secondary Analysis of a Randomized Controlled Trial,” *Journal of Medical Internet Research in Mental Health* 25, no. 3 (1) e10 (Feb., 2016): 8-9, <https://mental.jmir.org/2016/1/e10/>.

²⁴⁵ Reeva Lederman, Simon D'Alfonso, “The Digital Therapeutic Alliance: Prospects and Considerations,” *JMIR Ment Health* 8, no. 7, e31385 (July, 2021): 2, <https://mental.jmir.org/2021/7/e31385/>.

²⁴⁶ Το ερωτηματολόγιο Agnew Relationship Measure (βλ. “Agnew Relationship Measure – 5 (ARM-5),” NovoPsych, accessed in October 1, 2023, <https://novopsych.com.au/assessments/feedback-informed/agnew-relationship-measure-5-arm-5/>), αξιολογεί πέντε έννοιες που θεωρούνται βασικά συστατικά στοιχεία της θεραπευτικής σχέσης και συμμαχίας. Τα στοιχεία αυτά είναι: ο δεσμός (*bond*), δηλ. τα αμοιβαία συναισθήματα θετικής εκτίμησης μεταξύ θεραπευτή και θεραπευόμενου, η συνεργασία (*partnership*) μεταξύ τους, η εμπιστοσύνη (*confidence*) που απορρέει από τις ικανότητες του θεραπευτή, η προθυμία στην ειλικρινή και ελεύθερη έκφραση (*openness*) από τον θεραπευόμενο και η πρωτοβουλία (*client initiative*), που σχετίζεται με συναισθήματα ελέγχου και ενδυνάμωσης από την πλευρά του. Berry et al., “Assessing Therapeutic Alliance,” 3.

²⁴⁷ Lederman, D'Alfonso, “The Digital Therapeutic Alliance,” 2.

ερωτηματολόγιο ωστόσο, όπως και άλλα συναφή που έχουν προσαρμοστεί και χρησιμοποιηθεί για τον ίδιο σκοπό (λ.χ. το *Working Alliance Inventory - WAI*²⁴⁸), βασίζονται στα χαρακτηριστικά της παραδοσιακής θεραπευτικής σχέσης. Έτσι η όλη προσέγγιση μοιάζει ατελής για δύο λόγους. Αφενός, γιατί παραβλέπει το γεγονός πως τα χαρακτηριστικά της παραδοσιακής θεραπευτικής σχέσης δεν ισχύουν απαραίτητα για την ψηφιακή. Αφετέρου, γιατί δε λαμβάνονται υπόψη τα μοναδικά χαρακτηριστικά της δεύτερης, καθώς και οι ενδεχόμενες νέες διαστάσεις και μορφές σχέσεων, που μπορεί να υπάρξουν και τις οποίες δεν συναντούμε στις παραδοσιακές μορφές θεραπείας.²⁴⁹

Για παράδειγμα, έχοντας σχεδιαστεί για τις δια ζώσης θεραπείες με ανθρώπινους θεραπευτές, τα ερωτηματολόγια Agnew Relationship Measure (ARM) και Working Alliance Inventory (WAI), δε συμπεριλαμβάνουν στην αξιολόγηση τους χαρακτηριστικά των διαδικτυακών μεθόδων, όπως η ευελιξία στον χρόνο, τον τόπο και τη διάρκεια της αλληλεπίδρασης, η δυνατότητα πρόσβασης μέσω διαφορετικών μέσων επικοινωνίας, καθώς και η ανωνυμία. Τα χαρακτηριστικά αυτά είναι σημαντικά, κι όμως παραμένουν σε μεγάλο βαθμό ανεξερεύνητα.²⁵⁰ Όπως τόνισαν οι συμμετέχοντες στη μελέτη των Berry et al., η δυνατότητα δημιουργίας θεραπευτικής σχέσης με την εφαρμογή είναι υπαρκτή, αλλά αφορά σε ένα διαφορετικό είδος σχέσης.²⁵¹ Χρειάζεται λοιπόν περαιτέρω έρευνα, ώστε να εντοπιστούν τα στοιχεία εκείνα που θα επαναπροσδιορίσουν την έννοια της θεραπευτικής σχέσης σε συνεδρίες που επιστρατεύουν την Τ.Ν. και θα χρησιμοποιηθούν για την ανάπτυξη νέων μοντέλων αξιολόγησης.²⁵²

Αναμφισβήτητα, ένας παράγοντας που συνδέεται στενά με την ανάπτυξη της θεραπευτικής σχέσης είναι η ποιότητα. Σύμφωνα με το εργαλείο *Mobile Application Rating Scale (MARS)*,²⁵³ που μετρά την ποιότητα μιας ψηφιακής εφαρμογής, οι παράγοντες που ορίζουν την ποιότητα αυτή είναι, μεταξύ άλλων, η εξατομίκευση της εφαρμογής, η διαδραστικότητα της, η ευκολία στη χρήση και η αξιοπιστία των πληροφοριών που παρέχει. Οι θετικές αξιολογήσεις γύρω από την ποιότητα, σχετίζονται με τη συχνότητα και τη διάρκεια της χρήσης της εφαρμογής και μοιάζει αυτονόητο το ότι, εάν ένα άτομο δεν είναι ικανοποιημένο με την ποιότητα μιας εφαρμογής για την ψυχική υγεία, είναι δύσκολο να αναπτύξει μία θεραπευτική σχέση κατά την αλληλεπίδραση μαζί της. Η σύνδεση αυτή είναι σημαντικό να λαμβάνεται υπόψη, καθώς η έλλειψη ποιότητας καθιστά ιδιαίτερα πιθανό, ένα άτομο να

²⁴⁸ Το ερωτηματολόγιο Working Alliance Inventory (WAI) αξιολογεί την ποιότητα της θεραπευτικής συμμαχίας ανάμεσα σε ένα άτομο και στον ψυχοθεραπευτή του. Βλ. “WAI Home,” Working Alliance Inventory, accessed October 3, 2023, <https://wai.profhorvath.com/>; “Working Alliance Inventory – Short Revised (WAI-SR),” WAI, accessed October 3, 2023, <https://wai.profhorvath.com/sites/default/files/upload/WAI-SR%20Client%20Version.pdf>.

²⁴⁹ Lederman, D'Alfonso, “The Digital Therapeutic Alliance,” 2.

²⁵⁰ Clarke et al., “Therapeutic Alliance With a Fully Automated Mobile Phone and Web-Based Intervention,” 9.

²⁵¹ Berry et al., “Assessing Therapeutic Alliance,” 6.

²⁵² Lederman, D'Alfonso, “The Digital Therapeutic Alliance,” 2.

²⁵³ Stoyan Stoyanov, Leanne Hides, David Kavanagh, Oksana Zelenko, Dian Tjondronegoro, Madhavan Mani, “Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps,” *JMIR mHealth and uHealth* 3, no. 1 e27 (March, 2015): 1-9, https://www.researchgate.net/publication/273407216_Mobile_App_Rating_Scale_A_New_Tool_for_Assessing_the_Quality_of_Health_Mobile_Apps.

εγκαταλείπει τη θεραπεία.²⁵⁴ Κάτι τέτοιο δεν θα γίνει ίσως αντιληπτό, όπως θα συνέβαινε με έναν ανθρώπινο θεραπευτή. Ενέχεται έτσι ο κίνδυνος, κάποιος να μη λάβει την απαραίτητη φροντίδα ψυχικής υγείας, τη στιγμή που το έχει ανάγκη. Μπορούμε να υποθέσουμε πόσο προβληματικό θα ήταν κάτι τέτοιο για κάποιον που εκφράζει αυτοκτονικό ιδεασμό σε μία τηλεφωνική γραμμή βοήθειας και συμβουλευτικής υποστήριξης. Η δυσκολία στη χρήση ή η έλλειψη αξιοπιστίας είναι πιθανό να τον οδηγήσει στον τερματισμό της κλήσης και παραμένει άγνωστο το εάν θα αναζητήσει εναλλακτικούς τρόπους επικοινωνίας με κάποιον ειδικό.

Σε σχέση με αυτό, έχουμε ήδη δει την, μέχρι στιγμής, ανεπάρκεια των chatbot στο να κατανοήσουν πλήρως την ιδιαιτερότητα των αναγκών και των συναισθημάτων που εκφράζονται από τον χρήστη, καθώς και τις ανησυχίες των ειδικών γύρω από αυτήν. Μπορούμε επίσης, να συνυπολογίσουμε παράγοντες που ίσως εντείνουν την ανεπάρκεια αυτή, όπως το «Πρόβλημα του Πλαισίου» της T.N. (The Frame Problem of AI), στο οποίο αναφερθήκαμε στην Ενότητα 2.α. της παρούσης διατριβής. Όπως είδαμε, το «Πρόβλημα Πλαισίου» αφορά στην αδυναμία των προγραμματιστών να συγκεκριμενοποιήσουν τις πληροφορίες με τις οποίες θα τροφοδοτήσουν τα συστήματα της T.N., παραβλέποντας τις αναρίθμητες εκείνες πληροφορίες που δεν είναι συναφείς με το αντικείμενο μελέτης.²⁵⁵ Στην περίπτωση ενός chatbot στον τομέα της ψυχικής υγείας, οι αναλύσεις των ποικίλων δηλώσεων του χρήστη, μπορεί να συμπεριλάβουν μη συναφείς πληροφορίες και επομένως, να οδηγήσουν στην εσφαλμένη ερμηνεία του κυρίου νοήματος πίσω από τα λεγόμενά του.

Η βελτίωση της ποιότητας, λοιπόν, που είναι απαραίτητη για τη δημιουργία της θεραπευτικής σχέσης και συνεπώς, για την αποτελεσματικότητα της συνολικής συμβουλευτικής διαδικασίας, είναι ο κύριος λόγος για τον οποίο τα τελευταία χρόνια, οι προσπάθειες των προγραμματιστών εντείνονται συνεχώς. Με στόχο την κάλυψη των κενών που, αυτή τη στιγμή, εντοπίζονται στην προσομοίωση της ανθρώπινης αλληλεπίδρασης, οι προγραμματιστές εστιάζουν στη βελτίωση των παραμέτρων εκείνων, που σχετίζονται άμεσα με αυτήν.

Μια από τις παραμέτρους αυτές σχετίζεται με την ενίσχυση των ανθρώπινων χαρακτηριστικών των εφαρμογών. Διάφορες πειραματικές μελέτες έχουν δείξει πως όταν η μορφή των συστημάτων της T.N. πλησιάζει στην ανθρώπινη, τότε η αίσθηση της οικειότητας που βιώνει ένας χρήστης αυξάνεται, το ίδιο και η τάση του να αντιμετωπίζει το συγκεκριμένο σύστημα ως έμψυχο, ακόμη και αν ήδη γνωρίζει πως δεν είναι.²⁵⁶ Μια τέτοια διαδικασία έχει βαθιές ψυχολογικές προεκτάσεις. Η απόδοση ανθρώπινων χαρακτηριστικών στα ζώα, ή ακόμα και σε υλικά αντικείμενα αποτελεί μια φυσική τάση του ανθρώπου, που ονομάζεται *ανθρωπομορφισμός*.²⁵⁷ Ο ανθρωπομορφισμός πηγάζει από την βαθύτερη ανάγκη του ανθρώπου να κατανοήσει τον κόσμο που τον περιβάλλει και να δραστηριοποιηθεί μέσα σε αυτόν. Η αναγνώριση της ανθρώπινης αυτής ιδιότητας έχει κινητοποιήσει τις προσπάθειες των προγραμματιστών που τα τελευταία χρόνια, εστιάζουν στο να αυξήσουν το στοιχείο των ανθρώπινων χαρακτηριστικών, ώστε με τον τρόπο αυτό, να εκμαιεύσουν από τους

²⁵⁴ Berry et al., “Assessing Therapeutic Alliance,” 6.

²⁵⁵ “The Frame Problem.”

²⁵⁶ Κωστέλετος, *Τεχνητή Νοημοσύνη και Ηθική, Ενότητα 10^η* [πανεπιστημιακές σημειώσεις], ΕΚΠΑ.

²⁵⁷ “The Psychology of Chatbots: When interacting with bots our brain believes it's chatting with a human,” Psychology Today, accessed October 10, 2023, <https://www.psychologytoday.com/intl/blog/behind-online-behavior/201607/the-psychology-chatbots>.

χρήστες την ανθρωπομορφική αντιμετώπιση των συστημάτων της Τ.Ν.²⁵⁸ Έτσι, η απόδοση ονομάτων στα chatbot, ανθρώπινης μορφής (εικονικά *avatar*) ή και φωνής, δημιουργούν μια λανθασμένη νοητική αντίληψη της αλληλεπίδρασης, κάνοντας τον χρήστη να πιστεύει ότι συνομιλεί με κάποιον άλλον άνθρωπο και να αποδίδει στο chatbot την ιδιότητα νοητικών καταστάσεων ή επιπλέον ανθρώπινων χαρακτηριστικών, τα οποία στην πραγματικότητα δεν διαθέτει (λ.χ. ενσυναίσθηση).²⁵⁹

Τα παραδείγματα που απεικονίζουν την προσπάθεια ανθρωπομορφισμού των chatbot είναι πολλά. Έχουμε ήδη αναφερθεί στο ELIZA, το πρώτο chatbot στο ρόλο ενός ψυχοθεραπευτή, που δημιουργήθηκε το 1966 από τον Αμερικανό επιστήμονα της πληροφορικής και της Τ.Ν., Joseph Weizenbaum. Το σύστημα ELIZA στην προσπάθειά του να προσομοιώσει μια θεραπευτική συνεδρία ροτζεριανής θεραπείας, χρησιμοποίησε την τεχνική της αντιστοίχισης προτύπων και την αντικατάσταση αντωνυμιών προκειμένου να δώσει την ψευδαίσθηση κατανόησης, παρόλο που δεν είχε ενσωματωμένες γνώσεις.²⁶⁰ Η χρήση του, πρώτη κατέστησε εμφανή την τάση των ανθρώπων να αντιδρούν διαφορετικά στα συστήματα της Τ.Ν. που φέρουν ανθρώπινο όνομα και συστήνονται στο κοινό μαζί με σύντομες ιστορίες που αφορούν τη μέχρι εκείνη τη στιγμή ύπαρξή τους. Η τάση αυτή, αντιμετώπισης των συστημάτων της Τ.Ν. ως πρόσωπα και απόδοσης νοητικών καταστάσεων σε αυτά, περιγράφεται μέσα από τον όρο *ELIZA Effect*, ο οποίος έχει επικρατήσει μέχρι σήμερα.²⁶¹

Έκτοτε, γνωρίζουμε ένα σημαντικό αριθμό chatbot με ανθρώπινα ονόματα, όπως τα chatbot Alexa, Siri και Cortana, ενώ ενδεικτικά, στα εικονικά avatar της εφαρμογής The Digital People Platform²⁶² διατίθενται, προς επιλογή από το χρήστη, διαφορετικοί εικονικοί βοηθοί (*avatar*) με ανθρώπινη μορφή, που χαμογελούν ζεστά ή γνέφουν κατά τη διάρκεια της συνομιλίας, δίνοντας την εντύπωση ότι τον ακούν προσεκτικά. Μάλιστα, κατά την προφορική τους απάντηση, τα avatars προσφωνούν το όνομα του χρήστη, ένα στοιχείο που αυξάνει την αίσθηση της εξατομίκευσης της συνομιλίας και τις πιθανότητες της συναισθηματικής σύνδεσης με την εφαρμογή.

Θα είχε νόημα να σταθούμε στο σημείο αυτό, στο χαρακτηριστικό της φωνής, μιας και ένα chatbot που μελλοντικά θα μπορούσε να αντικαταστήσει έναν άνθρωπο σύμβουλο στις γραμμές τηλεφωνικής βοήθειας και υποστήριξης, θα έπρεπε να κατέχει το χαρακτηριστικό αυτό. Η φωνή είναι το εργαλείο εκείνο που, από τα πρώτα χρόνια της ζωής του, ένας άνθρωπος χρησιμοποιεί για να επικοινωνήσει με τους άλλους. Όταν λοιπόν, ένα σύστημα της Τ.Ν. διαθέτει το χαρακτηριστικό αυτό, τότε το περιβάλλον της συνομιλίας τείνει να βιώνεται ως φυσικό και φιλικό.²⁶³ Σύμφωνα με τους Rzepka et al.,²⁶⁴ η επικοινωνία μέσω φωνής συνοδεύεται από υψηλότερη ικανότητα αντίληψης, χαμηλότερη γνωστική προσπάθεια κατανόησης και μεγαλύτερο βαθμό ικανοποίησης,

²⁵⁸ Κωστελέτος, *Τεχνητή Νοημοσύνη και Ηθική, Ενότητα 10^η* [πανεπιστημιακές σημειώσεις], ΕΚΠΑ.

²⁵⁹ “The Psychology of Chatbots.”; Κωστελέτος, *Τεχνητή Νοημοσύνη και Ηθική, Ενότητα 10^η* [πανεπιστημιακές σημειώσεις], ΕΚΠΑ.

²⁶⁰ Sweeney et al., “Can Chatbots Help Support a Person’s Mental Health?” 2.

²⁶¹ Κωστελέτος, *Τεχνητή Νοημοσύνη και Ηθική, Ενότητα 10^η* [πανεπιστημιακές σημειώσεις], ΕΚΠΑ.

²⁶² “The Digital People Platform,” D-ID, accessed October 6, 2023, <https://chat.d-id.com/>.

²⁶³ Zhu, Wang, Pu, “I am chatbot, your virtual mental health adviser,” 4.

²⁶⁴ Christine Rzepka, Benedikt Berger, Thomas Hess, “Voice Assistant vs. Chatbot – Examining the Fit Between Conversational Agents’ Interaction Modalities and Information Search Tasks,” *Information Systems Frontiers* 24 no. 2 (June, 2022): 1-18, <https://link.springer.com/article/10.1007/s10796-021-10226-5>.

σε σύγκριση με την επικοινωνία μέσω κειμένου.²⁶⁵ Επιπλέον, η φωνή αποτελεί ένα ισχυρό φυσικό διεγερτικό ερέθισμα που μπορεί να προσελκύσει την προσοχή και να μας εμπλέξει συναισθηματικά, ακόμη και στο πλαίσιο μιας εφαρμογής T.N. που χρησιμοποιείται στον τομέα της ψυχικής υγείας.²⁶⁶

Υπάρχουν διάφορες έρευνες που επιβεβαιώνουν τα παραπάνω. Στη μελέτη των Mancone et al. γίνεται αναφορά σε πολλές εξ αυτών που αποδεικνύουν πως η κατάλληλη προσωδία στη φωνή, δηλαδή η ευχάριστη χροιά, ο σταθερός ρυθμός και ένταση, μπορούν να επηρεάσουν θετικά της ποιότητα της σχέσης που αναπτύσσεται. Συγκεκριμένα, η επικοινωνία μόνο μέσω φωνής επιτρέπει στον χρήστη να επικεντρωθεί στην επικοινωνία, εγείρει τη δημιουργία συναισθημάτων από πλευράς του και δύναται να αυξήσει τη συμμετοχή του μέσω της συναισθηματικής έκφρασης. Επιπλέον, η φωνητική ικανότητα ενός chatbot, δημιουργεί στον χρήστη την αίσθηση ότι αυτό τον κατανοεί, ενώ παράλληλα συμβάλλει στη βελτίωση της αντίληψης του περί τέτοιου είδους μεθόδους.²⁶⁷

Ομοίως, η εργασία των Niculescu et al.,²⁶⁸ η οποία μελέτησε ανθρώπινες εκφράσεις όπως η συμπάθεια και το χιούμορ, διαπίστωσε ότι αυτές επηρέασαν θετικά την ποιότητα της σχέσης ανάμεσα στον άνθρωπο και ένα ρομπότ κοινωνικής αλληλεπίδρασης (*Social Robot*), ενώ όταν στο τέλος του πειράματος, το ρομπότ προγραμματίστηκε να είναι πιο εξωστρεφές, εκδηλωτικό και αστείο, αυτό έτυχε ακόμη μεγαλύτερης αποδοχής από τους χρήστες, που αξιολόγησαν θετικά τη συνολική ποιότητα της αλληλεπίδρασης και αναγνώρισαν την ικανότητά του να διεξάγει μια διασκεδαστική και απολαυστική συνομιλία. Το ίδιο μπορεί να συμβεί και με άλλες συμπεριφορές που συμπεριλαμβάνονται στην ανθρώπινη επικοινωνία, όπως η ευγένεια, οι άτυπες σύντομες κουβέντες γύρω από μη σχετικά ζητήματα (*small talk*), η ειδική ορολογία (*expert jargon*), ο κοινωνικός σχολιασμός (*gossip*) κ.ο.κ., που ενσωματώνονται στις εφαρμογές της T.N. με στόχο την δημιουργία της εμπιστοσύνης από την πλευρά του χρήστη.²⁶⁹

Οφείλουμε ωστόσο, στο σημείο αυτό να κάνουμε μια αναφορά στο φαινόμενο του *Uncanny Valley* (παράξενη κοιλάδα), που το 1970, επεσήμανε ο Ιάπωνας ψυχολόγος Masahiro Mori, γνωστός για τη συνεισφορά του στον τομέα της ρομποτικής και της T.N.²⁷⁰ Σύμφωνα με τον Mori, η οικειότητα και συμπάθεια των ανθρώπων προς τα ρομπότ, αυξάνεται όσο αυξάνεται και η ανθρωπόμορφη εμφάνισή τους. Η τάση αυτή όμως, αντιστρέφεται όταν η ομοιότητα στον άνθρωπο φτάσει σε ένα υπερβολικά μεγάλο, όχι όμως, απόλυτο βαθμό. Τότε, συναισθήματα αποστροφής και ανασφάλειας προκαλούνται στους ανθρώπους που έρχονται σε επαφή με το ρομπότ. Τα συναισθήματα αυτά γίνονται ακόμη πιο έντονα, όταν το ρομπότ κινείται και κατά τον

²⁶⁵ Stefania Mancone, Pierluigi Diotaiuti, Giuseppe Valente, Stefano Corrado, Fernando Bellizzi, Guilherme Torres Vilarino, Alexandro Andrade, “The Use of Voice Assistant for Psychological Assessment Elicits Empathy and Engagement While Maintaining Good Psychometric Properties,” *Behavioral Sciences* 13, no. 7, 550 (July, 2023): 3, <https://www.mdpi.com/2076-328X/13/7/550>.

²⁶⁶ *Αυτόθι*, 2.

²⁶⁷ *Αυτόθι*, 13-14.

²⁶⁸ Andreea Niculescu, Betsy van Dijk, Anton Nijholt, Haizhou Li, Swee Lan See, “Making Social Robots More Attractive: The Effects of Voice Pitch, Humor and Empathy,” *International Journal of Social Robotics* 5 (Jan., 2013): 171–191, <https://link.springer.com/article/10.1007/s12369-012-0171-x>.

²⁶⁹ Mancone et al., “The Use of Voice Assistant,” 13-14.

²⁷⁰ “The Uncanny Valley: The Original Essay by Masahiro Mori: “The Uncanny Valley” by Masahiro Mori is an influential essay in robotics. This is the first English translation authorized by Mori,” IEEE Spectrum, accessed October 20, 2023, <https://spectrum.ieee.org/the-uncanny-valley>.

Mori, αυτό ίσως συμβαίνει επειδή η μεγάλη ομοιότητα δημιουργεί στον άνθρωπο συνειρμούς θανάτου, μιας και η μορφή του ρομπότ πλησιάζει κατά πολύ αυτή ενός κινούμενου πτώματος.²⁷¹ Η παραπάνω υπόθεση του Mori έχει αμφισβητηθεί από πολλούς ερευνητές,²⁷² που υποστηρίζουν μεταξύ άλλων, πως το φαινόμενο του Uncanny Valley εξαρτάται από πολιτισμικούς παράγοντες και έτσι, η εξομάλυνσή του με το πέρασμα του χρόνου είναι πολύ πιθανή.²⁷³ Γενικότερα, οι ερευνητές αυτοί υποστηρίζουν πως η εμπειρική απόδειξη της ύπαρξης του φαινομένου παραμένει αν όχι ανύπαρκτη, σημαντικά ασαφής.²⁷⁴

Στον αντίποδα, η ύπαρξη του Uncanny Valley έχει επιβεβαιωθεί μέσα από αρκετές μελέτες.²⁷⁵ Οι περισσότερες εξ αυτών αφορούν στον τομέα του εμπορίου, ενώ κατά την αναζήτηση, δε στάθηκε δυνατή η εύρεση ερευνών, σχετικών με τη χρήση των συστημάτων της Συνομιλητικής Τ.Ν. (Conversational AI) στον τομέα της ψυχικής υγείας.²⁷⁶ Πιθανολογείται ωστόσο, ότι ο ανθρωπομορφισμός της λεκτικής έκφρασης δεν θα προκαλέσει συναισθήματα φόβου ή απέχθειας στους χρήστες, όταν η επικοινωνία λαμβάνει χώρα μόνο μέσω φωνής. Αυτό διότι η απουσία εικονικής μορφής και οπτικών ανθρωπομορφικών χαρακτηριστικών θα αποτρέψει τους χρήστες από σκέψεις αναφορικά με την οντολογική φύση του συνομιλητή τους.²⁷⁷

Ολοκληρώνοντας, είμαστε σε θέση με βάση τα παραπάνω να κατανοήσουμε τον τρόπο, με τον οποίο η ποιότητα ή και η απόδοση ανθρωπίνων χαρακτηριστικών στα συστήματα της Τ.Ν. μπορεί να συμβάλει στη δημιουργία ενός οικείου περιβάλλοντος, εντός του οποίου το άτομο θα εκφραστεί με ειλικρίνεια και θα αισθανθεί ότι εμπλέκεται σε μία αυθεντική θεραπευτική σχέση. Τα ερωτήματα περί αυθεντικότητας της θεραπευτικής σχέσης με ένα σύστημα της Τ.Ν. εγείρονται από την ουσιώδη διαφορά που υπάρχει ανάμεσα στο να δείχνει κανείς νοήμων και στο να είναι πραγματικά. Εξ αυτής της διαφοράς προκύπτουν επίσης, ζητήματα περί της ασφάλειας ενός χρήστη, που διακυβεύεται από τις πιθανότητες εσφαλμένων εκτιμήσεων και εξαπάτησης στο πλαίσιο μιας μονομερούς σχέσης με ένα chatbot. Το δίπολο αυτό, αυθεντικότητας και εξαπάτησης, που είναι ζωτικής σημασίας για την αξιοπρέπεια και την αυτονομία ενός ανθρώπου θα μας απασχολήσει στο επόμενο κεφάλαιο περί Ηθικής.

Προτού φτάσουμε εκεί, διερευνώντας τη δυνατότητα επίτευξης της ουσίας και της αυθεντικότητας της σχέσης που αναπτύσσεται ανάμεσα στον άνθρωπο και σε ένα σύστημα της Τ.Ν., θα πρέπει να επαναφέρουμε στην συζήτησή μας τις δεξιότητες εκείνες που θα πρέπει να φέρει ο σύμβουλος κατά τη διεξαγωγή της συμβουλευτικής διαδικασίας. Οι δεξιότητες αυτές αναφέρθηκαν στο πρώτο κεφάλαιο της παρούσας

²⁷¹ Αλέξανδρος Λάγιος, «Ενσυναίσθηση και Τεχνητή Νοημοσύνη» (Μεταπτυχιακή διατριβή, Εθνικό Μετσόβειο Πολυτεχνείο, 2016), 82, <https://dspace.lib.ntua.gr/xmlui/handle/123456789/43823>.

²⁷² *Αυτόθι*, 83.

²⁷³ Harry Brenton, Marco Gillies, Daniel Ballin, David Chatting, “The uncanny valley: Does it exist?” Workshop on Human-Animated Character Interaction - HCI 2005 (Sep., 2005): 3, <http://www.davidchatting.com/research/uncanny-valley-hci2005.pdf>.

²⁷⁴ Maya Mathur, David Reichling, “Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley,” *Cognition* 146 (Jan., 2016): 22, <https://www.sciencedirect.com/science/article/pii/S0010027715300640>.

²⁷⁵ Λάγιος, «Ενσυναίσθηση και Τεχνητή Νοημοσύνη», 83.

²⁷⁶ Tingru Cui, Xinwei Wang, Jianzhong Qi, “Designing Anthropomorphic Therapeutic Conversational Agents: An Uncanny Valley Perspective,” *AIS Electronic Library (AISEL)* (June, 2021): 1-2, https://aisel.aisnet.org/ecis2021_rip/?utm_source=aisel.aisnet.org%2Fecis2021_rip%2F30&utm_medium=PDF&utm_campaign=PDFCoverPages.

²⁷⁷ *Αυτόθι*, 5.

διατριβής και συνοπτικά περιλαμβάνουν την ανατροφοδότηση, την ειλικρίνεια, την αυτο-αποκάλυψη, την άνευ όρων αποδοχή, την ενσυναίσθηση, το σεβασμό και την ύπαρξη ενός σταθερού συστήματος αξιών και ενός σταθερού κώδικα ηθικής συμπεριφοράς. Στην επόμενη ενότητα θα διερευνήσουμε το εάν οι δεξιότητες αυτές, δύναται να πληρούνται αντίστοιχα από ένα σύστημα της T.N που θα αναλάβει το ρόλο του συμβούλου.

3.δ. Οι δεξιότητες του συμβούλου κατά τη συμβουλευτική διαδικασία.

Με βάση τα όσα έχουμε δει έως τώρα, το να ξεκινήσουμε από την δεξιότητα της ανατροφοδότησης, μοιάζει να είναι το πιο ασφαλές. Με δεδομένη την ικανότητα ενός συστήματος να καταγράφει και να επεξεργάζεται δεδομένα, οι πληροφορίες που έχουν συλλεχθεί κατά τη διάρκεια της αλληλεπίδρασης είναι εφικτό να αποτυπωθούν, κατηγοριοποιημένες και σε συνοπτική μορφή, για παράδειγμα μέσα από γραφήματα, που συνοδεύονται από επιπλέον πληροφορίες, προτροπές ή και το σχεδιασμό περαιτέρω θεραπευτικών στόχων. Στην έρευνα των Clarke et al., κάτι τέτοιο αποδείχτηκε ιδιαίτερα βοηθητικό για ορισμένους από τους χρήστες της εφαρμογής myCompass, που σχολίασαν πως με τον τρόπο αυτό μπόρεσαν να προβούν σε αλλαγές, εντοπίζοντας τις αντιδράσεις τους σε συγκεκριμένες καταστάσεις και τα σημεία εκείνα, στα οποία θα έπρεπε να εστιάσουν στη συνέχεια. Κάποιοι άλλοι ωστόσο, ανέφεραν πως η ανατροφοδότηση που παρείχε η εφαρμογή, είτε δεν ήταν ακριβής, υποβαθμίζοντας κάποια συμπτώματα και υπερθεματίζοντας άλλα ανύπαρκτα, είτε τους δημιουργούσε ένα συναίσθημα ψυχολογικής πίεσης, λόγω των πολλών επισημάνσεων γύρω από τις διαδικασίες της θεραπείας που θα έπρεπε να φέρουν εις πέρας.²⁷⁸

Ίσως η δυσχέρεια αυτή, κάποιων συμμετεχόντων, να έχει τις ρίζες της στο γεγονός πως, αν και μια εφαρμογή T.N. μπορεί να ανατροφοδοτήσει το άτομο που τη χρησιμοποιεί, ως προς την εξέλιξή του στην συμβουλευτική διαδικασία, η ανατροφοδότηση αυτή συμβαίνει σε ένα επίπεδο πρακτικό, μετατρέποντας σε ποσοτικά και αντικειμενικά δεδομένα, υποκειμενικά ανθρώπινα χαρακτηριστικά, όπως τα συναισθήματα και οι πεποιθήσεις.²⁷⁹ Ναι μεν, η ποσοτικοποίηση αυτή μπορεί να βοηθήσει τους χρήστες να «ετικετοποιήσουν» τις υποκειμενικές τους εμπειρίες και να οργανώσουν τις σκέψεις τους,²⁸⁰ ωστόσο ενέχεται ο κίνδυνος να αποσπάσει την προσοχή τους από την ποιότητα και την ουσία των εσωτερικών υποκειμενικών τους καταστάσεων. Με άλλα λόγια, τα συναισθήματα είναι αυτά που, αξιολογώντας μια κατάσταση ή ένα περιβάλλον, καθοδηγούν την ανθρώπινη δράση και τη δημιουργία νοημάτων. Αντίθετα, τα ψηφιακά δεδομένα δεν είναι παρά αποσπασματικές εγγραφές, που είναι διαθέσιμες προς ανάλυση, αλλά δεν συνδέονται άμεσα με τις ανθρώπινες ενέργειες ή αξιολογήσεις. Ως εκ τούτου, η ταύτιση σε αριθμούς και δεδομένα, μπορεί να απειλήσει την προσωπική ακεραιότητα, απομακρύνοντας τον άνθρωπο από τις βιωμένες εμπειρίες και τη μοναδικότητα του ψυχισμού του. Κάτι τέτοιο αντιτίθεται στη γενικότερη ουσία της

²⁷⁸ Clarke et al., “Therapeutic Alliance With a Fully Automated Mobile Phone and Web-Based Intervention,” 8.

²⁷⁹ Sedlakova, Trachsel, “Conversational Artificial Intelligence in Psychotherapy,” 9.

²⁸⁰ Kretschmar et al., “Can Your Phone Be Your Therapist?” 3-4.

συμβουλευτικής διαδικασίας, όπου οι εσωτερικές καταστάσεις κατέχουν εξέχοντα ρόλο και αποτελούν την κινητήρια δύναμη για τη θεραπευτική αλλαγή.²⁸¹

Επιπλέον, η γνώση της διαρκούς παρακολούθησης και καταγραφής ενδέχεται να επηρεάσει το περιεχόμενο των δηλώσεων ενός χρήστη, στην προσπάθειά του να το διαμορφώσει με βάση το τι θα ήθελε να παρατηρήσει σε αυτόν ο ψηφιακός του θεραπευτής. Τίθενται έτσι, υπό αμφισβήτηση η αντικειμενικότητα των συμπερασμάτων που εξάγονται και οι θεραπευτικές μέθοδοι, που στη συνέχεια εφαρμόζονται.²⁸² Σύν τοις άλλοις όμως, η αίσθηση αυτή, μιας παρακολούθησης που εισέρχεται εις βάθος, αν αναλογιστούμε την ικανότητα των εργαλείων της Τ.Ν. να καταγράφουν από τις καθημερινές ασχολίες μέχρι την εγκεφαλική δραστηριότητα, θα μπορούσε να οδηγήσει το άτομο σε μία συνθήκη διαρκούς αυτο-επιτήρησης. Κατά την αναφορά μας στο Διαδίκτυο των Πραγμάτων (Internet of Things - IoT) στην ενότητα 2^α, είδαμε πως, όπως έχει επισημανθεί από τους ερευνητές, η αίσθηση της παρακολούθησης δύναται να επηρεάσει το άτομο, ώστε να διαφοροποιήσει την καθημερινή του συμπεριφορά για το διάστημα που αυτή καταγράφεται.²⁸³ Η αυτο-επιτήρηση αυτή μπορεί τελικά, να μεταβάλλει τον ποιοτικό τρόπο βίωσης εμπειριών και συναισθημάτων από το άτομο, ώστε να εντάσσεται στο πλαίσιο αυτού που ορίζεται ως «φυσιολογικό». Εγείρονται λοιπόν, σκέψεις περί Πανοπτισμού και ανησυχίες για την πιθανότητα παραβίασης της νοητικής ιδιωτικότητας του ατόμου και πρόκλησης βλάβης στην ψυχική του υγεία, μέσα από εσφαλμένες εκτιμήσεις και θεραπευτικές παρεμβάσεις. Το θέμα αυτό, το οποίο λαμβάνει ηθικές προεκτάσεις, θα προσεγγίσουμε στο 4^ο κεφάλαιο της παρούσας διατριβής, όπου θα εξετάσουμε μέσω Ηθικής την εφαρμογή των εν λόγω συστημάτων στον τομέα της ψυχικής υγείας.

Πέρα από τα παραπάνω όμως, η ευκολία με την οποία ένα σύστημα Τ.Ν. προβαίνει στην ακριβή καταγραφή και ανάλυση της ψυχικής υγείας ενός ατόμου, είναι πιθανό να ενισχύσει το φαινόμενο της *υπερδιάγνωσης* (*overdiagnosis*) ή της εξαγωγής διαγνώσεων, που δηλώνουν τάσεις εμφάνισης διαταραχών και έτσι, δεν επιδέχονται παρεμβάσεων.²⁸⁴ Με τον όρο υπερδιάγνωση, ορίζονται οι καταστάσεις εκείνες στις οποίες, ενώ ένα άτομο πληροί τα διαγνωστικά κριτήρια μιας διαταραχής (ή της προδιάθεσης εμφάνισης της), η εξακρίβωση αυτής μπορεί να έχει επηρεάσει αρνητικά την ποιότητα της ζωής του. Στην ουσία, η υπερδιάγνωση δεν αποτελεί ένα λανθασμένο πόρισμα, αλλά ένα πόρισμα που μπορεί να είναι σωστό σύμφωνα με τα διαγνωστικά κριτήρια, αλλά να μην έχει κανένα ιατρικό και ψυχολογικό όφελος για το άτομο και να του προκαλέσει βλάβη.²⁸⁵

Αναμφισβήτητα, η ευρεία διάδοση και χρήση των τεχνολογιών, που διευκολύνουν τη συλλογή μεγάλου όγκου δεδομένων αναφορικά με τη συμπεριφορά των ατόμων (λ.χ. smartphones, smartwatches), παρέχουν σημαντικές ευκαιρίες για την ανίχνευση

²⁸¹ Sedlakova, Trachsel, "Conversational Artificial Intelligence in Psychotherapy," 9.

²⁸² Το ίδιο έχει παρατηρηθεί και κατά τη συμπλήρωση ερωτηματολογίων, όπου στην προσπάθειά τους να λάβουν την αποδοχή των άλλων, οι ερωτώμενοι τείνουν να αποκρύπτουν τις αληθινές απόψεις και εμπειρίες τους (Social Desirability Bias). Ως αποτέλεσμα τίθεται υπό αμφισβήτηση η αντικειμενικότητα των πορισμάτων που προκύπτουν. Βλ. Nikolopoulou, "What is Social Desirability Bias?".

²⁸³ Rozgonjuk et al., "The association between problematic smartphone use, depression and anxiety symptom severity," 15.

²⁸⁴ Blasimme, Vayena, "The Ethics of AI in Biomedical Research," 710.

²⁸⁵ Daniel Capurro, Simon Coghlan, Douglas Pires, "Preventing Digital Overdiagnosis," *JAMA The Journal of the American Medical Association* 327, no. 6 (Feb., 2022): 525, https://www.researchgate.net/publication/358016994_Preventing_Digital_Overdiagnosis/citations.

ασθενειών σε πρώιμα στάδια και τη βελτίωση της υγείας σε επίπεδο πληθυσμιακό.²⁸⁶ Ωστόσο, η ευκολία με την οποία οι διαγνωστικοί αλγόριθμοι μπορούν να εφαρμοστούν σε εύκολα προσβάσιμα σύνολα δεδομένων και να δημιουργήσουν καταστάσεις υπερδιάγνωσης είναι προβληματική,²⁸⁷ καθώς μπορεί να οδηγήσει στην *ιατρικοποίηση* (*medicalization*) των καταστάσεων, μια διαδικασία κατά την οποία μη-ιατρικά προβλήματα ορίζονται και αντιμετωπίζονται ως νοσήματα ή διαταραχές.²⁸⁸

Συνέπειες όπως οι αδικαιολόγητες επισκέψεις στους ειδικούς, η διαρκής ανησυχία και η ψυχολογική επιβάρυνση που φέρει η επισήμανση ατόμων ως «ασθενείς», καθιστά την αποφυγή της υπερδιάγνωσης ηθική υποχρέωση. Μια λύση που προτείνεται αφορά στην καλύτερη εκπαίδευση των αλγορίθμων μηχανικής μάθησης για την ακριβέστερη ανίχνευση των κλινικά σημαντικών ασθενειών στον γενικό πληθυσμό (λ.χ. η πρώιμη ανίχνευση της νόσου του Parkinson μέσα από δεδομένα όπως το μοτίβο πληκτρολόγησης στον H/Y ή η ανίχνευση της κοιλιακής μαρμαρυγής μέσα από ρολόγια τύπου smartwatches), ή και στην αξιοποίηση των δεδομένων, προκειμένου να κατανοηθεί ο τρόπος της ταξινόμησης των ανθρώπων ως «ασθενείς» και να οριστούν τα υποσύνολα των ασθενών εκείνα, που δεν θα επωφεληθούν από την ταξινόμηση αυτή.²⁸⁹ Το βέβαιο είναι πως στο πλαίσιο της ψηφιακής επανάστασης που λαμβάνει χώρα στον τομέα της ψυχικής υγείας, η υπάρχουσα γνώση θα πρέπει να συμπεριλάβει συν τοις άλλοις, τις διαδικασίες πρόληψης της υπερδιάγνωσης, μέσα από την αξιολόγηση των διαγνωστικών αλγορίθμων ως προς την ασφάλη και ωφέλιμη για το κοινό, χρήση τους.²⁹⁰

Προχωρώντας στη δεξιότητα της ειλικρίνειας, θα λέγαμε πως εάν διερευνήσουμε την έννοια αυτή ως την απλή δήλωση αληθών δηλώσεων, τότε εν μέρει, ένα σύστημα της T.N. διαθέτει το χαρακτηριστικό αυτό. Σε σχετική δοκιμή με τα υπάρχοντα chatbot (όπως το ChatGpt ή το D-ID),²⁹¹ αυτά, σταθερά δηλώνουν πως είναι λογισμικά προγράμματα υπολογιστή, βασισμένα στην T.N., μη ενσώματα, που δεν έχουν φυσική ύπαρξη ή συναισθήματα και λειτουργούν μέσω υπολογιστικών διεργασιών προκειμένου να αλληλεπιδρούν με τους χρήστες. Από αυτή την άποψη λοιπόν, των δηλώσεων σχετικά με την ταυτότητα κάποιου, ένα chatbot είναι σίγουρα ειλικρινές. Αρκεί όμως αυτό;

Ας επανέλθουμε στην δεξιότητα της ανατροφοδότησης, μια έννοια στενά συνδεδεμένη με την ειλικρίνεια. Όπως είδαμε, στο πλαίσιο της συμβουλευτικής διαδικασίας, το μοίρασμα των σκέψεων, των συναισθημάτων και των προσωπικών βιωμάτων από τον σύμβουλο, είναι αυτό που θα ενθαρρύνει το συμβουλευόμενο να κάνει το ίδιο και θα θεμελιώσει μια σχέση εμπιστοσύνης ανάμεσα τους.²⁹² Η δεξιότητα της αυτο-

²⁸⁶ *Αυτόθι*, 525.

²⁸⁷ Για παράδειγμα, η χρήση αλγορίθμων βαθιάς μάθησης για να εντοπίσουμε τον κίνδυνο καρδιαγγειακής πάθησης μέσα από εικόνες του βυθού του αμφιβληστροειδούς θα ήταν χρήσιμη, αν λάβουμε υπόψη τις επιπτώσεις της πάθησης στον τρόπο ζωής των ασθενών. Όμως, η αντίστοιχη διαδικασία για τον εντοπισμό του κινδύνου εμφάνισης άνοιας δεν ωφελεί, με δεδομένη την απουσία στοιχείων που αποδεικνύουν την αποτελεσματικότητα των παρεμβάσεων στην επιβράδυνση της. Blasimme, Vayena, “The Ethics of AI in Biomedical Research,” 710.

²⁸⁸ Peter Conrad, Thomas Mackie, Ateev Mehrotra, “Estimating the costs of medicalization,” *Social Science & Medicine* 70, no. 12 (June, 2010): 1943, <https://www.sciencedirect.com/science/article/pii/S0277953610001784>.

²⁸⁹ *Αυτόθι*, 525.

²⁹⁰ *Αυτόθι*, 526.

²⁹¹ Βλ. “The Digital People Platform,”; “ChatGPT,” OpenAI, accessed October 6, 2023, <https://openai.com/chatgpt>.

²⁹² Μαλικιώση – Λοϊζου, *Συμβουλευτική Ψυχολογία*, 294-295.

αποκάλυψης όμως, απαιτεί την ύπαρξη όσων θα μοιραστούν και όπως αναφέρουν οι ίδιες οι εφαρμογές, ένα σύστημα της T.N. δεν διαθέτει εμπειρίες, σκέψεις και συναισθήματα. Αν και είναι λοιπόν, ειλικρινείς ως προς τα χαρακτηριστικά της ταυτότητάς τους, δεν διαθέτουν τη δυνατότητα εκείνη, της ειλικρινούς έκφρασης του εαυτού, που θα θεμελιώσει μια σχέση γνήσια και ουσιαστική. Κατανοούμε λοιπόν, με τον τρόπο αυτό πως αν προσεγγίσουμε την έννοια της ειλικρίνειας ως συστατικό μιας γνήσιας και αυθεντικής σχέσης, τότε τα πράγματα γίνονται ιδιαίτερα πολύπλοκα.

Πιο συγκεκριμένα, στην προσπάθεια επίδειξης κατανόησης, αποδοχής, ενδυνάμωσης ή και ενθάρρυνσης, ένα chatbot μπορεί να προβεί σε δηλώσεις όπως «κατανοώ ότι είναι δύσκολο αυτό που σου συμβαίνει» ή «είσαι πολύ δυνατός που μπορείς να μιλήσεις για αυτό το θέμα» και γενικότερα, να χρησιμοποιήσει εκφράσεις, σαν αυτές που λαμβάνουν χώρα κατά την ανθρώπινη αλληλεπίδραση, δίνοντας στον χρήστη την αίσθηση πως εμπλέκεται νοητικά στη διαδικασία της συνομιλίας.²⁹³ Πόσο αληθές όμως, είναι αυτό;

Επιπλέον, είδαμε στο πρώτο κεφάλαιο τη ζωτική σημασία της ενσυναίσθησης για τη συμβουλευτική διαδικασία. Η ενσυναίσθηση είναι το στοιχείο εκείνο που θα θεμελιώσει τη θεραπευτική συμμαχία ανάμεσα στον σύμβουλο και τον συμβουλευόμενο.²⁹⁴ Με άλλα λόγια, σε μία θεραπευτική σχέση, ένας σύμβουλος καλείται να σταθεί στο πλάι του συμβουλευόμενου, και να τον υποστηρίξει στην προσπάθειά του για προσωπική ανάπτυξη. Έτσι το «κάνω κάτι για τον πελάτη», μας λένε οι Strawbridge & Woolfe (2010), μετατρέπεται σε «είμαι με τον πελάτη», σε μια αληθινή θεραπευτική συμμαχία, που εκτυλίσσεται σε ένα «φιλικό και μη-απειλητικό περιβάλλον»,²⁹⁵ εντός του οποίου ο θεραπευόμενος βιώνει την απαραίτητη αίσθηση επάρκειας, αποδοχής και σεβασμού, που θα διευκολύνει την επίτευξη των προσωπικών του στόχων.²⁹⁶

Σε μία δήλωση όπως «αισθάνομαι λυπημένος», εφαρμογές όπως η Siri μπορεί να απαντήσουν «λυπάμαι που ακούω αυτό». Ωστόσο, η δυνατότητά τους για αληθινή έκφραση της ενσυναίσθησης παραμένει υπό αμφισβήτηση. Αυτό θα μπορούσε να θεωρηθεί ασήμαντο όταν συμβαίνει στο πλαίσιο αλληλεπίδρασης με εφαρμογές που έχουν σχεδιαστεί για διάφορες γενικές χρήσεις. Είναι όμως βαθιά προβληματικό όταν συμβαίνει κατά τη χρήση εφαρμογών που απευθύνονται σε πληθυσμούς με ευάλωτο ψυχισμό. Παρόλο που τα chatbot ψυχικής υγείας συχνά ζητούν από τους χρήστες να αποκαλύψουν πληροφορίες γύρω από ευαίσθητα προσωπικά θέματα, οι αντιδράσεις τους παραμένουν περιορισμένες, καθώς είτε αδυνατούν να κατανοήσουν πολύπλοκες δηλώσεις, υπονομεύοντας έτσι την αίσθηση του χρήστη ότι τον ακούνε προσεκτικά, είτε απλώς επιβεβαιώνουν τη λήψη του κειμένου και ευχαριστούν για αυτό (όπως το chatbot της Koko,²⁹⁷ που παραθέτει δηλώσεις όπως “Thanks for sharing...”). Άλλες φορές, οι απαντήσεις των chatbot χαρακτηρίζονται από ανούσιες έως και επιπόλαιες

²⁹³ Sedlakova, Trachsel, “Conversational Artificial Intelligence in Psychotherapy,” 10.

²⁹⁴ Robert Morris, Kareem Kouddous, Rohan Kshirsagar, Stephen Schueller, “Towards an Artificially Empathic Conversational Agent for Mental Health Applications: System Design and User Perceptions,” *Journal of Medical Internet Research* 20, no. 6, e10148 (June, 2018): 1-2, <http://www.jmir.org/2018/6/e10148/>.

²⁹⁵ Καλαντζή-Αζίζι, Μαλικιώτη-Λοΐζου, «Η θεραπευτική σχέση στην κλινική και συμβουλευτική ψυχολογία», 5.

²⁹⁶ *Αυτόθι*, 3-5.

²⁹⁷ “Free mental health support for millions of young people struggling online,” Koko, accessed October 12, 2023, <https://www.kokocares.org/>.

(π.χ. «Θα σου έδινα τον ώμο μου για να κλάψεις πάνω, αν είχα έναν»), κάτι που μπορεί να είναι προσβλητικό και κυρίως απογοητευτικό για ένα χρήστη σε δυσμενή ψυχολογική κατάσταση που αναζητά υποστήριξη. Ιδιαίτερα, σε αυτού του είδους τις άστοχες αποκρίσεις από τα chatbot γίνεται, περισσότερο από κάθε άλλη φορά, εμφανής η έλλειψη διεπιστημονικότητας κατά το σχεδιασμό των συστημάτων T.N. Η συμμετοχή όμως, εκπροσώπων από τα επιστημονικά πεδία της ψυχικής υγείας (Ψυχιατρική, Ψυχολογία, Συμβουλευτική κ.ο.κ.) είναι απαραίτητη για τα συστήματα αυτά, που απευθύνονται σε ανθρώπους με ευάλωτο ψυχισμό και ενδέχεται να προκαλέσουν ανυπολόγιστο κόστος κατά το λανθασμένο ή επιπόλαιο χειρισμό των περιπτώσεών τους.

Με βάση τις παραπάνω αδυναμίες, καθίσταται εμφανές πως μέχρι και σήμερα, τα chatbot στον τομέα της ψυχικής υγείας δεν είναι σε θέση να επιδείξουν βαθιά κατανόηση της συγκεκριμένης κατάστασης ενός χρήστη, να εμπλακούν σε ένα γνήσιο θεραπευτικό διάλογο και να αγγίξουν την πολυπλοκότητα αυτού, εκφράζοντας ενσυναίσθηση και λαμβάνοντας υπόψη την μοναδική κατάσταση στην οποία βρίσκεται.²⁹⁸

Μπορεί να φαίνεται πως η CAI (Conversational AI) αποτελεί μέρος μιας διαλεκτικής πρακτικής επειδή εμπλέκει τους χρήστες σε μια συνομιλία και παράγει γνώση και η αλήθεια είναι πως, όσο περισσότερο εμπλουτίζεται με ανθρώπινα χαρακτηριστικά, όπως η συναισθηματική ανταπόκριση, τόσο περισσότερο βιώνεται ως υποκείμενο από τους χρήστες. Απέχει όμως κατά πολύ από την ιδιότητα αυτή. Ένας καλός τρόπος για να κατανοήσουμε τις διαφορές της από έναν άνθρωπο σύμβουλο, είναι να ανατρέξουμε στη *πραγματιστική θεωρία του νοήματος*, όπου η γνώση και κατανόηση είναι αναπόσπαστα στοιχεία της αυτουργίας.²⁹⁹

Σύμφωνα με αυτήν, ένας λογικός αυτουργός χαρακτηρίζεται από το ότι είναι σε θέση να υιοθετήσει κανονιστικές θέσεις (όπως δεσμεύσεις ή δικαιώματα) ως προς τις αξιώσεις του, κατά τη διάρκεια της αλληλεπίδρασής του με τους άλλους. Έτσι, ένα άτομο που υποστηρίζει ότι *p*, έχει την υποχρέωση και το δικαίωμα να ενσωματώνει στο σύνολο των πεποιθήσεών του, νέες αξιώσεις που συνδέονται λογικά με το *p*. Αυτό είναι δυνατό, όταν ο λογικός αυτουργός κατανοεί σωστά τις διαφορετικές έννοιες και είναι σε θέση να τις εφαρμόσει.³⁰⁰

Επιπλέον, από πρακτική άποψη, ο λογικός αυτουργός οφείλει να προσφέρει λόγους και αιτιολογήσεις γύρω από τις αξιώσεις και το ίδιο, μπορεί να απαιτήσει και από το συνομιλητή του. Έτσι, στο πλαίσιο της συνομιλίας, οι συμμετέχοντες δεσμεύονται σε ένα κανονιστικό παιχνίδι κατάθεσης και απαίτησης λόγων, το οποίο τους καθιστά, πέρα από λογικούς, κοινωνικούς και ηθικούς αυτουργούς. Το παιχνίδι αυτό είναι διαλεκτικό, γιατί ανάμεσα στους συνομιλητές υπάρχει αμοιβαίος σεβασμός και αμοιβαία

²⁹⁸ Morris et al., "Towards an Artificially Empathic Conversational Agent for Mental Health Applications," 1-2.

²⁹⁹ Ορισμένα κλασσικά έργα που εστιάζουν σε αυτήν ανήκουν στον Γερμανό φιλόσοφο και κοινωνιολόγο Jürgen Habermas και είναι τα εξής: Jürgen Habermas, *The Theory of Communicative Action: Reason and the Rationalization of Society* (Vol. 1), trans. Thomas McCarthy (Cambridge: Polity Press, 2015); Jürgen Habermas, *The Theory of Communicative Action. Lifeworld and System: A Critique of Functionalist Reason* (Vol. 2), trans. Thomas McCarthy (Cambridge: Polity Press, 1989); Jürgen Habermas, *On the Pragmatics of Social Interaction: Preliminary Studies in the Theory of Communicative Action*, trans Barbara Fultner (Cambridge, Massachusetts: The MIT Press, 2001); Jürgen Habermas, *Truth and Justification*, trans. and ed. by Barbara Fultner (Cambridge, Massachusetts: The MIT Press, 2001).

³⁰⁰ Sedlakova, Trachsel, "Conversational Artificial Intelligence in Psychotherapy," 7-8.

αναγνώριση της λογικής αυτοουργίας του άλλου και της δύναμης του επί των αξιώσεών του.³⁰¹

Στο παιχνίδι αυτό της διαλεκτικής, ένας αυτοουργός είναι αυτόνομος, με την έννοια ότι μπορεί να αναλάβει την ευθύνη των αξιώσεών του, αλλά και γιατί αναγνωρίζεται ως αυτόνομος από τον άλλο. Την αυτονομία αυτή, μπορεί να ενισχύσει η απόκτηση νέας γνώσης, πράγμα που διαδραματίζει σημαντικό ρόλο στο περιβάλλον της συμβουλευτικής διαδικασίας. Ένας σύμβουλος δηλαδή, όντας ικανός να προσφέρει νέους λόγους, δύναται να διευρύνει το ορίζοντα σκέψης του συμβουλευόμενου και να αυξήσει την αυτονομία του. Αυτός, αναγνωρίζοντας τη λογική ικανότητα και την αυτονομία του συμβούλου, διευκολύνεται στο να είναι πιο ανοιχτός σε νέους τρόπους σκέψης και είναι αυτοί, που τελικά θα οδηγήσουν στη θεραπευτική αλλαγή.³⁰²

Οι παραπάνω προϋποθέσεις δεν πληρούνται από ένα σύστημα της Συνομιλητικής T.N. Ένα chatbot δεν είναι ικανό να κατανοήσει έννοιες, παρά το γεγονός πως προσομοιώνει μια ανθρώπινη συνομιλία. Έτσι, αφενός δεν μπορεί να προσφέρει λόγους για τις αξιώσεις του, με τον ίδιο τρόπο που το κάνει ένας άνθρωπος. Για παράδειγμα, στην ερώτηση του γιατί θα πρέπει κανείς να αντιμετωπίζει τις αρνητικές του σκέψεις, ένα chatbot θα μπορούσε να προσφέρει ορισμούς ή περιγραφές συγκεκριμένων τεχνικών για το σκοπό αυτό. Δεν θα μπορούσε ωστόσο, να κατανοήσει επαρκώς το ερώτημα του χρήστη, τις συνθήκες κάτω από τις οποίες το θέτει, καθώς και τις συγκεκριμένες εκείνες λεπτομέρειες της μοναδικής του ύπαρξης, ώστε η απάντηση να είναι εξατομικευμένη, κατανοητή, συσχετισμένη με τις αρχές της συμβουλευτικής διαδικασίας και να επιφέρει μια επιτυχημένη αλλαγή.³⁰³

Αυτό γιατί, η λειτουργία ενός chatbot, ως σύστημα της Συνομιλητικής T.N. (Conversational AI), καθοδηγείται από αλγορίθμους και βασίζεται μόνο σε ποσοτικά δεδομένα και τα εργαλεία επεξεργασίας αυτών.³⁰⁴ Είναι ένα σύστημα πολύπλοκης ανάπτυξης και ιδιαίτερα αποτελεσματικό στην επεξεργασία και την αξιολόγηση δεδομένων, στην εξαγωγή συμπερασμάτων γύρω από την ψυχολογική κατάσταση του χρήστη μέσα από την αναγνώριση μοτίβων και στον εντοπισμό των παρεμβάσεων εκείνων και τεχνικών που θα είναι πιο επιτυχημένες σε συγκεκριμένες καταστάσεις.³⁰⁵ Ένα καλό εργαλείο όμως, σημαίνει κάτι διαφορετικό από έναν καλό πράκτορα.³⁰⁶ Η δυνατότητά του chatbot να προσομοιώνει μια ανθρώπινη συνομιλία είναι μέρος μιας διαδικασίας εκτέλεσης ενός αλγορίθμου, αλλά δεν αποτελεί απόδειξη ύπαρξης της γνωστικής ικανότητας. Κατά τον John Searle κάτι τέτοιο δεν απαιτείται από ένα σύστημα της T.N. Το επιχείρημα που προβάλλει, γνωστό ως *Chinese Room Argument*, υποστηρίζει πως η επιτυχής σύνταξη φυσικών συμβόλων από το μηχάνημα δεν απαιτεί από το μηχάνημα να κατανοεί αυτά τα σύμβολα, αλλά απλώς να εφαρμόζει μια αντιστοίχιση εισόδου και εξόδου των δεδομένων, ανάλογη με αυτήν ενός ανθρώπου σε μια συγκεκριμένη εργασία. Συνεπώς, οι μηχανές δεν κατανοούν τις έννοιες και δεν

³⁰¹ Αυτόθι, 8.

³⁰² Αυτόθι, 8.

³⁰³ Αυτόθι.

³⁰⁴ Αυτόθι, 7.

³⁰⁵ Αυτόθι, 9.

³⁰⁶ Αυτόθι, 6.

είναι αυτονόητο ότι, ένα σύστημα που προσομοιώνει τη συμπεριφορά του ανθρώπου χρειάζεται να το κάνει.³⁰⁷

Θα ήταν αναληθές λοιπόν, το να ισχυριστούμε πως ένα chatbot είναι σε θέση να ανταποκριθεί στην πολυπλοκότητα της θεραπευτικής συνομιλίας, να υιοθετήσει τις κανονιστικές θέσεις ενός λογικού αυτουργού και να παρέχει αιτιολογήσεις, που θα βοηθήσουν τους χρήστες να κατανοήσουν καλύτερα τις μοναδικές ατομικές τους εμπειρίες. Συνεπώς, η διαλεκτική εκείνη διαδικασία της αμοιβαίας αναγνώρισης και του σεβασμού, που μπορεί να οδηγήσει σε αυξημένη αυτονομία και θεραπευτική αλλαγή, δεν αποτελούν χαρακτηριστικά που ένα chatbot κατέχει. Η διαφορά του λοιπόν, από τον ανθρώπινο επαγγελματία είναι προφανής. Το chatbot προσομοιώνει τη λογική ικανότητα, αλλά δεν τη διαθέτει. Ή αντίστοιχα, προσομοιώνει τη θεραπευτική συνομιλία, αλλά δεν μπορεί να τη διεξάγει ουσιαστικά.³⁰⁸

Αφήσαμε για το τέλος την έννοια του σεβασμού, καθώς ως μια σημαντική ηθική αξία, είναι αυτή που θα μας εισάγει στο πεδίο της ηθικής συμπεριφοράς που μπορούμε να αναμένουμε από ένα σύστημα της Τ.Ν. Είδαμε στο πρώτο κεφάλαιο το πόσο σημαντική είναι η επίδειξη σεβασμού από τον σύμβουλο, απέναντι στη μοναδικότητα και στην προσωπικότητα του ανθρώπου που στέκεται απέναντί του. Το ίδιο σημαντική είδαμε ότι είναι και η ύπαρξη ενός σταθερού συστήματος αξιών από πλευράς του και ενός κώδικα ηθικής που θα καθοδηγήσει τη δράση του ως σύμβουλος. Τίθεται λοιπόν εδώ το ερώτημα του εάν ένα αντίστοιχο σύστημα ηθικών αξιών και κανόνων θα ήταν εφικτό να ενσωματωθεί σε ένα σύστημα της Τ.Ν. και να διέπει την όποια λειτουργία του.

Στη δυνατότητα αυτή του προγραμματισμού ηθικών κανόνων στα συστήματα της Τ.Ν. εστιάζει το υπο-πεδίο έρευνας της Τ.Ν. με την ονομασία *Ηθική των Μηχανών* (*Machine Ethics*). Οι έρευνες στον επιστημονικό αυτό χώρο στοχεύουν στο να καταστούν τα συστήματα «ηθικώς δρώσες οντότητες» (*moral agents*), ικανές να ακολουθούν ηθικούς κανόνες, σε ένα φάσμα που θα περιλαμβάνει από απλούς κανονισμούς μέχρι και ολόκληρα ηθικά θεωρητικά συστήματα. Για τους ερευνητές του υπο-πεδίου της Ηθικής των Μηχανών, κάτι τέτοιο θεωρείται εφικτό για δύο λόγους. Πρώτον, γιατί τα συστήματα της Τ.Ν. είναι ικανά για την πολύ πιο ταχεία επεξεργασία μεγάλου όγκου δεδομένων σε σχέση με τον άνθρωπο και δεύτερον, γιατί δύνανται να τηρούν με απόλυτη σταθερότητα τους ηθικούς κανόνες, χωρίς κανέναν υποκειμενικό παράγοντα (π.χ. συναισθήματα) να επηρεάσει την κρίση και συμπεριφορά τους. Έτσι, όπως υποστηρίζουν οι ερευνητές της Ηθικής των Μηχανών, τα συστήματα της Τ.Ν. μπορούν να ενισχύσουν, να κατευθύνουν και να αναπληρώσουν την ηθική χρήση των ανθρώπων που τα χρησιμοποιούν.³⁰⁹

Θα πρέπει ωστόσο, να τονίσουμε τη διαφορά που υπάρχει ανάμεσα στο ηθικό δραν (*moral agency*) και το ηθικό καθεστώς (*moral status*). Στην πρώτη περίπτωση δεν απαιτείται από ένα σύστημα η κατανόηση των ηθικών κανόνων που θα ακολουθήσει. Αν και η δράση του μπορεί να έχει ηθικές συνέπειες, το να δρα απλά βάσει συγκεκριμένων προγραμματισμένων ηθικών κανόνων είναι αρκετό.³¹⁰ Κάτι τέτοιο

³⁰⁷ Alkis Gounaris, George Kosteletos, “Licensed to Kill: Autonomous Weapons as Persons and Moral Agents,” in *Personhood*, ed. Dragan Prole & Goran Rujević (Novi Sad: Filozofski Fakultet and Athens: The NKUA Applied Philosophy Research Lab Press, 2020): 155, https://www.researchgate.net/publication/348506490_Personhood.

³⁰⁸ *Αυτόθι*, 8-9.

³⁰⁹ Γούναρης, Κωστέλετος, *Τεχνητή Νοημοσύνη και Ηθική, Ενότητα 7^η* [πανεπιστημιακές σημειώσεις], ΕΚΠΑ; Anderson et al., “Towards moral machines,” 177-202.

³¹⁰ *Αυτόθι*.

είναι προβληματικό για το επιστημονικό πεδίο της Συμβουλευτικής. Αυτό, γιατί η απλή προσομοίωση της ηθικής συμπεριφοράς στη διάρκεια της συμβουλευτικής διαδικασίας, αν μη τι άλλο υποδηλώνει την απουσία κατανόησης των σημαντικών ηθικών αξιών και εννοιών που διαμορφώνουν τις θεραπευτικές σχέσεις.³¹¹

Η ανησυχία αυτή εντείνεται, αν λάβουμε υπόψη τον απώτερο στόχο της Ηθικής των Μηχανών, που είναι η επίτευξη ενός επιπέδου αυτονομίας των συστημάτων, στο οποίο τα ίδια, θα επιλέγουν ή θα διατυπώνουν νέους ηθικούς κανόνες που θα ακολουθούν. Αν όμως ένα σύστημα επιλέγει μόνο του τους ηθικούς κανόνες που θα τηρήσει, τότε υπάρχει ο κίνδυνος οι κανόνες αυτοί να μην ευθυγραμμίζονται με αυτούς των ανθρώπων. Ειδικότερα δε, αν συνυπολογίσουμε το πρόβλημα του μαύρου κουτιού (The Black Box Problem) ή της πιθανής μεροληψίας των συστημάτων (AI Bias), για τα οποία μιλήσαμε στην ενότητα 3.β., τότε η παροχή και η ακολούθηση συμβουλών με άγνωστο και ανεξέλεγκτο για τον άνθρωπο σκεπτικό, κάθε άλλο παρά ασφαλής είναι για την ψυχική ακεραιότητα του ανθρώπου.³¹²

Επιπλέον, το ηθικό καθεστώς (moral status) των συστημάτων της Τ.Ν., που αφορά στις ηθικές ευθύνες και τα δικαιώματά τους, παραμένει ασαφές. Προκύπτει λοιπόν το θέμα του καταλογισμού ηθικών ευθυνών στην περίπτωση που οι παρεχόμενες συμβουλές από ένα chatbot επιφέρουν βλάβη στον ανθρώπινο ψυχισμό.³¹³ Έχει νόημα στο σημείο αυτό, να κάνουμε μια αναφορά στον φιλόσοφο Robert Sparrow που ανέπτυξε ένα ισχυρό επιχείρημα κατά της ανάπτυξης στρατιωτικών συστημάτων Τ.Ν. Ο Sparrow παραλλήλισε τα συστήματα αυτά με τα ανήλικα παιδιά που επιστρατεύονται σε πολεμικές συρράξεις του τρίτου κόσμου. Όπως ανέφερε, τα παιδιά αυτά διαθέτουν δύο βασικά χαρακτηριστικά. Κατά πρώτον, δύναται να είναι αποτελεσματικά στη μάχη και επαρκώς αυτόνομα, ώστε δύσκολα να μπορούν να ελεγχθούν από τους ενήλικες επικεφαλής τους. Κατά δεύτερον, το ηθικό καθεστώς που φέρουν είναι ασαφές, ώστε να μπορούν να κατανοήσουν τις συνέπειες των πράξεων τους και να αναλάβουν την ευθύνη για αυτές. Τα δύο αυτά στοιχεία έχουν ως αποτέλεσμα την εκτέλεση εγκληματικών πράξεων, οι οποίες παραμένουν ατιμώρητες, καθώς ο καταλογισμός των ηθικών ευθυνών για αυτές δεν μπορεί να τεκμηριωθεί. Συνεπώς, η ευθύνη αφορά σε αυτούς που επιλέγουν να τοποθετήσουν τα παιδιά αυτά στο πεδίο της μάχης. Για τον Sparrow, τα ίδια αυτά χαρακτηριστικά φέρουν και τα στρατιωτικά συστήματα Τ.Ν. και για το λόγο αυτό, την ανάπτυξή τους θα πρέπει να αποτρέπει το ανεπίλυτο αυτό πρόβλημα του ηθικού καταλογισμού των ευθυνών.³¹⁴

Το θέμα αυτό μας μεταφέρει στο πεδίο της Ηθικής της Τ.Ν. (AI ethics), το οποίο και θα αναλύσουμε στο 4^ο κεφάλαιο της παρούσας διατριβής. Εκεί θα επιδιώξουμε να συσχετίσουμε τη σκέψη του Sparrow με τα συστήματα της Τ.Ν. που σταδιακά ενσωματώνονται στον τομέα της ψυχικής υγείας. Μεταξύ άλλων, θα διερευνήσουμε το κατά πόσο η χρήση των συστημάτων αυτών σε ευάλωτους πληθυσμούς και σε ένα αντικείμενο τόσο ευαίσθητο, όπως είναι το κομμάτι του ανθρώπινου ψυχισμού, είναι ηθικά αποδεκτή, τη στιγμή που το ηθικό καθεστώς των συστημάτων παραμένει απροσδιόριστο.

Για την ώρα, θα πρέπει να σταθούμε στο γεγονός πως τα υπάρχοντα συστήματα της Τ.Ν. στον τομέα της ψυχικής υγείας δεν είναι σε θέση να αναλάβουν τις ευθύνες των

³¹¹ Sedlakova, Trachsel, 'Conversational Artificial Intelligence in Psychotherapy,' 10.

³¹² Γούναρης, Κωστελέτος, *Τεχνητή Νοημοσύνη και Ηθική, Ενότητα 7^η* [πανεπιστημιακές σημειώσεις], ΕΚΠΑ.

³¹³ *Αυτόθι*.

³¹⁴ Κωστελέτος, *Τεχνητή Νοημοσύνη και Ηθική, Ενότητα 9^η* [πανεπιστημιακές σημειώσεις], ΕΚΠΑ.

λόγων που παρέχουν και τα καθήκοντα που αναλαμβάνει ένας άνθρωπος θεραπευτής. Ίσως αυτή η διαφορά να είναι και η πιο σημαντική. Η τεχνητή νοημοσύνη μπορεί να είναι κατάλληλη για εκπαιδευτικούς σκοπούς ή για τη χρήση συγκεκριμένων τεχνικών και δεξιοτήτων που βασίζονται στα ποσοτικά δεδομένα.³¹⁵ Όμως, δεν αποτελεί ηθικό υποκείμενο και έτσι, δεν δύναται να έχει τα καθήκοντα, τις ευθύνες και τις αρετές που διέπουν τις θεραπευτικές σχέσεις.

Κάτι τέτοιο θα πρέπει να γίνεται σαφές προς τους χρήστες. Διαφορετικά, η πεποίθηση πως ένα chatbot είναι ηθικό υποκείμενο, ενέχει τον κίνδυνο λανθασμένων πεποιθήσεων και προσδοκιών, ακόμη και τον κίνδυνο διατάραξης, αντί της διευκόλυνσης, της σχέσης μεταξύ συμβούλου και συμβουλευόμενου, κατά τη χρήση της ως εργαλείο.

Συχνά, τα ισχυρά ανθρώπινα χαρακτηριστικά των εφαρμογών μπορούν να δημιουργήσουν την ψευδαίσθηση ότι κάτι περισσότερο από μία συνομιλία με μία μηχανή μπορεί να συμβεί. Εάν όμως οι χρήστες πιστέψουν πως ένα chatbot είναι ικανό να τους φροντίσει ή να τους σεβαστεί, τότε αντίστοιχα θα αναμένουν τη δημιουργία μιας θεραπευτικής σχέσης, βασισμένη στην αλληλεπίδραση μεταξύ δύο ανθρώπων. Ο κίνδυνος αυτός της εξαπάτησης (ή αυταπάτης) των χρηστών, μπορεί να οδηγήσει σε μια σειρά αρνητικών επιπτώσεων, όπως η πρόκληση βλάβης της ψυχολογικής ακεραιότητας του χρήστη, η μείωση της αυτονομίας του, ακόμη και η εξάρτηση ή η τυφλή εμπιστοσύνη προς ένα chatbot, λόγω της 24/7 διαθεσιμότητας και της επιστημονικής του ανωτερότητας.³¹⁶ Κυρίως όμως θα επιφέρει τη ματαιώση, που θα επέλθει με τη συνειδητοποίηση πως αυτή η σχέση δεν είναι αληθινή. Μία καλή λύση που προτείνεται είναι οι συνεχείς υπενθυμίσεις προς τους χρήστες κατά την χρήση, ώστε να αποφεύγονται τέτοιου είδους παρανοήσεις.³¹⁷ Όπως προειδοποιεί το Woebot, «όσο έξυπνος κι αν φαίνομαι, δεν είμαι ικανός να κατανοήσω πραγματικά τι χρειάζεσαι».³¹⁸ Ωστόσο, μία υπενθύμιση στους ανθρώπους ίσως δεν αρκεί και δεν σημαίνει αυτομάτως ότι οι προσδοκίες τους για κάτι παραπάνω θα πάψουν να υπάρχουν. Χρειάζεται λοιπόν, κάτι περισσότερο, όπως η διασαφήνιση των στόχων και λειτουργιών της Τ.Ν. και η δημόσια συζήτηση και ευαισθητοποίηση του κοινού, γύρω από αυτή τη νέα τεχνολογία.³¹⁹

Για να δούμε στην πράξη τα όσα συζητήσαμε εδώ, στην ενότητα που ακολουθεί, θα επιχειρήσουμε την περιγραφή ενός υποθετικού περιστατικού που μπορεί να λάβει χώρα σε μία τηλεφωνική γραμμή βοήθειας και συμβουλευτικής υποστήριξης. Αρχικά θα δούμε το πως ένα τέτοιο περιστατικό αντιμετωπίζεται τη δεδομένη στιγμή από τους επαγγελματίες που στελεχώνουν μια τέτοια υπηρεσία. Στη συνέχεια, θα υποθέσουμε την εμπλοκή της Τ.Ν. και θα διερευνήσουμε το εάν στην πράξη, η εμπλοκή αυτή μπορεί να βελτιώσει ή όχι και πως, τους ήδη υπάρχοντες τρόπους δράσης.

3.ε. Περίπτωση μελέτης (case study)

Προτού ξεκινήσουμε, αξίζει να σημειωθεί πως το περιστατικό που ακολουθεί είναι υποθετικό. Ωστόσο, βάσει προσωπικής εργασιακής εμπειρίας σε μία από τις

³¹⁵ Sedlakova, Trachsel, "Conversational Artificial Intelligence in Psychotherapy," 10.

³¹⁶ *Αυτόθι*, 10-11.

³¹⁷ Sweeney et al., "Can chatbots help support a person's mental health?" 3.

³¹⁸ Kretzschmar et al., "Can Your Phone Be Your Therapist?" 3-4.

³¹⁹ Sedlakova, Trachsel, "Conversational Artificial Intelligence in Psychotherapy," 11.

τηλεφωνικές γραμμές βοήθειας και συμβουλευτικής υποστήριξης που υλοποιούνται στην Ελλάδα, είναι προσαρμοσμένο με τρόπο, ώστε να ανταποκρίνεται σε μεγάλο βαθμό στην πραγματικότητα. Τα υποθετικά χαρακτηριστικά που θα εξετάσουμε παρακάτω, ίσως συνυπάρχουν σε μία μόνο κλήση, ίσως και όχι. Ανήκουν όμως σίγουρα στο σύνολο εκείνο των χαρακτηριστικών των κλήσεων, που λαμβάνουν χώρα σε μια συνηθισμένη, θα λέγαμε και χωρίς εκπλήξεις, εργασιακή βάρδια.

Ας υποθέσουμε λοιπόν, πως σε μία γραμμή τηλεφωνικής υποστήριξης, ένα δεκάχρονο παιδί καλεί από κινητό smartphone και κλαίγοντας γοερά, αναφέρει πως χρειάζεται βοήθεια. Κατόπιν σχετικών ερωτήσεων, το παιδί αποκαλύπτει πως εδώ και αρκετό καιρό, βιώνει έντονες καταστάσεις εκφοβισμού και βίας στο χώρο του σχολείου, από ομάδα συνομήλικων του. Το τελευταίο περιστατικό σημειώθηκε περίπου μία ώρα πριν από την κλήση στη γραμμή και μετά από αυτό, το αγόρι εγκατέλειψε το χώρο του σχολείου, χωρίς να ενημερώσει κανέναν από τους συμμαθητές ή τους δασκάλους του. Μάλιστα, φεύγοντας τρομοκρατημένο, ακολούθησε άγνωστη διαδρομή και πλέον, έχει χάσει τον προσανατολισμό του, με αποτέλεσμα να μη γνωρίζει το πως θα επιστρέψει στο σχολείο, στην οικία του ή σε κάποιο άλλο γνώριμο σημείο. Το παιδί είναι ταραγμένο και φοβάται, τόσο λόγω του ότι είναι μόνο του σε ένα άγνωστο για αυτό μέρος, όσο και λόγω του ενδεχόμενου επίπληξης ή τιμωρίας από τους δασκάλους ή τους γονείς του.

Την κλήση αυτή, που διαδραματίζεται σε εργάσιμη μέρα και ώρα, καλείται να διαχειριστεί ένας επαγγελματίας κοινωνικός λειτουργός ή ψυχολόγος, που εκείνη τη στιγμή πραγματοποιεί τη βάρδιά του στη γραμμή. Αυτός, αφού προσπαθήσει να ηρεμήσει το παιδί, θα ακούσει όσα εκείνο έχει να του πει και με προτεραιότητά την ασφάλειά του, θα ζητήσει επιπλέον πληροφορίες (πόλη, συνοικία, τηλέφωνο), προκειμένου να το κατευθύνει. Λόγω της ανηλικότητας του παιδιού, της ψυχολογικής του κατάστασης, καθώς και του γεγονότος πως δεν βρίσκεται υπό την επίβλεψη ενήλικα, ο επαγγελματίας πιθανά θα ζητήσει από το παιδί να του αναφέρει κάποιο χαρακτηριστικό σημείο εντός του οπτικού του πεδίου (λ.χ. μια πινακίδα) και ίσως, το καθοδηγήσει να μιλήσει σε κάποιον άνθρωπο (λ.χ. τον ιδιοκτήτη κάποιου καταστήματος), προκειμένου να μπορέσει να εντοπίσει την ακριβή του τοποθεσία. Εάν αυτό είναι αδύνατο, τότε ο επαγγελματίας θα ενημερώσει την αστυνομία και τον αρμόδιο εισαγγελέα ανηλίκων, ώστε με εισαγγελική εντολή στον πάροχο κινητής τηλεφωνίας, να εντοπιστεί το στίγμα της κινητής συσκευής από την οποία καλεί το παιδί και στη συνέχεια, το ακριβές σημείο στο οποίο βρίσκεται. Μόλις ο εντοπισμός της τοποθεσίας ολοκληρωθεί, αστυνομικοί της Άμεσης Δράσης θα μεταβούν στο σημείο, θα παραλάβουν το παιδί και αφού αναζητήσουν τα στοιχεία της οικογένειάς του, θα το επιστρέψουν με ασφάλεια στην οικία του. Κατά τη διάρκεια της αναζήτησης αυτής και μέχρι την άφιξη της αστυνομίας, ο επαγγελματίας θα συνεχίσει να κρατά το παιδί στη γραμμή, προκειμένου να το καθησυχάζει, να του εξηγήει τη διαδικασία και να το στηρίζει ψυχολογικά, σε μια προσπάθεια να κερδίσει την εμπιστοσύνη του και να επιτύχει τη συνεργασία μαζί του. Εφόσον η παραλαβή του παιδιού διεξαχθεί ομαλά και ενημερωθεί από την αστυνομία, ο επαγγελματίας θα εκφράσει (μέσω αυτής) την περαιτέρω διαθεσιμότητά του προς το παιδί, την οικογένειά του ή και το σχολείο στο οποίο φοιτεί, για την παροχή συμβουλευτικής υποστήριξης γύρω από τον κατάλληλο τρόπο χειρισμού και επίλυσης των θεμάτων που εντόπισε κατά τη διαχείριση του περιστατικού (συμπεριφορές εκφοβισμού, ψυχολογική κατάσταση, φυγή).

Η παραπάνω περίπτωση ίσως μοιάζει σύνθετη, ωστόσο δυστυχώς, οι περιπτώσεις που φτάνουν σε μια γραμμή βοήθειας και συμβουλευτικής υποστήριξης, μπορεί να είναι πολύ πιο σοβαρές από αυτήν. Επιπλέον, η εξέλιξη του περιστατικού που υποτέθηκε, μπορεί να χαρακτηριστεί ως ιδιαίτερα ομαλή, καθώς άπειρες πιθανότητες θα

μπορούσαν να επιφέρουν ανατροπές και τελικώς, να στερήσουν ένα τέλος αίσιο, όπως το προαναφερθέν. Παρακάτω θα περιγράψουμε το ίδιο περιστατικό, υποθέτοντας πως αντί του επαγγελματία κοινωνικού λειτουργού ή ψυχολόγου, ένα σύστημα της Συνομιλητικής Τ.Ν. (Conversational AI) βρίσκεται στην άλλη άκρη της τηλεφωνικής γραμμής. Παράλληλα με την περιγραφή, θα επιδιώξουμε να διερευνήσουμε το εάν ένα chatbot θα μπορούσε να λειτουργήσει αποτελεσματικά, προληπτικά έναντι των πιθανών ανατροπών ή και το αντίθετο, δηλαδή να επιβαρύνει την όλη διαδικασία, επιφέροντας ανυπολόγιστους κινδύνους για την ασφάλεια ενός ανήλικου παιδιού.

Κατά την έναρξη της κλήσης λοιπόν, ο δεκάχρονος συνδέεται απευθείας με ένα εξελιγμένο chatbot, το οποίο και τον ρωτά πως θα μπορούσε να βοηθήσει. Στην αρχή της συνομιλίας, μέσω αυτοματοποιημένων μηνυμάτων, το chatbot συλλέγει τα απαραίτητα ατομικά στοιχεία του παιδιού (όνομα, ηλικία κ.α.) και καταγράφει τον τηλεφωνικό αριθμό της κλήσης. Στη συνέχεια και ενόσω το παιδί διηγείται τι του έχει συμβεί, το chatbot, ως σύστημα της Τ.Ν. επιστρατεύει τα εργαλεία εκείνα, που θα συμβάλλουν στην εξαγωγή περαιτέρω πληροφοριών. Για παράδειγμα, μέσω ψηφιακού φαινοτύπου, η ακριβής τοποθεσία του παιδιού δύναται να εντοπιστεί άμεσα (μέσω του GPS της συσκευής smartphone), ενώ το εργαλείο επεξεργασίας φυσικής γλώσσας (NLP), μπορεί να εξάγει συμπεράσματα γύρω από το τι ακριβώς αυτό αισθάνεται. Θα μπορούσαμε να σκεφτούμε παραπέρα, και να υποθέσουμε πως ένα smartwatch παρέχει δεδομένα γύρω από τα ζωτικά σημεία του παιδιού, που μπορούν με τη σειρά τους να δώσουν πληροφορίες, λ.χ. για το εάν το παιδί είναι στα πρόθυρα κρίσης πανικού, εάν έχει τραυματιστεί από την ομάδα των συνομήλικων που του επιτέθηκε και κινδυνεύει να χάσει τις αισθήσεις του (επομένως θα πρέπει να ενημερωθεί και το ΕΚΑΒ) κ.ο.κ. Επιπλέον, αν υποθέσουμε πως η τεχνολογία αναγνώρισης συναισθημάτων (Facial Emotion Technology) είναι ενσωματωμένη, τότε η αποστολή μιας στιγμιαίας φωτογραφίας από το παιδί, θα μπορούσε να προσφέρει πολύ περισσότερα, μέσω της ανάλυσης των μορφασμών του προσώπου, δίνοντας έτσι, ακόμη πιο ολοκληρωμένη και ακριβή εικόνα της ψυχολογικής κατάστασης στην οποία βρίσκεται.

Ας σταθούμε λοιπόν για λίγο στο σημείο αυτό, όπου ένα chatbot αποκρίνεται στην κλήση ενός παιδιού, αντί ενός επαγγελματία. Αρχικά, η απευθείας σύνδεση με την εφαρμογή και η δυνατότητά αυτής ως προς τον χειρισμό πολλών κλήσεων ταυτόχρονα, συνεπάγεται την καλύτερη διαχείριση του φόρτου εργασίας και την σχεδόν μηδενική αναμονή για το άτομο που καλεί. Επί του παρόντος, η μεγάλη αναμονή αποτελεί ένα ιδιαίτερα αρνητικό στοιχείο της λειτουργίας των γραμμών. Λόγω υποστελέχωσης, ένας μειωμένος αριθμός υπαλλήλων καλείται να διαχειριστεί έναν ιδιαίτερα μεγάλο όγκο κλήσεων. Ειδικότερα δε, στις μη εργάσιμες ημέρες και ώρες (π.χ. αργίες, νυχτερινές ώρες), συχνά οι γραμμές είναι ακόμη πιο υποστελεχωμένες, ενώ οι κλήσεις σε άλλα τμήματα του φορέα, πέραν της γραμμής, εκτρέπονται σε αυτήν με στόχο την εξυπηρέτηση των πολιτών. Επίσης, καθοριστική είναι και η παράμετρος της επικαιρότητας. Στις περιόδους εκείνες, όπου ένα περιστατικό ή μία είδηση λαμβάνει μεγάλη δημοσιότητα στα ΜΜΕ (λ.χ. θάνατος τριών παιδιών στην Πάτρα,³²⁰ υπόθεση Βαγγέλη Γιακουμάκη³²¹ κ.ο.κ.), εκατοντάδες κλήσεις αφορούν σε αυτό, με το κοινό να

³²⁰ «Πάτρα: Από τον πρώτο θάνατο στη σύλληψη της Ρούλας Πισπιρίγκου - Οι ημερομηνίες ορόσημα της υπόθεσης», CNN Greece, προσπελάστηκε 29 Σεπτεμβρίου 2023, <https://www.cnn.gr/ellada/story/306884/patra-apo-to-thanato-sti-syllipsi-tis-roylas-pispirigkoy-oi-imerominies-orosimo-tis-ypothesis>.

³²¹ «Βαγγέλης Γιακουμάκης: Το χρονικό της εξαφάνισης που συγκλόνησε όλη την Ελλάδα...», Μηχανή του Χρόνου, προσπελάστηκε 29 Σεπτεμβρίου 2023, <https://www.mixanitouxronou.com.cy/epikerotita/vangelis-giakoumakis-to-chroniko-tis-exafanisis-pou-sigklonise-oli-tin-ellada/>.

εκφράζει ανησυχία, να καταγγέλλει παρόμοια περιστατικά και να θέτει δεκάδες ερωτήματα στους ειδικούς γύρω από το πως θα προφυλαχθεί, θα αντιληφθεί, θα μιλήσει κ.ο.κ.³²² Το αποτέλεσμα είναι ακόμη περισσότερες κλήσεις για ελάχιστους εργαζομένους και συνεπώς, μεγαλύτερη αναμονή για το κοινό, πρόχειρη και μη ποιοτική διαχείριση των περιστατικών ή και μεγαλύτερη εργασιακή και ψυχολογική κόπωση για το προσωπικό της γραμμής. Έτσι, η δυνατότητα αυτή των chatbot για άμεση ανταπόκριση και καλύτερη διαχείριση των κλήσεων, θα λέγαμε πως αποτελεί σωτήρια λύση για τα προβλήματα που επιφέρει ο μεγάλος εργασιακός φόρτος.

Το ίδιο σημαντικό είναι και η δυνατότητα των εργαλείων της T.N., σαν αυτά που αναφέρθηκαν παραπάνω (ψηφιακός φαινότυπος, επεξεργασία φυσικής γλώσσας και τεχνολογία αναγνώρισης συναισθημάτων), να εξάγουν σε ελάχιστο χρόνο σημαντικά δεδομένα.³²³ Αυτό εξοικονομεί πολύτιμο χρόνο σε μία επείγουσα περίπτωση, παρέχει ακρίβεια (π.χ. ένα παιδί νηπιακής ηλικίας, ίσως δεν γνωρίζει την ονομασία της περιοχής στην οποία βρίσκεται και δεν ξέρει να διαβάσει τις πινακίδες γύρω του), καθώς και την εξασφάλιση του ότι όλα τα απαραίτητα στοιχεία θα συλλεχθούν. Για παράδειγμα, αν λόγω φόρτου εργασίας και εργασιακής κόπωσης σε μία δύσκολη βάρδια, ο επαγγελματίας δεν προλάβει ή παραβλέψει την καταγραφή του τηλεφωνικού αριθμού του ανήλικου και η κλήση τερματιστεί αιφνιδίως (π.χ. λόγω κακού σήματος), τότε η επανάκληση από τον ανήλικο και η επαναφορά της συνομιλίας μπορεί να καθυστερήσει σημαντικά λόγω της αναμονής. Εάν δε, η επανασύνδεση δεν σταθεί δυνατή (π.χ. λόγω αποφόρτισης της μπαταρίας της συσκευής του), τότε ο επαγγελματίας δεν θα διαθέτει τον τηλεφωνικό αριθμό του παιδιού και έτσι, δεν θα μπορέσει να κινητοποιήσει τις αρμόδιες αρχές για τον εντοπισμό του στίγματος της κινητής συσκευής. Συνεπώς, ένα ανήλικο παιδί θα παραμείνει εκτεθειμένο σε κίνδυνο και ο επαγγελματίας, αν και γνωρίζει την αναγκαιότητα της περίπτωσης, δεν θα έχει τη δυνατότητα καν να το προσεγγίσει, προκειμένου να το βοηθήσει. Λόγω της σοβαρότητας πολλών από τις περιπτώσεις που συναντούμε σε μία γραμμή βοήθειας και συμβουλευτικής υποστήριξης, κάτι τέτοιο μπορεί να αποβεί ακόμη και μοιραίο για ένα παιδί που βρίσκεται σε κίνδυνο και δυστυχώς, τέτοιου είδους σοβαρά περιστατικά δεν είναι λίγα.

Για την συνέχεια, μπορούμε να προχωρήσουμε εξετάζοντας δύο ενδεχόμενα εξέλιξης της κλήσης αυτής. Στην πρώτη περίπτωση, το chatbot λειτουργεί ως εργαλείο για τον επαγγελματία της τηλεφωνικής γραμμής. Έτσι, αφού συλλέξει τα απαραίτητα στοιχεία για το παιδί που κάλεσε και τις συνθήκες υπό τις οποίες πραγματοποίησε την κλήση αυτή, προβαίνει στη σύνδεση του παιδιού με τον επαγγελματία. Αυτός, μπορεί άμεσα να προχωρήσει στις απαραίτητες ενέργειες, χωρίς να χάνει πολύτιμο χρόνο στη συλλογή των απαραίτητων πληροφοριών και χωρίς την ανασφάλεια ότι οι πληροφορίες αυτές μπορεί να μην είναι ακριβείς. Η ανακατεύθυνση αυτή της κλήσης, θα πραγματοποιηθεί εφόσον το αίτημα που εκφράζεται από τον καλόντα δεν μπορεί να καλυφθεί μέσω αυτοματοποιημένων μηνυμάτων. Για παράδειγμα, αν ο καλών επιθυμεί απλά, να μάθει τις ώρες λειτουργίας της γραμμής και τις πληθυσμιακές ομάδες στις οποίες απευθύνεται, τότε το αίτημα αυτό μπορεί να καλυφθεί από το chatbot και έτσι,

³²² Tim Rens de Boer, Saskia Mirelle, Sandjai Bhulai, Renske Gilissen, Rob van der Mei, “Forecasting call and chat volumes at online helplines for mental health,” *BMC Public Health* 23, no. 984 (May, 2023): 6, <https://doi.org/10.1186/s12889-023-15887-2>; “More calls to men's helpline after Whispering Wall suspected murder-suicide that killed baby Kobi,” ABC NEWS, accessed November 17, 2023, <https://www.abc.net.au/news/2021-04-23/more-calls-to-mens-helpline-after-whispering-wall-murder-suicide/100090324>.

³²³ Βλ. σχετικώς: Montag, Sindermann, Baumeister, “Digital phenotyping,” 19-20; D’Alfonso, “AI in mental health,” 113; Bedi et al., “Automated analysis of free speech,” 1-7.

ο επαγγελματίας να μην επιβαρυνθεί με μία επιπλέον κλήση. Παρομοίως, αν στη λειτουργία του chatbot έχουν ενσωματωθεί εργαλεία για την ανίχνευση ψεύδους και διαπιστωθεί πως πρόκειται για φάρσα, τότε η κλήση θα τερματιστεί αυτόματα, καθιστώντας ακόμη πιο αποτελεσματική τη διαχείριση των κλήσεων.

Αξίζει να σημειωθεί πως οι περιπτώσεις της τηλεφωνικής φάρσας δεν είναι σπάνιες και γενικότερα, ένας σημαντικός αριθμός των κλήσεων δεν αφορά σε κοινωνικά περιστατικά. Λόγω του ότι οι κλήσεις είναι δωρεάν και ανώνυμες, πολλά άτομα μπορεί να καλούν για ερωτήσεις σχετικά με το φορέα που υλοποιεί τη γραμμή βοήθειας (τηλέφωνα, διευθύνσεις κ.λ.π.), μη σχετικές ερωτήσεις (π.χ. τι ώρα είναι), ακόμη και προκειμένου να παραπλανήσουν τον επαγγελματία με μία τηλεφωνική φάρσα, αδιαφορώντας (ή αγνοώντας, αν πρόκειται για παιδιά) για το ότι με τον τρόπο αυτό, στερούν από τους ανθρώπους που βρίσκονται σε άμεση ανάγκη την προτεραιότητα που απαιτείται. Συμπεραίνουμε λοιπόν, πως η διαμεσολάβηση ενός chatbot ανάμεσα σε έναν άνθρωπο που χρειάζεται βοήθεια και έναν εξειδικευμένο επαγγελματία είναι ιδιαίτερα πολύτιμη. Όπως φάνηκε και στη Γερμανία με το chatbot Conbot, το φιλτράρισμα των κλήσεων μπορεί να μειώσει σημαντικά το φόρτο εργασίας των εργαζομένων και με τον τρόπο αυτό να αυξήσει σημαντικά την απόδοσή τους, την ποιότητα της ενασχόλησής τους με το κάθε περιστατικό και τη γενικότερη αποτελεσματικότητα της γραμμής σε όλα τα επίπεδα.³²⁴

Αυτό ωστόσο, δεν αποκλείει σημαντικούς κινδύνους. Είδαμε στα προηγούμενα κεφάλαια τη δυσκολία των συστημάτων της Συνομιλητικής Τ.Ν. στο να κατανοήσουν επαρκώς τα αιτήματα που διατυπώνονται προφορικά. Την αδυναμία αυτή μπορεί να εντείνουν παράγοντες όπως, το γοερό κλάμα του παιδιού, η κακή άρθρωση, το πολύπλοκο λεξιλόγιο ή και η περιορισμένη ικανότητα ομιλίας.³²⁵

Είδαμε, βέβαια, πως στην περίπτωση του Conbot, τα αιτήματα αυτά που δεν έγιναν κατανοητά, ανακατευθύνθηκαν στους εργαζόμενους της γραμμής.³²⁶ Έτσι αναλόγως και στην περίπτωση μας, μία τέτοια αδυναμία θα μπορούσε να επιλυθεί με τον ίδιο τρόπο. Πώς θα μπορούσε, ωστόσο, να επιλυθεί το «Πρόβλημα του Πλαισίου» (The Frame Problem of AI)³²⁷ ή το φαινόμενο της υπερδιάγνωσης (overdiagnosis)³²⁸; Στην πρώτη περίπτωση, ο επαγγελματίας θα κατέληγε με έναν τεράστιο όγκο πληροφοριών, που θα καθιστούσαν δύσκολη την κατανόηση του ουσιαστικού νοήματος πίσω από τα λεγόμενα του παιδιού και του κυρίου αιτήματος που εκφράζεται. Επιπλέον, οι αναρίθμητες πληροφορίες θα καθιστούσαν ιδιαίτερα πολύπλοκο το χειρισμό μιας περίπτωσης που απαιτεί άμεσες ενέργειες. Στη δεύτερη περίπτωση, ο επαγγελματίας θα μπορούσε να λαμβάνει πληροφορίες για πιθανότητες, όπως η τάση για εμφάνιση παχυσαρκίας με βάση τον *Δείκτη Μάζας Σώματος (ΔΜΣ)*, που θα παρέχεται ως δεδομένο μέσα από το smartwatch του ανήλικου ή την ύπαρξη μιας μαθησιακής δυσκολίας του παιδιού (που θα εντοπίζεται μέσω επεξεργασίας φυσικής γλώσσας NLP). Οι πληροφορίες αυτές, όμως, είναι περιττές κατά τη στιγμή εκείνη που ο επαγγελματίας χρειάζεται να εκτελέσει έναν αριθμό απαραίτητων ενεργειών, ενώ και η αναφορά σε αυτές, στο πλαίσιο μιας ενδεχόμενης συνεδρίας με τους γονείς μετά την

³²⁴ Voelskow et al., “Prospective mixed-methods study,” 4.

³²⁵ Βλ. Palanica et al., “Physicians' Perceptions of Chatbots in Health Care,” 1-10; Kretzschmar et al., “Can Your Phone Be Your Therapist?,” 7; Voelskow et al., “Prospective mixed-methods study,” 9-11.

³²⁶ Voelskow et al., “Prospective mixed-methods study,” 4.

³²⁷ Βλέπε υπ.¹⁵ στην ενότητα 2.α. με τίτλο: «Τα οφέλη και τα επιστημολογικά - μεθοδικά προβλήματα βάσει των σημερινών δεδομένων», καθώς και “The Frame Problem.”

³²⁸ Βλ. ενότητα 3.δ. με τίτλο: «Οι δεξιότητες του συμβούλου κατά τη συμβουλευτική διαδικασία», καθώς και Capurro, Coghlan, Pires, “Preventing Digital Overdiagnosis,” 525.

ολοκλήρωση του περιστατικού, δεν θα επιφέρει τίποτα περισσότερο από την περαιτέρω ψυχολογική επιβάρυνση των ίδιων και του παιδιού τους.

Συν τοις άλλοις, τα πράγματα θα μπορούσαν να εξελιχθούν ακόμη χειρότερα, στην περίπτωση μεροληψίας του συστήματος (AI bias). Αν, για παράδειγμα, υποθέσουμε ότι βάσει των δεδομένων που συλλέγονται, το σύστημα της T.N. εξάγει το συμπέρασμα πώς το μεγαλύτερο ποσοστό των τηλεφωνικών φαρσών πραγματοποιείται από αγόρια ηλικίας εννέα με έντεκα ετών, τότε αντίστοιχα μία κλήση από ένα αγόρι αυτής της ηλικίας, θα ήταν πιθανό να αντιμετωπιστεί ως φάρσα, χωρίς να είναι, και να τερματιστεί αυτόματα. Ο τρόπος αυτός, με τον οποίο η μεροληψία ενός συστήματος θα μπορούσε να στερήσει σε ένα άτομο την πρόσβαση στις παροχές της γραμμής, έρχεται σε απόλυτη σύγκρουση με την αρχή της ισότιμης πρόσβασης³²⁹ που διέπει την παροχή της Πρωτοβάθμιας Φροντίδας Υγείας.³³⁰

Γενικότερα, το πρόβλημα του πλαισίου, η υπερδιάγνωση και η μεροληψία, ανήκουν στα προβλήματα εκείνα που οι προγραμματιστές των συστημάτων της T.N. εστιάζουν, με σκοπό την αντιμετώπισή τους. Όσο, όμως, τα προβλήματα αυτά είναι υπαρκτά, κατανοούμε πώς η ένταξη των συστημάτων της T.N. σε μία τηλεφωνική γραμμή βοήθειας και συμβουλευτικής υποστήριξης, απαιτεί ιδιαίτερη σκέψη προτού αποφασιστεί. Μπορεί το Conbot να επέφερε σημαντική ευκολία στις τηλεφωνικές γραμμές του υγειονομικού συστήματος της Γερμανίας, ωστόσο στην περίπτωση αυτή, το αντικείμενο της ενασχόλησης ήταν πολύ συγκεκριμένο και έτσι, το ρίσκο απέναντι στην ψυχολογική κατάσταση ή και την ίδια τη ζωή των καλούντων ήταν περιορισμένο. Αντίθετα, στην περίπτωση μιας τηλεφωνικής γραμμής βοήθειας και συμβουλευτικής υποστήριξης, όπως αυτές που υλοποιούνται στην Ελλάδα, τα περιστατικά που μπορούν να φτάσουν είναι ποικίλα και συνήθως, καλύπτουν ένα ευρύ φάσμα των ανησυχιών μιας μεγάλης πληθυσμιακής ομάδας (λ.χ. έφηβοι, παιδιά, γονείς). Παραδείγματος χάριν, ένας γονέας μπορεί να καλέσει σε μία συμβουλευτική γραμμή γονέων, προκειμένου να λάβει συμβουλές για το πώς θα μιλήσει στο παιδί του γύρω από την ομοφυλοφιλία, ένα θέμα γενικό που δεν κατατάσσεται στις επείγουσες περιπτώσεις. Μπορεί, όμως, να αιτηθεί συμβουλευτική υποστήριξη μετά από μία απόπειρα αυτοκτονίας που το ίδιο παιδί έχει κάνει. Αναμφισβήτητα, η δεύτερη περίπτωση είναι εντελώς διαφορετική και η αλήθεια είναι πως, ο βαθμός σοβαρότητας μιας κλήσης δεν δύναται να είναι γνωστός ή να προβλεφθεί, προτού αυτή απαντηθεί.

Τα πράγματα γίνονται ακόμη πιο πολύπλοκα αν υποθέσουμε πως το chatbot δεν διαμεσολαβεί ανάμεσα στον καλούντα και στον επαγγελματία αλλά αντίθετα, αναλαμβάνει το ρόλο του συμβούλου και χειρίζεται εξολοκλήρου μία κλήση στη γραμμή.³³¹ Στην δεύτερη αυτή περίπτωση, τα παραπάνω πλεονεκτήματα που φέρει η συμμετοχή του chatbot παραμένουν τα ίδια. Το ίδιο και οι ανησυχίες που

³²⁹ Για την Ελλάδα, η αρχή της ισότιμης πρόσβασης στις μονάδες της Πρωτοβάθμιας Φροντίδας Υγείας κατοχυρώνεται μέσα από το Άρθρο 1 παρ. 3 (*Γενικές Αρχές Πρωτοβάθμιας Φροντίδας Υγείας*) του νόμου υπ' αριθμ. 4486 (ΦΕΚ Α' 115/07.08.2017). Βλ. «ΝΟΜΟΣ ΥΠ' ΑΡΙΘΜ. 4486 ΦΕΚ Α' 115/07.08.2017», [Kodiko.gr](https://www.kodiko.gr), προσπελάστηκε στις 14 Νοεμβρίου 2023, <https://www.kodiko.gr/nomothesia/document/274963/nomos-4486-2017>. Επιπλέον, σε ευρωπαϊκό επίπεδο, η αρχή επιβεβαιώνεται μέσα από τη *Σύμβαση για τα ανθρώπινα δικαιώματα και την Βιοϊατρική* (Σύμβαση του Οβιέδο, Article 3–Equitable access to health care). Βλ. “Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine,” Council of Europe, accessed November 17, 2023, <https://rm.coe.int/168007cf98>.

³³⁰ Την ηθική εδραίωση αυτής της αρχής, θα δούμε στο επόμενο κεφάλαιο, όπου θα αναλύσουμε τις τέσσερις αρχές της Βιοϊατρικής Ηθικής, μια από τις οποίες είναι η αρχή της Δικαιοσύνης.

³³¹ Μεταβαίνει δηλ. σε μία κατάσταση Human off the loop (βλ. Ενότητα 2.β., *υπ.*¹⁵⁰), ενός επιπέδου δηλ. πλήρους αυτονομίας που εγείρει τελικά, για κάθε εφαρμογή της T.N., τα πιο σημαντικά ερωτήματα.

εξακολουθούν να υπάρχουν. Στην περίπτωση αυτή, όμως, προστίθενται κι άλλες, ακόμη πιο σοβαρές.

Ίσως σε πρακτικό επίπεδο ένα σύστημα της Συνομιλητικής Τ.Ν. να μπορεί να βοηθήσει ουσιαστικά στη λειτουργία μιας γραμμής. Σε ποιοτικό επίπεδο, ωστόσο, δεν δύναται να φτάσει την ουσιαστική επαφή που χαρακτηρίζει μία γνήσια ανθρώπινη θεραπευτική σχέση. Όπως είδαμε έως τώρα, μη κατέχοντας την ιδιότητα του υποκειμένου, το chatbot δεν έχει την ικανότητα να κατανοήσει ζωτικές έννοιες.³³² Στην περίπτωσή μας επομένως, μπορεί να παράγει χιλιάδες δεδομένα γύρω από την έννοια *φόβος* ή τον όρο *σχολικός εκφοβισμός*. Δεν θα κατανοήσει, όμως, την ουσία των εννοιών αυτών και δεν θα μπορέσει να αντιληφθεί μια σειρά σκέψεων και συναισθημάτων με τις οποίες συνδέονται, ώστε να επιδείξει ουσιαστική κατανόηση, ενσυναίσθηση και να αναπτύξει μια γνήσια θεραπευτική αλληλεπίδραση κατά τη συνομιλία του με το παιδί. Έτσι, αν και μπορεί να προβεί σε δηλώσεις όπως «καταλαβαίνω πως είναι δύσκολο αυτό που περνάς», οι δηλώσεις αυτές δεν παύουν να είναι μέρος μιας προσομοίωσης της ανθρώπινης συμπεριφοράς από ένα σύστημα που δείχνει νοήμον, χωρίς να είναι πραγματικά.³³³

Αυτή άλλωστε, όπως είδαμε, είναι και η βασική διαφορά ενός συστήματος της Τ.Ν. από έναν άνθρωπο και εξ αυτής προκύπτουν αφενός, ο κίνδυνος ψυχικής βλάβης του παιδιού λόγω της ματαιώσης που θα νιώσει όταν αντιληφθεί πως η σχέση που έχει αναπτύξει με το chatbot δεν είναι αληθινή και αφετέρου, κίνδυνοι για την ασφάλεια του, λόγω πιθανοτήτων εσφαλμένης κατανόησης της κατάστασης που βιώνει και λανθασμένων κατευθύνσεων που μπορούν να του δοθούν. Για παράδειγμα, εάν το chatbot αντιμετωπίσει τη συμπεριφορά της φυγής από τον ανήλικο μεμονωμένα και ξέχωρα από τις καταστάσεις σχολικής βίας και εκφοβισμού που βιώνει, τότε, δεν θα κατανοήσει ποτέ τη σύνδεση αυτών των δύο παραμέτρων και την ιδιότητα της πρώτης ως αντίδραση στη δεύτερη. Με τον ίδιο τρόπο, το chatbot δεν θα μπορέσει αναλόγως να κατευθύνει σωστά τους γονείς του παιδιού ή και το σχολείο. Έτσι, ο από μέρους τους λανθασμένος χειρισμός της συμπεριφοράς του παιδιού, θα μπορούσε να του επιφέρει ένα αίσθημα ανασφάλειας που ίσως το ωθήσει σε παρόμοιες συμπεριφορές ή και επιπλέον συναισθηματική πίεση, που θα μπορούσε να οδηγήσει ακόμη και σε εκφράσεις αυτοκτονικού ιδεασμού.

Στο σημείο αυτό, προκύπτει το θέμα του καταλογισμού των ηθικών ευθυνών. Όπως είδαμε στην ενότητα που προηγήθηκε, το ηθικό καθεστώς ενός συστήματος³³⁴ παραμένει μέχρι και σήμερα ασαφές.³³⁵ Μπορεί ο προγραμματισμός των συστημάτων της Τ.Ν., ώστε να ακολουθούν με συνέπεια κανόνες ηθικής συμπεριφοράς να είναι εφικτός. Μη έχοντας όμως την ιδιότητα του ηθικού υποκειμένου, δεν είναι, επί του παρόντος, σε θέση να αναλάβουν τις ευθύνες και τα καθήκοντα ενός ανθρώπινου θεραπευτή ή να κατανοήσουν τις ουσιαστικές εκείνες ηθικές αξίες που διέπουν τις θεραπευτικές σχέσεις.³³⁶

³³² Βλ. ενότητα 3.δ. με τίτλο: «Οι δεξιότητες του συμβούλου κατά τη συμβουλευτική διαδικασία» και Sedlakova, Trachsel, "Conversational Artificial Intelligence in Psychotherapy," 7-8.

³³³ Έχουμε δει στην ενότητα 3.δ. της εργασίας αυτής, το επιχείρημα του John Searle, γνωστό ως "Chinese Room Argument," σύμφωνα με το οποίο, η προσομοίωση της ανθρώπινης συμπεριφοράς μέσα από την επιτυχή σύνταξη φυσικών συμβόλων, δεν απαιτεί την κατανόηση των συμβόλων αυτών. Βλ. Gounaris, Kosteletos, "Licensed to Kill," 155.

³³⁴ Για μια γενικότερη συζήτηση περί του προβλήματος διασάφησης του ηθικού καθεστώτος των οντοτήτων Τ.Ν., βλ. Gounaris, Kosteletos, "Licensed to Kill," 137-189.

³³⁵ Γούναρης, Κωστελέτος, *Τεχνητή Νοημοσύνη και Ηθική, Ενότητα 7^η* [πανεπιστημιακές σημειώσεις], ΕΚΠΑ.

³³⁶ Sedlakova, Trachsel, "Conversational Artificial Intelligence in Psychotherapy," 10.

Ολοκληρώνοντας, δεν θα μπορούσαμε να παραβλέψουμε την έλλειψη της δυνατότητας ενός συστήματος της Τ.Ν. να προσφέρει λόγους και νέους τρόπους σκέψης, που θα ενισχύσουν την αυτονομία του ατόμου και θα το ωθήσουν στην επιθυμητή αλλαγή. Αυτοί οι νέοι τρόποι σκέψης, στην περίπτωση του δεκάχρονου, θα μπορούσαν να είναι πολύ σημαντικοί. Για παράδειγμα, στο άκουσμα των διαδικασιών που εξελίσσονται από τον επαγγελματία (π.χ. μια κλήση στην Άμεση Δράση, η ενημέρωση εισαγγελίας), ένα δεκάχρονο παιδί μπορεί να πανικοβληθεί και να τερματίσει την κλήση ή να φύγει και πάλι, με τον κίνδυνο η αντίδραση της φυγής κάθε φορά που αισθάνεται φόβο, να εγκατασταθεί στη συμπεριφορά του. Είναι σημαντικό λοιπόν, να κατανοήσει το γιατί απαιτείται μια τέτοια κινητοποίηση και να αισθανθεί την απαραίτητη ασφάλεια. Μόνο έτσι, θα συνεργαστεί, αρχικά με τον επαγγελματία και τις αρμόδιες αρχές και μετέπειτα με επαγγελματίες συμβούλους που θα συντελέσουν στην αναπλαισίωση των εμπειριών του, δηλαδή στην αλλαγή οπτικής προς τα γεγονότα που βιώνει και στην εύρεση νέων τρόπων διαχείρισης και αντιμετώπισής τους. Αυτό ακριβώς συνιστά, τη θεραπευτική αλλαγή, που αποτελεί το βασικό στόχο της θεραπευτικής διαδικασίας και απουσία της δυνατότητας αυτής, φαίνεται πως για την ώρα, η αλλαγή αυτή, μέσω της αλληλεπίδρασης με chatbot, δεν είναι εφικτή.

Διαπιστώνουμε λοιπόν, με βάση τα παραπάνω, πως τόσο τα θεωρητικά δεδομένα που αναλύσαμε στις προηγούμενες ενότητες, όσο και η υποθετική αναπαράστασή τους στην πράξη εγείρουν σοβαρές ανησυχίες και προβληματισμούς γύρω από αδυναμίες, που εάν δεν αντιμετωπιστούν, θα μπορούσαν να θέσουν σε σημαντικό κίνδυνο την ασφάλεια ενός ατόμου και την ψυχολογική του ακεραιότητα. Όπως αναφέραμε και νωρίτερα, σε μία γραμμή τηλεφωνικής βοήθειας και συμβουλευτικής υποστήριξης, τα περιστατικά τα οποία καλείται να διαχειριστεί ένας επαγγελματίας μπορεί να είναι πολύ πιο σοβαρά από αυτό που περιγράψαμε στην παρούσα ενότητα. Σίγουρα η διαθεσιμότητα των συστημάτων της Τ.Ν. δεν είναι άνευ αξίας και πράγματι, η χρήση τους ως εργαλείο στα χέρια ενός επαγγελματία θα μπορούσε να αποφέρει σημαντικά οφέλη και να βελτιώσει το επίπεδο ποιότητας και λειτουργικότητας των γραμμών. Ακόμη και αυτή η χρήση όμως, απαιτεί περαιτέρω βελτίωση των συστημάτων, ώστε να επιτευχθεί η ασφάλεια όσων αναζητούν βοήθεια με αυτόν τον τρόπο και η βεβαιότητα πως στη διαχείριση ενός σοβαρού περιστατικού, το chatbot θα είναι αποτελεσματικό, τουλάχιστον στο βαθμό εκείνο, που είναι μέχρι σήμερα ένας ανθρώπινος επαγγελματίας. Από την άλλη, η εξ ολοκλήρου αντικατάσταση ενός επαγγελματία από ένα εξελιγμένο σύστημα της Συνομιλητικής Τ.Ν., για την ώρα φαίνεται να μην είναι εφικτή. Σίγουρα, οι συνεχείς εξελίξεις που σημειώνονται στον τομέα της Τ.Ν. δεν αποκλείεται να ανατρέψουν την άποψη αυτή. Τη δεδομένη στιγμή ωστόσο, τα εμπόδια που θα πρέπει να ξεπεραστούν είναι ιδιαίτερα σημαντικά, θέτοντας σε σοβαρούς κινδύνους τόσο ανθρώπους με ευάλωτο ψυχισμό, όσο και την ακεραιότητα των συστημάτων παροχής Πρωτοβάθμιας Φροντίδας Υγείας.

Κεφάλαιο 4°: Η εξέταση μέσω ηθικής

4.α. Οι τέσσερις αρχές της Βιοηθικής.

Η ψυχική υγεία αποτελεί έναν τομέα που, από τη φύση του, εγείρει ιδιαίτερα ηθικά ζητήματα γύρω τη φροντίδα και το χειρισμό καταστάσεων που σχετίζονται με τον ευάλωτο ανθρώπινο ψυχισμό. Η αυξανόμενη εφαρμογή της Τ.Ν. στον τομέα αυτό, έρχεται να δημιουργήσει τις δικές της ηθικές προκλήσεις και ανησυχίες και επαναφέροντας στο προσκήνιο έννοιες όπως η δικαιοσύνη, η διαφάνεια, το απόρρητο και η ασφάλεια, να επισημάνει την ανάγκη εξέτασης και ρύθμισης της εφαρμογής αυτής.³³⁷

Τα ηθικά ζητήματα που προκύπτουν, θα εξετάσουμε στην ενότητα αυτή, υπό το πρίσμα των τεσσάρων αρχών της Βιοηθικής, όπως αναλύθηκαν από τους Αμερικανούς φιλοσόφους James Childress και Thomas Beauchamp.³³⁸ Νωρίτερα όμως, θα ήταν χρήσιμη μια αναφορά στο ίδιο το πεδίο της Βιοηθικής, το στενά συνδεδεμένο τόσο με την ηθική φιλοσοφία και πρακτική όσο και με τα ανθρώπινα δικαιώματα.³³⁹

Η Βιοηθική λοιπόν, ως όρος πρωτοεμφανίστηκε πριν τριάντα χρόνια και σήμερα, αποτελεί κλάδο της εφαρμοσμένης φιλοσοφίας.³⁴⁰ Ένας από τους ορισμούς που την περιγράφει είναι αυτός του Βέλγου καθηγητή Φιλοσοφίας Gilbert Hottois, που όρισε ως Βιοηθική το «σύνολο ερευνών, λόγων και πρακτικών, γενικά διεπιστημονικών, που έχει ως αντικείμενο να αποσαφηνίσει, ή και να λύσει ζητήματα που έχουν ηθική σπουδαιότητα και προκύπτουν από την πρόοδο και την εφαρμογή των βιοϊατρικών τεχνοεπιστημών».³⁴¹ Στην ουσία, λαμβάνοντας υπόψη τις θετικές πλευρές της εξέλιξης και τις επιπτώσεις της στον άνθρωπο, η Βιοηθική καλείται να «υποδείξει τις ασφαλιστικές δικλίδες που θα διασφαλίσουν τον σεβασμό της ανθρώπινης αξιοπρέπειας, αυτονομίας και αξιοκρατικής συμβίωσης».³⁴²

Στην επιδίωξη της ισορροπίας ανάμεσα στα οφέλη και τις ζημιές, η Βιοηθική βασίζεται σε τέσσερις αρχές, όπως αυτές διατυπώθηκαν από τους Beauchamp και Childress³⁴³ και οι οποίες, ισχύουν ανεξάρτητα από τις ιδεολογίες και τους πολιτισμούς.³⁴⁴ Οι αρχές αυτές είναι:

α. Η αρχή της αυτονομίας (*The principle of autonomy*), που σχετίζεται με την ικανότητα του ατόμου να σκεφτεί ένα σχέδιο δράσης για τον εαυτό του και να το

³³⁷ D'Alfonso, "AI in mental health," 114.

³³⁸ "Theory and Bioethics," Stanford Encyclopedia of Philosophy, accessed November 2, 2023, <https://plato.stanford.edu/entries/theory-bioethics>.

³³⁹ Μυρτώ Δραγώνα-Μονάχου, «Ηθική και βιοηθική», *Επιστήμη και Κοινωνία: Επιθεώρηση Πολιτικής και Ηθικής Θεωρίας* 8 (2015): 1, <https://doi.org/10.12681/sas.715>.

³⁴⁰ *Αυτόθι*, 6.

³⁴¹ *Αυτόθι*, 10.

³⁴² *Αυτόθι*, 13.

³⁴³ Το σχετικό έργο: Tom L. Beauchamp and James F. Childress, *Principles of Biomedical Ethics*, 8th Edition (England: Oxford University Press, 2019).

³⁴⁴ Θεοδώρα Πλακιά, «Η μείωση της βλάβης στη δημόσια υγεία. Βασικές Αρχές Βιοηθικής» (Διπλωματική διατριβή, Πανεπιστήμιο Δυτικής Αττικής, 2022): 34, <https://scholar.archive.org/work/5d3ckdxecbcwvouiimp5ztyfjy>.

εφαρμόσει, χωρίς τον έλεγχο των άλλων και με επαρκές επίπεδο κατανόησης, ώστε να παρέχεται η δυνατότητα για νοσηματοδότηση της επιλογής.³⁴⁵ Έτσι, ως αρχή, η δικαιοσύνη είναι συνυφασμένη με τις έννοιες της συναίνεσης, της ικανότητας, της επιθυμίας και των πεποιθήσεων.³⁴⁶

β. Η αρχή της αγαθοποιίας (*The principle of beneficence*), που αφορά στην συμβολή της ανθρώπινης ευημερίας και κατά τους Beauchamp και Childress, περιλαμβάνει τόσο την θετική αγαθοποιία (*positive beneficence*), δηλαδή την παροχή οφελών για τον ασθενή, όσο και τη χρησιμοθηρική αγαθοποιία (*utility*), η οποία απαιτεί την, από μέρους του επαγγελματία, ανάλυση των οφελών έναντι των κινδύνων, ώστε να παράγει το καλύτερο δυνατό αποτέλεσμα.³⁴⁷ Δεδομένου ότι ο ορισμός του οφέλους μπορεί να είναι διαφορετικός για τον καθένα, η αρχή της αγαθοποιίας είναι στενά συνδεδεμένη με το σεβασμό της ανθρώπινης αυτονομίας και την έννοια της συναίνεσης.³⁴⁸

γ. Η αρχή της μη βλάβης (*The principle of nonmaleficence*), δηλαδή της αποφυγής πρόκλησης ζημίας στους ασθενείς.³⁴⁹ Η προτεραιότητα αυτή, της αποφυγής πρόκλησης βλάβης στην ιατρική ενσαρκώνεται στη λατινική φράση "*primum non nocere*" («πρωτίστως, μη βλάψεις»), μια θεμελιώδης ηθική αρχή στην ιατρική.³⁵⁰ Για πολλούς ερευνητές οι αρχές της αγαθοποιίας και της μη βλάβης αποτελούν τις δύο όψεις του ίδιου νομίσματος. Αυτό γιατί, στην ανάλυση κινδύνου-οφέλους, η επίτευξη του οφέλους απαιτεί παράλληλα την αποφυγή του κινδύνου της βλάβης.³⁵¹

δ. Η αρχή της δικαιοσύνης (*The principle of justice*), που θέτει το θεμελιώδες ερώτημα του εάν το δικαίωμα στην υγειονομική περίθαλψη είναι καθολικό. Αφορά στην ισότιμη κατανομή των παροχών, στην εξασφάλιση του δικαιώματος πρόσβασης και γενικότερα, στην ισότητα και στον σεβασμό των νόμων.³⁵²

Στον πλαίσιο της εφαρμογής της T.N. στον τομέα της ψυχικής υγείας συγκεκριμένα, οι παραπάνω αρχές εκδηλώνονται μέσα από διαδικασίες όπως η συγκατάθεση, η προστασία του απορρήτου, η αυτόνομη λήψη αποφάσεων κ.ο.κ., που θα πρέπει να λαμβάνονται υπόψη και να εφαρμόζονται σε κάθε υλοποίηση. Στο πλαίσιο αυτής της ηθικής διακυβέρνησης, ο τελικός στόχος δεν είναι η παύση της τεχνολογικής εξέλιξης, αλλά η εξασφάλιση της προστασίας των συμφερόντων όλων των υποκειμένων, παράλληλα με την αξιοποίηση των δυνατοτήτων που η T.N. μας παρέχει.³⁵³

³⁴⁵ Dana J. Lawrence, "The Four Principles of Biomedical Ethics: A Foundation for Current Bioethical Debate," *Journal of Chiropractic Humanities* 14 (December, 2007): 35, [https://doi.org/10.1016/S1556-3499\(13\)60161-8](https://doi.org/10.1016/S1556-3499(13)60161-8).

³⁴⁶ Πλακιά, «Η μείωση της βλάβης στη δημόσια υγεία», 39.

³⁴⁷ Lawrence, "The Four Principles of Biomedical Ethics," 35.

³⁴⁸ *Αυτόθι*, 36.

³⁴⁹ Πλακιά, «Η μείωση της βλάβης στη δημόσια υγεία», 39.

³⁵⁰ Raanan Gillon, "Primum non nocere and the principle of non-maleficence," *British Medical Journal (Clinical research ed.)* 13, no. 291 (6488) (July, 1985): 130, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1416256/pdf/bmjcred00456-0042.pdf>.

³⁵¹ Lawrence, "The Four Principles of Biomedical Ethics," 36.

³⁵² *Αυτόθι*, 36.

³⁵³ Jian Guan, "Artificial Intelligence in Healthcare and Medicine: Promises, Ethical Challenges and Governance," *Chinese Medical Sciences Journal* 34, no. 2 (June 2019): 76-83, https://www.researchgate.net/publication/334574521_Artificial_Intelligence_in_Healthcare_and_Medicine_Promises_Ethical_Challenges_and_Governance/citations.

Οι παραπάνω αρχές είναι πάντοτε έγκυρες και δεσμευτικές, χωρίς καμία να έχει προτεραιότητα έναντι των υπολοίπων.³⁵⁴ Μάλιστα συχνά, τις βλέπουμε να συνυπάρχουν σε μία κατάσταση και για το λόγο αυτό, δεν θα τις δούμε ξεχωριστά, αλλά σε ένα σύνολο συνθηκών που μπορεί να τις πληρούν ή και να τις διακυβεύουν. Μια κατηγοριοποίηση που θα μπορούσε να μας βοηθήσει εδώ, είναι αυτή που κάνουν στο άρθρο τους οι Kretzschmar et al.³⁵⁵ Όπως αναφέρουν, η ηθικά υπεύθυνη εφαρμογή της Τ.Ν. στην υποστήριξη της ψυχικής υγείας θα πρέπει να πληροί πρωτίστως το σεβασμό στην ιδιωτικότητα του χρήστη, την εμπειρική στοιχειοθέτηση της αποτελεσματικότητας των διαφορετικών εργαλείων και εφαρμογών και κυρίως, την προστασία της ασφάλειας του. Με βάση αυτή την κατηγοριοποίηση λοιπόν, θα προχωρήσουμε στη συνέχεια.

4.α.1. Απόρρητο και ιδιωτικότητα

Σε μία δημοκρατική κοινωνία, το δικαίωμα στον αυτοκαθορισμό και στην ανάπτυξη της προσωπικότητας είναι άρρηκτα συνδεδεμένα με το απαραβίαστο της ιδιωτικής ζωής του ατόμου και της πληροφοριακής του αυτοδιάθεσης.³⁵⁶ Σε αυτά ακριβώς συνίσταται το δικαίωμα στην προστασία των προσωπικών δεδομένων, που νοείται ως «το δικαίωμα κάθε ανθρώπου να μη μετατρέπεται από κανέναν άλλο σε πληροφοριακό αντικείμενο και να συμμετέχει ο ίδιος ενεργά στον καθορισμό εκείνων των πληροφοριών που τον αφορούν και επιθυμεί να καταστούν γνωστές στο περιβάλλον του».³⁵⁷

Σε μια εποχή όπου η συλλογή και επεξεργασία των προσωπικών δεδομένων από τα συστήματα της Τ.Ν. είναι αλόγιστη και παράλληλα, πολλές εταιρείες κερδίζουν τα χρήματά τους μέσω της εκμετάλλευσης των δεδομένων αυτών, τα ζητήματα που εγείρονται γύρω από τη συμβατότητα της τεχνολογίας με το δικαίωμα προστασίας προσωπικών δεδομένων και πληροφοριακού αυτοκαθορισμού είναι έντονα. Ο μεγάλος όγκος των δεδομένων αυτών και η δυσκολία να τεθεί υπό τον έλεγχο του υποκειμένου που αφορούν, θίγει το δικαίωμα της αυτονομίας και της αυτοδιάθεσης του, ενώ παράλληλα, δημιουργεί σοβαρές ανησυχίες για την ιδιωτική του ζωή.³⁵⁸ Εάν μάλιστα, λάβουμε υπόψη το ευαίσθητο περιεχόμενο των πληροφοριών που διαμοιράζονται στις εφαρμογές της ψυχικής υγείας (chatbot), και το οποίο σχετίζεται με τα συναισθήματα, τις σκέψεις και τις πεποιθήσεις του ατόμου, τότε κατανοούμε τον τρόπο με τον οποίο διακυβεύεται η κάθε μορφή της ιδιωτικότητας, όπως η συναισθηματική και η νοητική.³⁵⁹

³⁵⁴ Lawrence, “The Four Principles of Biomedical Ethics,” 36.

³⁵⁵ Kretzschmar et al., “Can Your Phone Be Your Therapist?” 7.

³⁵⁶ Σταυρούλα Γεωργίου, «Αυτόματη αναγνώριση προσώπου στους δημόσιους χώρους: Μεταξύ πραγματικότητας και δυστοπίας», Syntagmawatch.gr, προσπελάστηκε στις 4 Νοεμβρίου 2023, <https://www.syntagmawatch.gr/trending-issues/aytomati-anagnwrisi-proswpou-stous-dhmosious-xwrous-metaxy-pragmatikothtas-kai-dystopias/>

³⁵⁷ Φερενίκη Παναγοπούλου-Κουτνατζή, *Τεχνητή Νοημοσύνη: Ο δρόμος προς έναν ψηφιακό συνταγματισμό. Μια ηθικο-συνταγματική θεώρηση*. (Αθήνα: Εκδόσεις Παπαζήση, 2023), 85.

³⁵⁸ *Αυτόθι*, 86.

³⁵⁹ Βλ. Gary Marchant, Lucille Nalbach Tournas, “Matter Over Mind: Liability Considerations Surrounding Artificial Intelligence in Neuroscience,” in *Artificial Intelligence in Brain and Mental Health: Philosophical, Ethical & Policy Issues*, ed. by Fabrice Jotterand, Marcello Ienca (Switzerland: Springer, 2021): 243; Giulio Mecacci, Pim Haselager, “Identifying Criteria for the Evaluation of the Implications of Brain Reading for Mental Privacy,” *Sci Eng Ethics* 25 (Dec., 2019): 443-461,

Οι ανησυχίες αυτές, γύρω από την κακή χρήση των προσωπικών δεδομένων, σχετίζονται με τον τρόπο συλλογής τους, αποθήκευσης, εκμετάλλευσης και σύνδεσής τους με άλλα δεδομένα, που θα μπορούσε να οδηγήσει ακόμη και στην εξακρίβωση της ταυτότητας ενός χρήστη. Παράλληλα, ένας ιδιαίτερα σημαντικός κίνδυνος που ελλοχεύει είναι η διάχυση των δεδομένων,³⁶⁰ ή και η υποκλοπή τους από εγκληματίες του κυβερνοχώρου,³⁶¹ που θα μπορούσε να επιφέρει ανυπολόγιστες συνέπειες, από την έκθεση του απορρήτου μέχρι και την χειραγώγηση της κοινής γνώμης.³⁶² Επί αυτού, το πρόσφατο σκάνδαλο της Cambridge Analytica,³⁶³ όπου τα δεδομένα εκατομμυρίων χρηστών του Facebook δόθηκαν στην εν λόγω εταιρία για πολιτικούς σκοπούς χωρίς τη ρητή συγκατάθεσή τους, απέδειξε πως η χειραγώγηση της κοινής γνώμης εν αναμονή των εκλογών είναι κάθε άλλο παρά αδύνατη, πράγμα που υπονομεύει τις δημοκρατικές διαδικασίες και την ελεύθερη βούληση των πολιτών.³⁶⁴

Οι ηθικές προκλήσεις που δημιουργούνται λοιπόν, είναι πολλές, κυρίως εάν λάβουμε υπόψη πως τα τελευταία χρόνια, οι έννοιες της ψυχικής υγείας και γενικότερα, του ψυχικού *eu ζην*, έχουν επεκταθεί σημαντικά σε πεδία όπως η ποινική δικαιοσύνη, η εκπαίδευση, η πολιτική και οι επιχειρήσεις, πεδία που σταδιακά ενσωματώνουν στη λειτουργία τους τις ψηφιακές τεχνολογίες και την Τ.Ν. Υπό το πρίσμα των αρχών της αγαθοποιίας και της μη-βλάβης, αυτό μπορεί να επιφέρει θετικές επιπτώσεις, καθώς οι παράγοντες εκείνοι που επηρεάζουν την υγεία του πληθυσμού είναι πολλοί (εισόδημα, εκπαίδευση, περιβάλλον κ.ο.κ.) και πολλές φορές, εντοπίζονται εκτός των υγειονομικών συστημάτων. Έτσι, η χρήση των αλγορίθμων μπορεί να συμβάλλει στον εντοπισμό των ατόμων που χρήζουν φροντίδας, κάτι που μακροπρόθεσμα θα μπορούσε να βελτιώσει σημαντικά την ψυχική υγεία σε επίπεδο πληθυσμιακό, μέσα από τη στοχευμένη κατανομή των πόρων και το σχεδιασμό των κατάλληλων παρεμβάσεων (πρόληψη ή αντιμετώπισης), φτάνοντας ακόμη και στους πιο απομακρυσμένους πληθυσμούς.³⁶⁵

Χαρακτηριστικά, στον τομέα της εκπαίδευσης, η χρήση των αλγορίθμων της μηχανικής μάθησης σε αρκετά εκπαιδευτικά ιδρύματα, στοχεύει στην παρακολούθηση των δεδομένων και την υποστηρικτική παρέμβαση στις περιπτώσεις των μαθητών που

<https://link.springer.com/article/10.1007/s11948-017-0003-3>; Pim Haselager, Giulio Mecacci, “Importance of Expiry Dates: Evaluating the Societal Impact of AI-Based Neuroimaging,” in *Artificial Intelligence in Brain and Mental Health: Philosophical, Ethical & Policy Issues*, ed. by Fabrice Jotterand, Marcello Ienca (Switzerland: Springer, 2021): 208.

³⁶⁰ Λόγου χάρι, σε εφαρμογές όπως το Joy (βλ. “Hi, I’m Joybot,” Facebook, accessed November 6, 2023, <https://www.facebook.com/hellojoyai/>), η πολιτική απορρήτου αναφέρει ότι με τη χρήση της υπηρεσίας οι χρήστες παραχωρούν στην εταιρεία την άδεια πλήρους εκμετάλλευσης του περιεχομένου της συνομιλίας, πράγμα που είναι άκρως ανησυχητικό. Kretzschmar et al., “Can Your Phone Be Your Therapist?” 4.

³⁶¹ Sebastian Sitaru, Georg Bramm, Alexander Zink, Matthias Hiller, “Cybersecurity in digital healthcare-challenges and potential solutions,” *Dermatologie* 74 (Feb., 2023): 213-217, <https://doi.org/10.1007/s00105-023-05117-6>.

³⁶² Simon Coghlan, Kobi Leins, Susie Sheldrick, Marc Cheong, Piers Gooding, Simon D’Alfonso, “To chat or bot to chat: Ethical issues with using chatbots in mental health,” *Digital Health* 9 (June, 2023): 7, <https://journals.sagepub.com/doi/10.1177/20552076231183542>.

³⁶³ «Τι μάθαμε από το σκάνδαλο της Cambridge Analytica για τα προσωπικά μας δεδομένα», Quantum Business & IT Solutions, προσπελάστηκε στις 10 Νοεμβρίου 2023, <https://www.quantum.gr/qblog/personal-data/>.

³⁶⁴ Montag, Sindermann, Baumeister, “Digital phenotyping,” 21.

³⁶⁵ Blasimme, Vayena, “The Ethics of AI in Biomedical Research,” 711-712.

φαίνεται να βρίσκονται σε κρίση ψυχικής υγείας.³⁶⁶ Με βάση αυτό, αν λάβουμε υπόψη το υποθετικό περιστατικό που περιγράψαμε στο προηγούμενο κεφάλαιο, η χρήση των συστημάτων της T.N. θα μπορούσε να εντοπίσει τις περιπτώσεις σχολικού εκφοβισμού σε ένα σχολείο και μέσω ενημερωτικών δράσεων, να αποτρέψει την εξάπλωσή τους και έτσι, την ψυχική βλάβη του συνόλου των μαθητών.

Οι στόχοι όμως παρόμοιων διαδικασιών δεν είναι πάντοτε υπέρ των ανθρώπων. Πολλές φορές η βασική επιδίωξη αφορά στο οικονομικό κέρδος, μια επιδίωξη που μπορεί να εσωκλείει στις μεθόδους της, όχι μόνο την κλοπή ή την παράνομη διακίνηση των δεδομένων, αλλά και την χρήση τους ως εργαλείο διακρίσεων και ανισοτήτων. Ως προς την περίπτωση κλοπής και παράνομης διακίνησης δεδομένων για παράδειγμα, το 2019 η διεθνής εφημερίδα Wall Street Journal, αποκάλυψε πληροφορίες γύρω από το "Project Nightingale", κατά το οποίο η Ascension, ένας από τους μεγαλύτερους ιδιωτικούς υγειονομικούς φορείς στις Ηνωμένες Πολιτείες Αμερικής, αποσκοπούσε με τη βοήθεια της Google και της προσβασιμότητας που αυτή διαθέτει σε λεπτομερή στοιχεία που αφορούν την υγεία εκατοντάδων ασθενών, στην αύξηση των εσόδων της, μέσω της συνταγογράφησης συγκεκριμένων φαρμάκων και της προώθησης περισσότερων ιατρικών εξετάσεων.³⁶⁷ Σε ένα ακόμη περιστατικό στο Ηνωμένο βασίλειο, η εθνική βάση ιατρικών δεδομένων με την ονομασία Care.data scheme, έκλεισε το 2016, αφού διαπιστώθηκε ότι πωλούσε δεδομένα ασθενών σε φαρμακευτικές και ασφαλιστικές εταιρείες.³⁶⁸ Τέτοιου είδους κίνδυνοι δεν είναι απίθανοι, εάν υποθέσουμε πως η διάχυση των δεδομένων που ένας χρήστης θα μοιραστεί με ένα chatbot ψυχικής υγείας, θα μπορούσαν να καταλήξουν στα αρχεία μιας ασφαλιστικής εταιρείας, η οποία στο βωμό του κέρδους θα επιχειρήσει να εκμεταλλευτεί τις σκέψεις και τα συναισθήματα του.

Πέρα από αυτό όμως, υπάρχουν αρκετά παραδείγματα που αποδεικνύουν πως η διάχυση και εκμετάλλευση των δεδομένων μπορούν να γίνουν η αιτία διακρίσεων και ανισοτήτων. Για παράδειγμα, στον ασφαλιστικό τομέα κάποιων χωρών, το ύψος των ασφαλιστρών φαίνεται να καθορίζουν παράγοντες, όπως η καταναλωτική συμπεριφορά και η ψυχική κατάσταση των ατόμων, που σκιαγραφούνται μέσα από τα δεδομένα που συλλέγονται και τα ψηφιακά τους αποτυπώματα στο διαδίκτυο.³⁶⁹ Κάτι ανάλογο θα μπορούσε να συμβεί και στο σύστημα της υγείας, με την προτεραιότητα νοσηλείας να δίνεται σε ασθενείς που θα αυξήσουν τα έσοδα των ιδιωτικών κλινικών. Πέρα από το κίνητρο του κέρδους, οι διακρίσεις και οι ανισότητες θα μπορούσαν να είναι το αποτέλεσμα των ενυπαρχουσών προκαταλήψεων.³⁷⁰ Λόγου χάρι, στον τομέα των τραπεζών, ένα λογισμικό πιστοληπτικής ικανότητας θα μπορούσε να θέσει κριτήρια σχετικά με το φύλο ή την καταγωγή των υποψηφίων δανειοληπτών,³⁷¹ ενώ και στον εργασιακό χώρο, τα προσωπικά δεδομένα ψυχικής υγείας στα χέρια των

³⁶⁶ Gooding, "Mapping the rise of digital mental health technologies," 1-2.

³⁶⁷ Χρυσή Χατζηβασιλείου, «Προσωπικά δεδομένα, Τεχνητή Νοημοσύνη, υπολογιστική νέφος και Διαδίκτυο των Πραγμάτων στον τομέα της Υγείας» (Διπλωματική διατριβή, Πανεπιστήμιο Μακεδονίας, Δημοκρίτειο Πανεπιστήμιο Θράκης, 2020): 42, <https://dspace.lib.uom.gr/handle/2159/24152>.

³⁶⁸ Gooding, "Mapping the rise of digital mental health technologies," 6-7.

³⁶⁹ *Αυτόθι*, 1-2.

³⁷⁰ Οι προκαταλήψεις αυτές μπορεί να απορρέουν από αναπόδεικτες πεποιθήσεις των ανθρώπων που λαμβάνουν τις διάφορες αποφάσεις (λ.χ. οι εργοδότες), έχουμε δει ωστόσο, πως η μεροληψία αποτελεί μια σημαντική αδυναμία των ίδιων των συστημάτων της T.N. (AI bias). Το θέμα αυτό, της μεροληψίας των συστημάτων, θα δούμε στην επόμενη ενότητα που αφορά στην ερευνητική διαδικασία.

³⁷¹ Χατζηβασιλείου, «Προσωπικά δεδομένα», 74-75.

εργοδοτών μπορούν κάλλιστα να εγείρουν ζητήματα εργατικού δικαίου, όπως το δικαίωμα πρόσβασης στην εργασία.³⁷²

Παραδείγματα όπως τα παραπάνω, φέρνουν στο προσκήνιο της συζήτησης την αρχή της δικαιοσύνης. Οι άνθρωποι που πάσχουν από ψυχικές ασθένειες βρίσκονται συχνά σε δυσμενή θέση λόγω αυτού. Όχι μόνο ταλαιπωρούνται ψυχικά από την κατάσταση τους, αλλά μπορεί και να αντιμετωπίζουν δυσκολίες σε όλους τους τομείς της ζωής τους, όπως η εύρεση και διατήρηση εργασίας, οι ανθρώπινες σχέσεις κ.ο.κ. Ο στιγματισμός τους μέσα από την παραβίαση του προσωπικού τους απορρήτου μπορεί να χαρακτηριστεί από άδικο έως και κακοποιητικός, το ίδιο και η άνιση αντιμετώπισή τους μέσω της χρήσης των εργαλείων της Τ.Ν.³⁷³ Το ίδιο ισχύει και για κάποιον που δεν πάσχει απαραίτητα από ψυχικό νόσημα, μα η ψυχολογική του κατάσταση είναι επιβαρυνμένη, λόγω δύσκολων καταστάσεων που βιώνει σε μία συγκεκριμένη περίοδο της ζωής του. Το ενδεχόμενο αυτό, η αναζήτηση συναισθηματικής ανακούφισης μέσω ενός chatbot ψυχικής υγείας να μην τύχει της σωστής αντιμετώπισης, ή ακόμη παραπέρα, να σταθεί η αιτία μετέπειτα δυσκολιών στην εύρεση εργασίας ή στην επιβίωση, φαίνεται να είναι πιθανό. Η γνώση της πιθανότητας αυτής, αρκεί για να αποτρέψει την αναζήτηση βοήθειας, παραβιάζοντας έτσι τόσο την αρχή της δικαιοσύνης και το καθολικό δικαίωμα πρόσβασης στην περίθαλψη, όσο και τις αρχές της αγαθοποιίας και μη βλάβης, μέσα από το ανυπολόγιστο ψυχικό κόστος που θα επιφέρει η περιθωριοποίηση και ο κοινωνικός αποκλεισμός.

Πέρα από τις περιπτώσεις κλοπής ή διάχυσης των προσωπικών δεδομένων για λόγους εξυπηρέτησης συμφερόντων, έχουν καταγραφεί περιπτώσεις όπου ο διαμοιρασμός των προσωπικών δεδομένων λαμβάνει χώρα με τρόπο επίσημο και κατόπιν κρατικών πρωτοβουλιών, με στόχο την επίτευξη της προστασίας των πολιτών. Ένα χαρακτηριστικό παράδειγμα συνέβη στην Φλόριντα των ΗΠΑ, εκεί όπου μετά από μακελειό σε σχολείο, η κυβέρνηση ενέκρινε τη δημιουργία βάσης δεδομένων των μελών της πολιτείας, ως μέρος ενός νομοσχεδίου για την πρόληψη της ένοπλης βίας. Το 2019, η βάση περιελάμβανε περισσότερα από δύο εκατομμύρια αρχεία με στοιχεία όσων υποβλήθηκαν σε ψυχιατρικές εξετάσεις, ενώ το επόμενο βήμα ήταν να συμπεριληφθούν αρχεία για περισσότερα από εννέα εκατομμύρια παιδιά σε αναδοχή, άτομα με ιστορικό στη χρήση ουσιών, αρχεία μαθητών που υπέστησαν εκφοβισμό και παρενόχληση λόγω της φυλής ή του σεξουαλικού τους προσανατολισμού, ενώ το ίδιο νομοσχέδιο υποχρέωνε τους γονείς να ενημερώνουν το σχολείο κατά την εγγραφή των παιδιών τους, για το εάν αυτά έχουν παραπεμφθεί σε υπηρεσίες ψυχικής υγείας. Ένα παράδειγμα σαν αυτό, δείχνει τον τρόπο με τον οποίο τα ευαίσθητα δεδομένα μπορούν να χρησιμοποιηθούν για την διευκόλυνση της εφαρμογής της αστυνομικής εξουσίας, χωρίς να έχει προηγηθεί η απαραίτητη ερευνητική τεκμηρίωση τέτοιων πρακτικών ή να παρέχεται η νομική και ηθική αιτιολόγησή τους.³⁷⁴

Κάτι ανάλογο δύναται να συμβεί και στον τομέα της ποινικής δικαιοσύνης. Σε αυτόν, οι πληροφορίες της ψυχικής υγείας μπορούν να παρέχουν μια ξεκάθαρη εικόνα γύρω από την ψυχική κατάσταση των ατόμων και έτσι, να επηρεάσουν τις ποινικές διαδικασίες (απόδοση ενοχής, μετριάσμός της ποινής κ.λπ.). Αυτό είναι σημαντικό, καθώς πράγματι, η ψυχική κατάσταση ενός ατόμου θα μπορούσε να ευθύνεται για την παραβατική του συμπεριφορά κι έτσι, η συμπερίληψη αυτών των πληροφοριών θα

³⁷² Gooding, “Mapping the rise of digital mental health technologies,” 8-9.

³⁷³ Coghlan et al. “To chat or bot to chat,” 6.

³⁷⁴ Gooding, “Mapping the rise of digital mental health technologies,” 8.

μπορούσε να επισημάνει την αναγκαιότητα της θεραπευτικής παρέμβασης ή και να κατευθύνει την απόφαση ενός δικαστηρίου προς αυτήν, έναντι της όποιας ποινής.³⁷⁵

Ωστόσο, οι πληροφορίες αυτές θα μπορούσαν να χρησιμοποιηθούν για την ενεργοποίηση ενός, εξουσιοδοτημένου από το κράτος, καταναγκασμού, με τη μορφή νοσοκομειακής κράτησης, υποχρεωτικής θεραπείας³⁷⁶ ή και συνεχούς παρακολούθησης. Τέτοιες μέθοδοι εξαναγκασμού γενικότερα, όμως, παραβιάζουν τη θεμελιώδη αρχή της αυτονομίας, πράγμα που θίγει εν τέλει την ανθρώπινη αξιοπρέπεια και ελευθερία, μέσα από την επιβολή μεθόδων στις οποίες το άτομο δεν συναινεί και οι οποίες βλάπτουν περαιτέρω τον ήδη ευάλωτο ψυχισμό του, αντί να τον ωφελούν. Επομένως, μια τέτοια τακτική, έρχεται σε αντίθεση και με τις αρχές της αγαθοποιίας και της μη βλάβης του ατόμου, καθώς απέχει από την έννοια της περίθαλψης και κλονίζει την εμπιστοσύνη του ανθρώπου στους φορείς φροντίδας.

Ήδη, σε περιοχές όπως το Κουίνσλαντ και η Νότια Αυστραλία, η νομοθεσία εξουσιοδοτεί τις υπηρεσίες υγείας να επιβάλλουν υποχρεωτικές συνθήκες παρακολούθησης σε ιατροδικαστικούς ψυχιατρικούς ασθενείς, μέσω εργαλείων, όπως το GPS ή ηλεκτρονικές συσκευές που, στερεωμένες στο ανθρώπινο σώμα, δύνανται να καταγράφουν και να μεταδίδουν τακτικά, δεδομένα σχετικά με την τοποθεσία ενός ατόμου, τις ποσότητες αλκοόλ στο αίμα του κ.ο.κ. και μέσω αλγορίθμων να προβλέπουν ακόμη και τις μελλοντικές του συμπεριφορές. Αν και οι σχετικές μελέτες αναφέρουν σημαντική μείωση των παραβατικών συμπεριφορών, το αίσθημα αυτό, της συνεχούς παρακολούθησης είναι αρκετό από μόνο του για να αποτρέψει τους ανθρώπους, ειδικά όσους ανήκουν σε μειονότητες ή περιθωριοποιημένες ομάδες, από το να αισθάνονται πραγματικά ελεύθεροι να ασκούν τα ανθρώπινα δικαιώματά τους.³⁷⁷

Πάνω σε παραδείγματα όπως αυτό αλλά και στο γενικότερο φαινόμενο της συνεχούς καταγραφής των προσωπικών δεδομένων της υγείας, αρκετοί ερευνητές κάνουν λόγο για έναν αναδυόμενο «πανοπτικό ψηφιακής ιατρικής», που στοχεύει σε ευάλωτες και περιθωριοποιημένες κοινότητες, συμπεριλαμβανομένων των ατόμων που χρήζουν ψυχικής φροντίδας και που η ανάγκη τους αυτή, τελικά τους φέρνει σε σύγκρουση με την ιατρική, τη δημόσια υγεία και το νόμο.³⁷⁸

Την έννοια του πανοπτισμού συναντούμε στη θεωρία του Michel Foucault, που εμπνευσμένος από το *πανοπτικόν* του Jeremy Bentham, αμφισβήτησε την άποψη ότι οι τεχνολογίες αναδύονται από ουδέτερες τακτικές ή τεχνολογική εμπειρογνωμοσύνη, αλλά αντίθετα αποσκοπούν στην ενίσχυση ενός συστήματος ευρύτερης επιτήρησης των πολιτών σε μία πειθαρχική κοινωνία εξουσίας και συνεχούς ελέγχου.³⁷⁹

Αναλυτικότερα, το 1785, ο Άγγλος Jeremy Bentham σχεδίασε το *πανοπτικόν*,³⁸⁰ ένα αρχιτεκτονικό σχέδιο κυκλικού κτιρίου φυλακών. Στο κέντρο του είχε σχεδιαστεί ένας πύργος, μέσα από τον οποίο ο φύλακας θα μπορούσε να παρακολουθεί το σύνολο των

³⁷⁵ *Αυτόθι*, 1.

³⁷⁶ *Αυτόθι*.

³⁷⁷ *Αυτόθι*, 7.

³⁷⁸ *Αυτόθι*, 2.

³⁷⁹ “Michel Foucault”, Stanford Encyclopedia of Philosophy, accessed November 3, 2023, <https://plato.stanford.edu/entries/foucault/>.

³⁸⁰ Η έννοια του πανοπτικού (*panopticon*), ως θεσμικό κτίριο (λ.χ. φυλακή, εργοστάσια) περιγράφεται από τον Jeremy Bentham στο βιβλίο του με τίτλο “Panopticon: Or the Inspection House”. Βλ. Bentham, Jeremy. *Panopticon: Or the Inspection House*. Dublin: Payne, 1791. https://www.google.gr/books/edition/Panopticon_Or_The_Inspection_House/SYc5AAAAcAAJ?hl

κρατούμενων. Λόγω της δομής του κτιρίου, ο φύλακας δεν θα ήταν αντιληπτός από τους κρατούμενους και επομένως, αυτοί δεν θα γνώριζαν τη στιγμή που παρακολουθούνται.³⁸¹ Το σχέδιο του πανοπτικού βασίστηκε στη θεωρία του Bentham γύρω από τις έννοιες της *ηδονής* και του *πόνου*, τις δύο κυρίαρχες δυνάμεις, όπως αναφέρει, στις οποίες υποτάσσεται η φύση του ανθρώπου. Κάθε ανθρώπινη ενέργεια, συνεχίζει, ορίζεται από την υποταγή αυτή και στοχεύει απόλυτα στην αύξηση της αίσθησης της ηδονής και στην αποφυγή της αίσθησης του πόνου.³⁸²

Έτσι, οι κρατούμενοι σε ένα πανοπτικό κτίριο φυλακών, στην προσπάθεια να αποφύγουν την αίσθηση του πόνου που θα δημιουργούσε η οποιαδήποτε ποινή, θα οδηγούνταν στη συμμόρφωση της συμπεριφοράς τους με τους κανόνες της φυλακής αυτής ή και στην επιτήρηση τόσο των άλλων, όσο και του ίδιου τους του εαυτού (αυτο-επιτήρηση).³⁸³ Με άλλα λόγια, η υποψία και μόνο της παρακολούθησης θα ήταν αρκετή, ώστε οι άνθρωποι να ενεργούν διαφορετικά από τον τρόπο που θα ήθελαν και θα μπορούσαν. Για τον Foucault,³⁸⁴ το πανοπτικό ήταν πολλά περισσότερα από το σχέδιο ενός κτιρίου. Ήταν το σύμβολο της πειθαρχικής εξουσίας, βασισμένο στο αξίωμα ότι «όσο πιο αυστηρά μας παρακολουθούν, τόσο καλύτερα συμπεριφερόμαστε».³⁸⁵

Σε μια προσπάθεια σύνδεσης της θεωρίας του πανοπτισμού με την αδιάκοπη καταγραφή δεδομένων, κατανοούμε την ηθική σημασία που αποκτά το ζήτημα του απορρήτου και το πώς αυτό διακυβεύεται από τη χρήση των αλγορίθμων. Η αίσθηση της διαρκούς παρακολούθησης, καθώς και η αβεβαιότητα για τη μελλοντική χρήση ή κατάχρηση των καταγεγραμμένων πληροφοριών, μπορεί να οδηγήσει το άτομο στο να τροποποιεί συνεχώς τη συμπεριφορά του, με βάση τους κοινωνικούς κανόνες και προσδοκίες και στην αδυναμία να λειτουργήσει και να εκφραστεί με ελευθερία.

Εργαλεία όπως το DeepMood,³⁸⁶ που επιτρέπουν τον εντοπισμό της διάθεσης με βάση τη δυναμική πληκτρολόγησης στο κινητό τηλέφωνο, είναι πράγματι ελπιδοφόρα σε πεδία όπως η πρόγνωση και η ακριβής διάγνωση. Ωστόσο, η διαρκής παρακολούθηση και καταγραφή δεδομένων που αφορούν από τη συναισθηματική κατάσταση ενός ατόμου, μέχρι το μοτίβο της αναπνοής ή τον τόνο της φωνής του, μπορεί να είναι τουλάχιστον επιβλαβείς για τον ψυχισμό του, λόγω της ψυχολογικής πίεσης που θα επιφέρει αυτή η μορφή ψυχολογικού ελέγχου, που αντιβαίνει στις αξίες των πολυπολιτισμικών φιλελεύθερων κοινωνιών.³⁸⁷ Ας αναλογιστούμε τα συστήματα της T.N. ως εργαλεία στα χέρια ενός θεραπευτή ή μια κοινωνία, όπου οι θεραπευτές ψυχικής υγείας έχουν πλήρως αντικατασταθεί από τα συστήματα αυτά. Η έλλειψη ελευθερίας στην έκφραση λόγω του φόβου μελλοντικής χρήσης των προσωπικών δεδομένων ψυχικής υγείας εναντίον του ίδιου του ατόμου, θα μπορούσε να αποτρέψει

³⁸¹ “Michel Foucault.”

³⁸² Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation* (Canada: Batoche Books, 2000), 33-35.

³⁸³ “Michel Foucault.”

³⁸⁴ Το σκεπτικό αυτό ο Foucault αναλύει στο έργο του με τίτλο «Επιτήρηση και Τιμωρία: Η Γέννηση της Φυλακής». Βλ. Φουκώ, Μισέλ, *Επιτήρηση και Τιμωρία: Η Γέννηση της Φυλακής*, μτφ. Κ. Χατζηδημού & Ι. Ράλλη (Αθήνα: Εκδόσεις Ράππα, 1989), <https://pernoampariza.files.wordpress.com/2013/12/cf86cebfcf855cebaccf8e-ceb5cf80ceb9cf844ceaecf81ceb7cf83ceb7-cebaceb1ceb9-cf84ceb9cebccf89cf81ceafceb1.pdf>.

³⁸⁵ “Michel Foucault.”

³⁸⁶ “Emotions matter,” DeepMood®, accessed November 10, 2023, https://deepmood.io/en_US/.

³⁸⁷ Blasimme, Vayena, “The Ethics of AI in Biomedical Research,” 710-711.

την ειλικρινή έκφραση σκέψεων και συναισθημάτων ή εξ ολοκλήρου, την αναζήτηση υποστήριξης. Κάτι τέτοιο κλονίζει στο σύνολό της την έννοια της ψυχικής φροντίδας, καθώς περιορίζει το βασικό χαρακτηριστικό της που είναι το μοίρασμα των συναισθημάτων.

Επιπλέον, δημιουργούνται φόβοι ότι κράτη και εταιρίες που έχουν πρόσβαση στα προσωπικά δεδομένα ψυχικής υγείας θα προβούν στη χρήση αυτών, με στόχο την τροποποίηση των ανθρώπινων συμπεριφορών για την μαζική επιβολή της εξουσίας και την προώθηση συγκεκριμένων συμφερόντων ή πρακτικών.³⁸⁸ Αναφερθήκαμε ήδη στο παράδειγμα της Cambridge Analytica. Σε αυτό, η μηχανική μάθηση και οι αλγόριθμοι της T.N. χρησιμοποιήθηκαν συγκεκριμένα, για τη σύνθεση ενός προφίλ των χρηστών και στη συνέχεια, την άσκηση ψυχολογικής επιρροής σε αυτούς (λ.χ. μέσω πολιτικών διαφημίσεων και εικόνων) με στόχο την πρόκληση των «κατάλληλων» συναισθημάτων, με έμφαση στους χρήστες εκείνους, που εντοπίστηκαν ως οι πιο επιρρεπείς σε τέτοιου είδους επηρεασμούς.³⁸⁹

Αυτή η μαζική συναισθηματική χειραγώγηση που έλαβε χώρα, δημιουργεί σοβαρούς κινδύνους για τους δημοκρατικούς θεσμούς και το κράτος δικαίου,³⁹⁰ ενώ παράλληλα, εγείρει σημαντικές ανησυχίες γύρω από την ενδεχόμενη συγκέντρωση της δύναμης στα πεδία των εξουσιών, μέσα από την υπερβολική εξάπλωση της τεχνολογίας, χωρίς την ανάλογη βάση αποδεικτικών στοιχείων ή ερωτήματα σχετικά με το σύνολο των επιπτώσεων στις ζωές των ατόμων.³⁹¹

Επιπλέον, αναμφισβήτητα έρχεται σε ρήξη με τις αρχές της Βιοηθικής, όταν αντιλαμβανόμαστε την πιθανότητα, το σημείο εκκίνησής της να τοποθετείται στον τομέα της ψυχικής υγείας και των θεραπευτικών chatbot. Συγκεκριμένα, η συνεχής παρακολούθηση των συναισθημάτων, μπορεί να επιφέρει την ψυχική βλάβη του χρήστη, ενώ η αναγκαιότητα της αυτο-επιτήρησης αποτελεί πλήγμα για την αρχή της αυτονομίας, με την έννοια της ελευθερίας στην έκφραση και δράση. Πέρα από αυτό, η στροφή της τεχνολογικής εξέλιξης προς συμφέροντα και εξουσιαστικές πρακτικές θίγει την αρχή της αγαθοποιίας, καθώς θέτει σκοπούς πέρα από την ουσιαστική φροντίδα των ατόμων με ευάλωτο ψυχισμό, αλλά και την αρχή της δικαιοσύνης, αφενός λόγω των νομικών θεμάτων που εγείρει η παραβίαση της ανθρώπινης ιδιωτικότητας και αφετέρου, λόγω του ότι η έλλειψη εμπιστοσύνης από τα ίδια τα άτομα στους φορείς φροντίδας, ως προς τη διαχείριση των προσωπικών τους πληροφοριών, καθιστά την απόφαση αναζήτησης βοήθειας δύσκολη και επίφοβη. Ακόμη, η υπερβολική χρήση των συστημάτων της T.N. έστω ως βοηθητικά εργαλεία στα χέρια του θεραπευτή, εμπεριέχει τον κίνδυνο αλλοίωσης της θεραπευτικής σχέσης και, λόγω των παραπάνω πιθανοτήτων, υπονομεύει την ικανότητα του θεραπευόμενου να παραμείνει αυτόνομος και να διατηρήσει ένα αίσθημα αυτοδιάθεσης απέναντι στο θεραπευτή του.³⁹²

Είναι βέβαιο πως υπάρχει μεγάλος ενθουσιασμός γύρω από την ιδέα της χρήσης της T.N. στη δημόσια υγεία, με τους φορείς χρηματοδότησης και τους διεθνείς οργανισμούς ήδη να επενδύουν σε δραστηριότητες δημόσιας υγείας σε χώρες με χαμηλό και μέσο εισόδημα.³⁹³

³⁸⁸ *Αυτόθι*, 712.

³⁸⁹ Παναγοπούλου-Κουτνατζή, *Τεχνητή Νοημοσύνη: Ο δρόμος προς έναν ψηφιακό συνταγματισμό*, 293.

³⁹⁰ *Αυτόθι*, 294.

³⁹¹ Gooding, “Mapping the rise of digital mental health technologies,” 9.

³⁹² Blasimme, Vayena, “The Ethics of AI in Biomedical Research,” 710-711.

³⁹³ *Αυτόθι*, 712.

Για να είναι όμως το βήμα αυτό προς όφελος των ανθρώπων, θα πρέπει πρωτίστως οι προσπάθειες και οι επενδύσεις να στραφούν στη δημιουργία ενός πλαισίου που θα διασφαλίζει την ιδιωτική ζωή του ατόμου και ταυτόχρονα θα βελτιώνει την ευημερία του συστήματος υγειονομικής περίθαλψης.³⁹⁴ Στο πλαίσιο αυτό, η προστασία των δεδομένων θα πρέπει να εξετάζεται πέρα από τα όρια της ψυχικής υγείας και να επεκτείνεται μεταξύ άλλων, στους διαχειριστές των συστημάτων υγείας, στις ιδιωτικές εταιρείες συλλογής προσωπικών δεδομένων ή και στις υπηρεσίες που τα διαχειρίζονται.³⁹⁵ Βασικά στοιχεία ενός τέτοιου σκοπού είναι η διαφάνεια των διαδικασιών συλλογής και χρήσης, η ανωνυμία και η συγκατάθεση του χρήστη.

Η διαφάνεια είναι σημαντική, καθώς πολλές φορές οι χρήστες έχουν άγνοια των τρόπων με τον οποίο λειτουργούν οι εφαρμογές της Τ.Ν. Λόγω των κινδύνων που εμπεριέχει η συλλογή, αποθήκευση και χρήση προσωπικών δεδομένων, οι τρόποι αυτοί λειτουργίας θα πρέπει να γίνουν διαφανείς προς στους χρήστες.³⁹⁶ Με τον τρόπο αυτό, θα γίνει σεβαστή η αυτονομία τους και θα επέλθει μια ουσιαστική ισορροπία ανάμεσα στα οφέλη και στους κινδύνους που εμπεριέχει η χρήση των εφαρμογών. Κυρίως όμως, η πλήρης ενημέρωση και διαφάνεια, θα αυξήσει το αίσθημα της εμπιστοσύνης των χρηστών, μέσα από την πίστη ότι τα καταγεγραμμένα δεδομένα δεν θα χρησιμοποιηθούν ποτέ εναντίον τους με ένα τρόπο κρυφό και πιθανά παράνομο.³⁹⁷ Η ενημέρωση των χρηστών και οι συχνές υπενθυμίσεις σχετικά με τις ρυθμίσεις του απορρήτου, θα πρέπει να γίνονται με προσοχή, ιδιαίτερα προς τους ανθρώπους εκείνους που, προβλήματα ψυχικής υγείας ή άλλες ευαλωτότητες, παρεμποδίζουν τον βαθμό κατανόησης τους, ενώ ιδιαίτερα χρήσιμη θα ήταν η εκπαίδευση των χρηστών πάνω σε θέματα προστασίας των δεδομένων, με ένα τρόπο που θα διασφαλίζει την ισότητα πρόσβασης όλων στη γνώση αυτή.³⁹⁸ Τα κομμάτια αυτά βέβαια, της ενημέρωσης ή και εκπαίδευσης δεν στερούνται των δικών τους δυσκολιών. Συχνά οι αλγόριθμοι ξεπερνούν το δημιουργό τους και έτσι, η αδυναμία των ίδιων των προγραμματιστών να κατανοήσουν τον τρόπο λειτουργίας τους δεν είναι πάντοτε δυνατή,³⁹⁹ με αποτέλεσμα το ίδιο δύσκολη να είναι και η ενημέρωση των ανθρώπων που κάνουν χρήση των αντίστοιχων εφαρμογών, σχετικά με τη λειτουργία των αλγορίθμων και τους τρόπους που συλλέγονται και επεξεργάζονται τα δεδομένα.⁴⁰⁰ Ίσως μια λύση για αυτό να αποτελεί ο τακτικός διαμοιρασμός της σχετικής γνώσης και η κάθε τόσο επισήμανση των κινδύνων που ελλοχεύουν.

Επιπλέον, είναι σημαντικό, τα προσωπικά δεδομένα της ψυχικής υγείας, εάν συλλέγονται να διατηρούνται με εχεμύθεια, ενώ εάν διαμοιράζονται, αυτό να γίνεται πάντοτε με τη συγκατάθεση των χρηστών και να διατηρείται η ανωνυμία τους. Ακόμη και αν με τη συγκατάθεση ενός χρήστη, τα δεδομένα κοινοποιούνται για σκοπούς έρευνας ή βελτίωσης των υπαρχουσών υπηρεσιών, τότε είναι πολύ σημαντικό να

³⁹⁴ Montag, Sindermann, Baumeister, “Digital phenotyping,” 22.

³⁹⁵ Gooding, “Mapping the rise of digital mental health technologies,” 1.

³⁹⁶ Coghlan et al. “To chat or bot to chat,” 4.

³⁹⁷ Montag, Sindermann, Baumeister, “Digital phenotyping,” 21.

³⁹⁸ Coghlan et al. “To chat or bot to chat,” 7-8.

³⁹⁹ Πρόκειται για το λεγόμενο πρόβλημα του μαύρου κουτιού (The Black Box Problem). Βλ. ενότητα 3.β. της παρούσας διατριβής.

⁴⁰⁰ Παναγοπούλου-Κουτνατζή, *Τεχνητή Νοημοσύνη: Ο δρόμος προς έναν ψηφιακό συνταγματισμό*, 86-87.

αφαιρούνται τα αναγνωριστικά εκείνα στοιχεία που ο χρήστης διαμοιράζεται κατά τη συνομιλία, όπως ονόματα, διευθύνσεις email και τηλεφωνικοί αριθμοί.⁴⁰¹

Μια ανησυχία που συναντούμε στην έννοια της συγκατάθεσης από το χρήστη, αφορά στο ότι λόγω της συνεχώς αυξανόμενης εξέλιξης και ενσωμάτωσης της Τ.Ν. στον τομέα της ψυχικής υγείας και όχι μόνο, είναι όλο και πιο δύσκολο να προβλεφθεί η μελλοντική χρήση των δεδομένων, αλλά και σε ποιους θα δοθεί η δυνατότητα πρόσβασης, για ποιους σκοπούς και υπό ποιες συνθήκες. Οι πιθανότητες επαναχρησιμοποίησης των δεδομένων και η σύνδεση τους με άλλα σύνολα δεδομένων στο μέλλον, καθιστούν την έννοια της συγκατάθεσης, ως μέτρο ασφαλείας και τρόπο προστασίας της ανθρώπινης αυτονομίας, ρευστή και αδύναμη. Ο δε κίνδυνος, τεράστιες ποσότητες πληροφοριών να βρεθούν στα χέρια των κυβερνήσεων ή των ιδιωτικών οργανώσεων, που συνάπτουν συμφωνίες με εθνικές κυβερνήσεις για τη συλλογή και ανάλυση δεδομένων από εκατομμύρια πολίτες, καθιστά τις παραπάνω πιθανότητες ιδιαίτερα προβληματικές.⁴⁰²

Συμπεραίνουμε με βάση τα παραπάνω, πως τα ζητήματα του απορρήτου και της προστασίας των δεδομένων είναι κρίσιμα στον ορίζοντα της έρευνας γύρω από τον μεγάλο όγκο πληροφοριών που συγκεντρώνονται στον τομέα ψυχικής υγείας. Για την ώρα, η νομοθεσία που τα διέπει προέρχεται από τον κανονισμό της Ευρωπαϊκής Ένωσης 2016/679 (*Γενικός Κανονισμός Προστασίας Δεδομένων - ΓΚΠΔ*),⁴⁰³ που συνιστά το κυριότερο νομοθετικό κείμενο αναγκαστικού δικαίου σε ισχύ.⁴⁰⁴

Αυτό που επιδιώκει ο ΓΚΠΔ, δεν είναι η απαγόρευση της συλλογής, επεξεργασίας και χρήσης των δεδομένων αλλά η επιτέλεσή τους με τρόπο θεμιτό, σε ένα θεσμικό νομοθετικό πλαίσιο που θα πληροί τις τέσσερις βασικές αρχές της Βιοηθικής.⁴⁰⁵ Μεταξύ άλλων, ο ΓΚΠΔ 2016/679 που τέθηκε σε εφαρμογή το Μάιο του 2018, επιδιώκει τη διασφάλιση των δεδομένων,⁴⁰⁶ μέσα από θεμελιώδεις αρχές (λ.χ. η αρχή της νομιμότητας,⁴⁰⁷ του χρονικού περιορισμού της περιόδου αποθήκευσης,⁴⁰⁸ της διαφάνειας στη διαδικασία⁴⁰⁹ κ.ο.κ.) και χωρίς αμφιβολία, οι έννοιες στις οποίες εστιάζει είναι καίριες και ιδιαίτερα κρίσιμες για τα θέματα που συζητήσαμε έως τώρα.

Ωστόσο, η συνεχής και ταχύτατη εξέλιξη της Τ.Ν. και η αντίστοιχη αύξηση των σκοπών που απαιτούν ολοένα και μεγαλύτερη συλλογή και επεξεργασία δεδομένων, καθιστά επιτακτική τη συνεχή επικαιροποίηση και αναθεώρηση των όποιων κανονισμών. Στην ουσία, αυτό που απαιτείται, για να επιτευχθεί η υπεύθυνη και ηθική εφαρμογή της Τ.Ν. στον τομέα της ψυχικής υγείας και όχι μόνο, είναι μια διακυβέρνηση που εξελίσσεται παράλληλα με τον ίδιο τον τομέα, εξερευνώντας νέους

⁴⁰¹ Kretzschmar et al., “Can Your Phone Be Your Therapist?” 4.

⁴⁰² Blasimme, Vayena, “The Ethics of AI in Biomedical Research,” 706.

⁴⁰³ «Κανονισμός (ΕΕ) 2016/679 του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου της 27ης Απριλίου 2016».

⁴⁰⁴ Παναγοπούλου-Κουτνατζή, *Τεχνητή Νοημοσύνη: Ο δρόμος προς έναν ψηφιακό συνταγματισμό*, 149.

⁴⁰⁵ *Αυτόθι*, 85.

⁴⁰⁶ Ως «δεδομένα που αφορούν την υγεία», ορίζονται τα δεδομένα εκείνα που σχετίζονται μεταξύ άλλων, με την σωματική ή ψυχική υγεία ενός φυσικού προσώπου και αποκαλύπτουν πληροφορίες σχετικά με την κατάσταση της υγείας του, συμπεριλαμβανομένης τις υπηρεσίες υγειονομικής φροντίδας που του έχουν παρασχεθεί (άρθ. 4, εδ. 15). Βλ. Χατζηβασιλείου, «Προσωπικά δεδομένα», 53-54; «Κανονισμός (ΕΕ) 2016/679 του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου της 27ης Απριλίου 2016».

⁴⁰⁷ Παναγοπούλου-Κουτνατζή, *Τεχνητή Νοημοσύνη: Ο δρόμος προς έναν ψηφιακό συνταγματισμό*, 156.

⁴⁰⁸ *Αυτόθι*, 165.

⁴⁰⁹ *Αυτόθι*, 167.

μηχανισμούς εποπτείας και αξιολόγησης για την αντιμετώπιση των ηθικών προκλήσεων που προκύπτουν και τις αμφιλεγόμενες πτυχές της καινοτομίας, ώστε να διασφαλιστεί πως οι νεότεροι κίνδυνοι δεν θα περάσουν απαρατήρητοι.⁴¹⁰

Ένα τέτοιο πλαίσιο διακυβέρνησης, που θα συμμορφώνεται με τις θεμελιώδεις αρχές την Βιοηθικής, είναι πιθανό να αυξήσει τον αριθμό των ανθρώπων εκείνων, που θα είναι πρόθυμοι να μοιραστούν πληροφορίες για την ψυχική τους διάθεση αλλά και την ποσότητα ή ποιότητα των πληροφοριών αυτών. Έτσι, θα καταστήσει τα εργαλεία που βασίζονται στην Τ.Ν. πιο ισχυρά στην ουσιαστική αντιμετώπιση των αναγκών της ψυχικής υγείας των ανθρώπων και στην αξιοπιστία των παρεχόμενων υπηρεσιών από τα συστήματα της Πρωτοβάθμιας Φροντίδας Υγείας.⁴¹¹

4.α.2. Έρευνα και εφαρμογή

Πέρα από το σεβασμό στην ανθρώπινη ιδιωτικότητα, η ηθικώς αποδεκτή χρήση των συστημάτων της Τ.Ν. στον τομέα της ψυχικής υγείας προϋποθέτει την πρότερη επιβεβαίωση της εγκυρότητας, της αξιοπιστίας και της αποτελεσματικότητάς τους, σε ένα πλαίσιο αξιολόγησης που θα ευθυγραμμίζεται με τις τέσσερις αρχές της Βιοηθικής. Η επιβεβαίωση αποτελεί μια σημαντική επιδίωξη σε όλες τις φάσεις της ανάπτυξης ενός συστήματος, (από το σχεδιασμό μέχρι την υλοποίησή του) και υποκινείται από την ανάγκη και τις απαιτήσεις μιας ανθρωποκεντρικής προσέγγισης της Τ.Ν. (*Human Centered AI*). Αυτή η προσέγγιση αναθέτει ευθύνες στους ανθρώπους σχεδιαστές και παρόχους και επικεντρώνεται, όχι στην ίδια την τεχνολογία, αλλά στους τρόπους με τους οποίους οι άνθρωποι τη «σχεδιάζουν και τη χρησιμοποιούν (ή την καταχρώνται)».⁴¹²

Ειδικότερα κατά το στάδιο του σχεδιασμού, τα τελευταία χρόνια, οι μελετητές της εφαρμοσμένης ηθικής εστιάζουν στην εύρεση τρόπων που διασφαλίζουν πρωτίστως το σεβασμό των ανθρωπίνων δικαιωμάτων (ιδιωτικότητα, ασφάλεια κ.ο.κ.), από τα πρώτα κιόλας στάδια της ανάπτυξης ενός συστήματος. Ένας αυξανόμενος αριθμός ερευνητικών εργασιών έχει αφιερωθεί στις ηθικές αξίες που οι προγραμματιστές επιδιώκουν να προστατεύσουν μέσω του σχεδιασμού των συστημάτων και παράλληλα, στην ενίσχυση της ικανότητας αυτών να ενισχύουν τη δικαιοσύνη, την ευθύνη και τη διαφάνεια, μέσα από τεχνικές που θα εξαλείψουν τις αδικαιολόγητες διακρίσεις και θα παρέχουν στους χρήστες καλύτερη κατανόηση του τρόπου με τον οποίο λειτουργούν. Στην επιδίωξη των παραπάνω στόχων, συμπεριλαμβάνονται οι διάφορες μορφές αξιολόγησης, δοκιμών, εκτιμήσεων και λογοδοσίας, ώστε, πριν τα συστήματα επικυρωθούν, να διαπιστωθεί εάν πράγματι ανταποκρίνονται στα απαραίτητα ηθικά πρότυπα και απαιτήσεις.⁴¹³

Όπως αναφέρουν οι Yeung et al., το σχεδιασμό ενός συστήματος θα πρέπει να συνοδεύει η διαβούλευση, δηλαδή η συζήτηση μεταξύ των προγραμματιστών και των ενδιαφερόμενων πλευρών (λ.χ. θεραπευτές, ασθενείς) και η ανταλλαγή απόψεων σχετικά με τον πιθανό αντίκτυπο του προτεινόμενου συστήματος σε ατομικό και

⁴¹⁰ Blasimme, Vayena, “The Ethics of AI in Biomedical Research,” 714.

⁴¹¹ Kretzschmar et al., “Can Your Phone Be Your Therapist?” 5.

⁴¹² David J. Gunkel, “Mind the gap: responsible robotics and the problem of responsibility,” *Ethics and Information Technology* 22 (Dec., 2020): 309, <https://doi.org/10.1007/s10676-017-9428-2>, (η μετάφραση δική μου).

⁴¹³ Karen Yeung, Andrew Howes, Ganna Pogrebna, “AI Governance by Human Rights–Centered Design, Deliberation, and Oversight. An End to Ethics Washing,” in *The Oxford Handbook of Ethics of AI*, ed. Markus D. Dubber, Frank Pasquale, Sunit Das (USA: Oxford University Press, 2020),86-87.

συλλογικό επίπεδο. Η συζήτηση αυτή είναι που θα συμβάλλει στην ευθυγράμμιση του σχεδιασμού με τις δημοκρατικές αξίες και στην επίλυση συγκρούσεων μεταξύ ατομικών δικαιωμάτων και συλλογικών συμφερόντων. Επιπλέον, θα δώσει την ευκαιρία της επανεξέτασης ή και της έγκαιρης τροποποίησης ενός συστήματος ώστε αυτό να ευθυγραμμιστεί πλήρως με τις ηθικές αρχές και αξίες.⁴¹⁴ Η άποψη αυτή βρίσκει σύμφωνους τους Blasimme και Vayena, που τονίζουν την ανάγκη να συμπεριληφθούν στις συζητήσεις και τις διαδικασίες λήψης αποφάσεων σχετικά με τη χρήση δεδομένων και αλγορίθμων σε συγκεκριμένους τομείς, όλα τα εμπλεκόμενα μέρη και ιδιαίτερα, οι φορείς και οι κοινότητες εκείνες που φέρουν ευαλωτότητες και που γενικότερα, είναι αποκλεισμένες από τους κύκλους της εξουσίας και των αποφάσεων (λ.χ. μειονότητες, ομάδες ασθενών).⁴¹⁵ Με τον τρόπο αυτό, οι σχεδιαστές των συστημάτων θα μπορούσαν να λάβουν γνώση τόσο περί των επιστημονικών ή οικονομικών πτυχών της χρήσης ενός συστήματος T.N. αλλά και περί των απόψεων και ανησυχιών των ατόμων εκείνων, που θα είναι οι τελικοί χρήστες ενός συστήματος και οι άμεσα επηρεαζόμενοι από αυτό.⁴¹⁶

Μια τέτοια τακτική ευθυγραμμίζεται πλήρως με τις τέσσερις αρχές της Βιοηθικής που συζητούμε στην παρούσα ενότητα. Αρχικά, η συμπερίληψη του τελικού χρήστη στις συζητήσεις αυτές, γύρω από τα οφέλη και τους κινδύνους που απορρέουν από τη χρήση ενός συστήματος, συμβαδίζει με τις αρχές της αγαθοποιίας και μη βλάβης, καθώς επικεντρώνει το σχεδιασμό και την ανάπτυξη των συστημάτων στη βελτίωση των παρεχόμενων υπηρεσιών ψυχικής φροντίδας (με βάση τις πραγματικές ανάγκες της κοινότητας) και στην αποφυγή των κινδύνων εκείνων που θα μπορούσαν να είναι επιβλαβείς για τον ανθρώπινο ψυχισμό. Επιπλέον, η ενεργή συμμετοχή των χρηστών στις συζητήσεις, κατά το σχεδιασμό και την ανάπτυξη των συστημάτων, επιδεικνύει τον απαραίτητο σεβασμό στην ανθρώπινη αυτονομία, καθώς το ίδιο το άτομο έχει λόγο στο είδος της ψυχικής φροντίδας που του παρέχεται, ενώ ταυτόχρονα, η συμπερίληψη των αποκλεισμένων ομάδων και η ενσωμάτωση των διαφορετικών απόψεων, ευθυγραμμίζεται με την αρχή της δικαιοσύνης, μέσα από τον εκδημοκρατισμό της T.N. και το σχεδιασμό της με σημείο αναφοράς την ποικιλομορφία και τη διαφορετικότητα.

Στο στάδιο που ακολουθεί το σχεδιασμό, αυτό της ενσωμάτωσης ενός συστήματος της T.N. στον τομέα της ψυχικής υγείας, τα βασικά χαρακτηριστικά που συνθέτουν το ηθικό πλαίσιο της εφαρμογής του είναι σαφώς, η αξιοπιστία και η εγκυρότητά του. Ένας από τους τρόπους εξασφάλισης αυτών, είναι η θεωρητική τους πλαισίωση με βάση αποδεδειγμένα επιστημονικά δεδομένα. Η διάθεση στο κοινό, για παράδειγμα, ενός chatbot ψυχικής υγείας μέσω του οποίου ένα άτομο θα αναζητήσει συμβουλευτική υποστήριξη, αν μη τι άλλο προϋποθέτει, οι τεχνικές και οι παρεμβάσεις που αυτό εφαρμόζει, να βασίζονται σε θεωρητικές προσεγγίσεις οι οποίες έχουν στοιχειοθετηθεί, ως προς την αποτελεσματικότητά τους, μέσω των αντίστοιχων επιστημονικών ερευνών.⁴¹⁷

Η σημασία αυτή της καθοδήγησης από τη θεωρία έχει ιδιαίτερη αξία. Λαμβάνοντας υπόψη τις μέχρι τώρα επιστημολογικές και μεθοδολογικές αδυναμίες των συστημάτων να ανταποκριθούν στις ανάγκες ψυχικά ευάλωτων ατόμων ή και σε καταστάσεις

⁴¹⁴ Yeung, Howes, Pogrebna, “AI Governance by Human Rights–Centered Design,” 89.

⁴¹⁵ Blasimme, Vayena, “The Ethics of AI in Biomedical Research,” 715.

⁴¹⁶ Coghlan et al., “To chat or bot to chat,” 8; Guan, “Artificial Intelligence in Healthcare and Medicine,” 78-79.

⁴¹⁷ Kretschmar et al., “Can Your Phone Be Your Therapist?” 5.

έκτακτης ανάγκης (λ.χ. αυτοκτονικός ιδεασμός), η διαχείριση των περιπτώσεων άνευ θεωρητικού υπόβαθρου, μπορεί να αποβεί από ακατάλληλη έως και επικίνδυνη για το άτομο,⁴¹⁸ θίγοντας έτσι τις αρχές της αγαθοποιίας και της μη βλάβης όσων αναζητούν βοήθεια.

Η αλήθεια είναι πως μέχρι σήμερα, τα περισσότερα chatbot ψυχικής υγείας (όπως το Woebot) βασίζονται στην προσέγγιση της γνωσιακής συμπεριφορικής θεραπείας (Cognitive Behavioral Therapy - CBT).⁴¹⁹ Παρά το γεγονός αυτό, όμως, έχουμε δει πως δεν είναι ακόμα σε θέση να αναπαράγουν επαρκώς την θεραπευτική αλληλεπίδραση που λαμβάνει χώρα μεταξύ των θεραπευτών και των θεραπευόμενων τους ή και να αντικαταστήσουν τον άνθρωπο θεραπευτή. Επιπλέον, αν λάβουμε υπόψη την πιθανότητα αυτή η μορφή προσέγγισης να μην είναι η κατάλληλη για όλα τα άτομα, τότε κατανοούμε την ανάγκη της παράλληλης επίβλεψης από κλινικά εκπαιδευμένο προσωπικό, στην προσπάθεια αποφυγής πρόκλησης περαιτέρω ψυχικής βλάβης στο άτομο ή και απαξίωσης του ρόλου των υπηρεσιών ψυχικής φροντίδας λόγω της αναξιοπιστίας και της αναποτελεσματικότητάς τους.⁴²⁰

Πέραν τούτων, ιδιαίτερη σημασία ενδέχεται να έχει η εμπειρία που αποκτάται μέσω της δοκιμής συστημάτων, όπως τα chatbot και η οποία θα μπορούσε να επιβεβαιώσει ή όχι, το ότι πράγματι, αυτά συμμορφώνονται με τα πρότυπα των ανθρωπίνων δικαιωμάτων και τις αρχές της Βιοηθικής. Η δοκιμαστική αυτή υλοποίηση σε πραγματικό χρόνο, αποτελεί το κατάλληλο έδαφος, που θα αναδείξει τους κινδύνους που ελλοχεύουν για την ασφάλεια και τα δικαιώματα του χρήστη καθώς και το εάν στην πράξη, η λειτουργία των συστημάτων ανταποκρίνεται στους στόχους που εξ' αρχής έχουν τεθεί ή αυτά χρήζουν άμεσης επανεξέτασης και τροποποίησης προτού διατεθούν στο ευρύτερο κοινό.⁴²¹

Ιδανικά, μια πειραματική εφαρμογή ενός chatbot, θα πρέπει να περιλαμβάνει τυχαίες ελεγχόμενες δοκιμές σε δείγματα ατόμων με ποικιλομορφία στα χαρακτηριστικά (όπως στο φύλο και στην ηλικία), καθώς τα χαρακτηριστικά αυτά είναι σημαντικά στη διαμόρφωση του τρόπου με τον οποίο οι άνθρωποι βιώνουν την τεχνολογία.⁴²² Μάλιστα, το στοιχείο αυτό της διαφορετικότητας είναι εξέχουσας σημασίας, εάν αναλογιστούμε μια ήδη υπάρχουσα αδυναμία των συστημάτων της Τ.Ν. που είναι η πιθανότητα μεροληψίας από πλευράς τους (AI bias).

Συγκεκριμένα, η έλλειψη αντιπροσωπευτικότητας των διαφορετικών πληθυσμιακών ομάδων⁴²³ στα δεδομένα που εκπαιδεύουν τους αλγορίθμους μηχανικής μάθησης μπορεί να οδηγήσει στην υπερεκπροσώπηση κάποιων ομάδων έναντι άλλων και επομένως σε προκαταλήψεις και αυξημένες υγειονομικές ανισότητες.⁴²⁴ Για παράδειγμα, στον τομέα της υγείας, συχνά η κατάσταση της υγείας ανθρώπων χαμηλού βιοτικού και μορφωτικού επιπέδου εξελίσσεται αρνητικά ή δεν ανταποκρίνεται

⁴¹⁸ Montag, Elhai, “A new agenda,” 131-132.

⁴¹⁹ Fitzpatrick, Darcy, Vierhile, “Delivering Cognitive Behavior Therapy,” 1, 9.

⁴²⁰ Coghlan et al. “To chat or bot to chat,” 5.

⁴²¹ Yeung, Howes, Pogrebnia, “AI Governance by Human Rights–Centered Design,” 89-90.

⁴²² Coghlan et al. “To chat or bot to chat,” 8; Kretzschmar et al., “Can Your Phone Be Your Therapist?” 5.

⁴²³ Η ελλιπής εκπροσώπηση συγκεκριμένων ομάδων θα μπορούσε να επηρεάσει ακόμη και τον τόπο με τον οποίο κατανέμεται η χρηματοδότηση της εθνικής έρευνας, εγείροντας θέματα ανισότητας εις βάρος τους. Βλ. Coghlan et al. “To chat or bot to chat,” 6.

⁴²⁴ Blasimme, Vayena, “The Ethics of AI in Biomedical Research,” 705-706.

επαρκώς στις θεραπευτικές παρεμβάσεις, λόγω της καθυστερημένης αναζήτησης υγειονομικής περίθαλψης, εξαιτίας παραγόντων όπως η οικονομική αδυναμία, τα ωράρια εργασίας, η άγνοια ή αδυναμία αυτο-φροντίδας (λ.χ. μοναχικοί ηλικιωμένοι, χαμηλό νοητικό επίπεδο) και οι ελλείψεις δομών στην περιοχή που ζουν. Σε μια τέτοια περίπτωση, ένας αλγόριθμος της Τ.Ν., μη λαμβάνοντας υπόψη τις κοινωνικές αυτές παραμέτρους, θα μπορούσε να συμπεράνει πως άνθρωποι με αυτά τα χαρακτηριστικά, θα εμφανίζουν πάντοτε κακή ανταπόκριση στη θεραπεία και έτσι, να συνιστά τη μη παροχή φροντίδας σε αυτούς εξ αρχής, εντείνοντας τις ανισότητες στη φροντίδα και στην προσπάθεια επίτευξης της ίασης.⁴²⁵

Τέτοιου είδους λανθασμένα συμπεράσματα θα μπορούσαν να εξαχθούν και μέσω των δεδομένων που συλλέγονται από ένα chatbot ψυχικής υγείας στο οποίο ο χρήστης έχει απευθυνθεί. Εκεί, τα δεδομένα χρησιμοποιούνται για να διαμορφώσουν τις απαντήσεις προς τους χρήστες και έτσι, απρόσκοπτες προκαταλήψεις μπορεί να προκύψουν εναντίον ανθρώπων με παρόμοια υπερ-εκπροσωπούμενα στις βάσεις δεδομένων χαρακτηριστικά, κάτι που θα επιφέρει αποτυχία στην παροχή σωστών κατευθύνσεων και συμβουλών και τη γενικότερη μη αξιόπιστη ανταπόκριση στο αίτημα που εκφράζεται από το χρήστη.⁴²⁶ Ας φανταστούμε την περίπτωση μιας μητέρας συγκεκριμένης καταγωγής, που καλεί σε μία γραμμή βοήθειας για να ζητήσει κατευθύνσεις σχετικά με το γονεϊκό της ρόλο. Η ενδεχόμενη συχνή καταγραφή περιστατικών κακοποίησης σε οικογένειες αυτής της καταγωγής, θα μπορούσε να έχει ως συνέπεια, η μητέρα που καλεί να αναγνωριστεί αυτόματα από το σύστημα της Συνομιλητικής Τ.Ν. ως κακοποιητική και παραμελητική προς τα παιδιά της, χωρίς να είναι. Τέτοιου είδους παραδείγματα, δημιουργούν ανησυχίες για τους τρόπους με τους οποίους η Τ.Ν. μπορεί να ενισχύσει νέες μορφές ταξινόμησης βασισμένες στον συσχετισμό μεταξύ βιολογικών, προσωπικών, συμπεριφορικών και κοινωνικών χαρακτηριστικών, προκαλώντας το στιγματισμό και τις διακρίσεις μεταξύ των ανθρώπων και εν συνεχεία, την ακατάλληλη διαχείριση των περιστατικών.⁴²⁷

Προς αποφυγή της μεροληψίας λοιπόν, η ισότιμη εκπροσώπηση των διαφορετικών ομάδων κατά τις δοκιμές των συστημάτων και όχι μόνο, τυγχάνει ιδιαίτερης αξίας. Η συμπερίληψη όσο το δυνατόν περισσότερων χαρακτηριστικών κατά το σχεδιασμό ενός συστήματος είναι σύμφωνη με τις αρχές της δικαιοσύνης και της αγαθοποιίας, καθώς, επιτυγχάνει την αποτροπή των προκαταλήψεων και την ισότητα των διαφορετικών ομάδων απέναντι στην εξέλιξη της τεχνολογίας και συνεπώς, την ωφέλεια για τις περισσότερες αυτές ομάδες του πληθυσμού.⁴²⁸

Πέρα από την πιθανότητα εμφάνισης μεροληψίας από τα συστήματα της Τ.Ν., μια ακόμη αδυναμία τους στην οποία και αξίζει να αναφερθούμε είναι και αυτή της υπερδιάγνωσης (overdiagnosis) ή της ανίχνευσης τάσεων εμφάνισης διαταραχών, που τη δεδομένη στιγμή δεν επιδέχονται παρεμβάσεις.⁴²⁹ Αναφερθήκαμε στο φαινόμενο αυτό στην Ενότητα 3.δ. της παρούσης διατριβής και είδαμε τον τρόπο με τον οποίο μπορεί να επηρεάσει αρνητικά την ποιότητα της ζωής ενός ατόμου, μέσα από την ψυχική επιβάρυνση που μπορούν να επιφέρουν οι επιπλέον ανησυχητικές πληροφορίες για την υγεία του ή και η υπερβολική ιατροκοποίηση (medicalization) των

⁴²⁵ *Αυτόθι*, 713.

⁴²⁶ Coghlan et al. “To chat or bot to chat,” 7.

⁴²⁷ Blasimme, Vayena, “The Ethics of AI in Biomedical Research,” 707.

⁴²⁸ Coghlan et al. “To chat or bot to chat,” 8.

⁴²⁹ Blasimme, Vayena, “The Ethics of AI in Biomedical Research,” 710.

πληροφοριών που εξάγονται από το πλήθος των δεδομένων.⁴³⁰ Προς αποφυγή της ψυχικής αυτής βλάβης του ατόμου λοιπόν, η εμπειρική δοκιμή των συστημάτων της Τ.Ν. κρίνεται απαραίτητη σε ένα πλαίσιο πειραματισμού και δοκιμής που θα μπορούσε να εντοπίσει έγκαιρα την αδυναμία αυτή του συστήματος και να την αποτρέψει, μέσα από την αξιολόγηση του τρόπου που λειτουργεί και την βελτιστοποίησή του εκεί που χρειάζεται.

Η αξιολόγηση αυτή βέβαια, δεν αφορά μόνο σε μία διαδικασία που ακολουθεί την πειραματική εφαρμογή ενός συστήματος αλλά και σε μία επαναλαμβανόμενη τακτική που απορρέει από την υποχρέωση των παρόχων των υπηρεσιών ψυχικής υγείας στην συνεχή βελτίωση της ποιότητας των υπηρεσιών αυτών.⁴³¹

Όπως ισχυρίζονται οι Yeung et al., οι τακτικές αξιολογήσεις είναι απαραίτητες για την καταγραφή των αρνητικών επιπτώσεων που προκύπτουν, συμπεριλαμβανομένων των πιθανών παραβιάσεων των ανθρωπίνων δικαιωμάτων, με βάση τις αναφορές που γίνονται από τους χρήστες ή το ευρύτερο κοινό. Τέτοιες αναφορές υποχρεώνουν τους παροχείς του συστήματος να επανεξετάσουν τον σχεδιασμό και να προβούν στις απαραίτητες τροποποιήσεις, η δημοσιοποίηση των οποίων ενισχύει το απαραίτητο αίσθημα εμπιστοσύνης από τα άτομα που κάνουν χρήση της εν λόγω εφαρμογής ή όσους διστάζουν να αναζητήσουν βοήθεια με αυτό τον τρόπο. Έτσι, μια τέτοια πολιτική επανεξέτασης και αλλαγών όπου χρειάζεται, αν μη τι άλλο ενισχύει την αποτελεσματικότητα και την αξιοπιστία των συστημάτων, υπηρετώντας τις αρχές της αγαθοποιίας και μη βλάβης, μέσα από την βελτίωση των παρεχόμενων υπηρεσιών και την ελαχιστοποίηση των διαφόρων κινδύνων.⁴³²

Η τακτική αυτή αξιολόγηση των βλαβών και των οφελών ενός συστήματος είναι βασικό συστατικό της ηθικά υπεύθυνη πρακτικής εφαρμογής τους και είναι ιδιαίτερα ανησυχητικό το γεγονός πως, μέχρι στιγμής, αρκετές από τις εφαρμογές της Συνομιλητικής Τ.Ν. δεν προβαίνουν σε αυτήν.⁴³³ Αυτή η ασυμμετρία, όμως, ανάμεσα στην ήδη εκτεταμένη χρήση των chatbot ψυχικής υγείας και στην ελλιπή επιστημονική ανάλυση και αξιολόγηση των επιπτώσεων τους,⁴³⁴ μπορεί να αποτελέσει σημαντικό πλήγμα για τις αρχές της Βιοηθικής και επιπλέον, να καταστήσει δύσκολη την έγκαιρη και σαφή διατύπωση των κατάλληλων ρυθμιστικών πλαισίων (νομικών και ηθικών) που θα διέπουν την χρήση τους στη Συμβουλευτική. Κατά συνέπεια, η ανάγκη αυτή για εκτενέστερη επιστημονική έρευνα είναι επιτακτική για τη στήριξη των κανονιστικών πλαισίων που θα προστατέψουν τις ανάγκες και τα δικαιώματα των ψυχικά ευάλωτων πληθυσμών.

Επιπλέον, στον εντοπισμό των προβληματικών πτυχών της χρήσης των chatbot θα μπορούσαν να συμβάλλουν οι κατάλληλοι μηχανισμοί εποπτείας, με κύριο καθήκον την παρακολούθηση και την επιβολή ηθικών απαιτήσεων γύρω από τον σχεδιασμό, τον έλεγχο και την αξιολόγηση της λειτουργίας των εφαρμογών, σε ένα νομικό πλαίσιο διαβούλευσης που θα περιλαμβάνει, εκτός από τους προγραμματιστές των συστημάτων, τους ενδιαφερόμενους φορείς της ψυχικής υγείας και εκπροσώπους των

⁴³⁰ Conrad, Mackie, Mehrotra, “Estimating the costs of medicalization,” 1943.

⁴³¹ Yeung, Howes, Pogrebna, “AI Governance by Human Rights–Centered Design,” 89-90; Kretzschmar et al., “Can Your Phone Be Your Therapist?” 5.

⁴³² Yeung, Howes, Pogrebna, “AI Governance by Human Rights–Centered Design,” 89-90; Coghlan et al. “To chat or bot to chat,” 7.

⁴³³ Kretzschmar et al., “Can Your Phone Be Your Therapist?” 7.

⁴³⁴ Sedlakova, Trachsel, “Conversational Artificial Intelligence in Psychotherapy,” 5.

ομάδων στις οποίες απευθύνονται τα συστήματα. Ένα τέτοιο σύστημα «επιτήρησης της Τ.Ν.», υποστηρίζουν οι Yeung et al., είναι απαραίτητο για τη διασφάλιση ότι τα συστήματα της Τ.Ν. στον τομέα της ψυχικής υγείας παρέχουν αξιόπιστες υπηρεσίες και συμμορφώνονται με τις αρχές της Βιοηθικής.⁴³⁵

Σίγουρα, η εποπτεία της ανάπτυξης μιας Τ.Ν. επικεντρωμένης στον άνθρωπο και στα δικαιώματά του μπορεί να αυξήσει το κόστος του σχεδιασμού ή και να επιβραδύνει την διάθεσή των συστημάτων στην κοινότητα. Μακροπρόθεσμα όμως, αυτή η προσέγγιση θα διευκολύνει τη σταδιακή προσαρμογή των κανονιστικών πλαισίων στις εξελίξεις της έρευνας και της τεχνολογίας και θα αυξήσει την αποτελεσματικότητα και την ακρίβεια εφαρμογών όπως τα chatbot, επιτυγχάνοντας έτσι, την ασφαλή και αποτελεσματική ενσωμάτωση τους στον τομέα της Συμβουλευτικής.

4.α.3. Ασφάλεια του χρήστη

Πολλά από αυτά στα οποία έχουμε αναφερθεί έως τώρα, συνδέονται άμεσα με το θέμα της προστασίας των ανθρώπων που θα αναζητήσουν βοήθεια μέσω ενός chatbot ψυχικής υγείας. Πιθανοί κίνδυνοι που περιγράψαμε, όπως η παραβίαση της ιδιωτικότητας, η υπερδιάγνωση, η μεροληψία των συστημάτων της Τ.Ν. ή και η έλλειψη αξιοπιστίας και αποτελεσματικότητάς τους, αν μη τι άλλο διακυβεύουν την ασφάλεια ενός χρήστη σε ευάλωτη ψυχολογική κατάσταση και επισημαίνουν την αναγκαιότητα, αυτή η ασφάλεια, να αποτελέσει προτεραιότητα, κατά την ανάπτυξη και εφαρμογή της τεχνολογίας. Ή πιο συγκεκριμένα, να αποτελέσει τη νομική υποχρέωση των προγραμματιστών και των παρόχων των διαφόρων εφαρμογών, σε ένα πλαίσιο διακυβέρνησης της Τ.Ν. που θα αποτρέψει τον κίνδυνο παραβίασης των ανθρωπίνων δικαιωμάτων, μέσα από τη θέσπιση ανθεκτικών και συστηματικών μηχανισμών διαφάνειας και εποπτείας, την παροχή επαρκούς ενημέρωσης προς το κοινό και την αποσαφήνιση του ρόλου της ανθρώπινης ευθύνης για τις επιπτώσεις των συστημάτων και την ποιότητα των παρεχόμενων υπηρεσιών.⁴³⁶

Ένα επιπλέον σημαντικό θέμα στο οποίο θα σταθούμε εδώ, αφορά στη σχέση που αναπτύσσεται ανάμεσα σε ένα άτομο και μια εφαρμογή της Τ.Ν., στην περίπτωση μας ένα chatbot ψυχικής υγείας. Η σχέση αυτή ξεκινά από το ίδιο το άτομο, που λόγω της ευάλωτης ψυχολογικής κατάστασης στην οποία βρίσκεται, εκφράζει την ανάγκη του για ψυχολογική υποστήριξη σε ένα chatbot, το οποίο και αναμένεται να του την παρέχει. Στη θεραπευτική αυτή αλληλεπίδραση που λαμβάνει χώρα μεταξύ τους, θα λέγαμε πως οι ηθικές απαιτήσεις που δημιουργούνται για ένα chatbot μπορούν να συμπυκνωθούν σε τρία βασικά πλαίσια συζήτησης: την κατανόηση των αναγκών του χρήστη, την ικανότητα επαρκούς ανταπόκρισης στις ανάγκες αυτές και τη δημιουργία μιας σχέσης ικανής να οδηγήσει στη θεραπευτική αλλαγή, που άλλωστε αποτελεί και το ζητούμενο της συμβουλευτικής διαδικασίας.

Αναφορικά με το πρώτο, με βάση τα όσα αναλύσαμε στα κεφάλαια που προηγήθηκαν, καταλήγουμε εύκολα στο συμπέρασμα πως η συμβολή της Τ.Ν. στο πεδίο της ψυχικής υγείας είναι ιδιαίτερα σημαντική και ελπιδοφόρα. Μέσα από την ταχύτερη επεξεργασία ενός μεγάλου όγκου δεδομένων, τη σαφή επιστημονική ανωτερότητά της σε σχέση με τον άνθρωπο και την αξιοποίηση εργαλείων, όπως ο ψηφιακός φαινότυπος και η επεξεργασία της φυσικής γλώσσας, τα συστήματά της Τ.Ν. μπορούν σε παρέχουν έναν αριθμό σημαντικών πληροφοριών για την ψυχολογική κατάσταση στην οποία

⁴³⁵ Yeung, Howes, Pogrebnna, “AI Governance by Human Rights–Centered Design,” 91.

⁴³⁶ *Αυτόθι*,” 93.

βρίσκεται ένα άτομο με τρόπο αντικειμενικό και εξατομικευμένο, χωρίς καμία συνδρομή από το ίδιο.⁴³⁷ Τέτοιου είδους δυνατότητες αν μη τι άλλο είναι σύμφωνες με την αρχή της αγαθοποιίας στο πεδίο της ψυχικής υγείας, καθώς συμβάλλουν στην έγκαιρη ανίχνευση και διάγνωση των ψυχικών διαταραχών, ενώ μακροπρόθεσμα μπορούν να βελτιώσουν την ποιότητα των παρεχόμενων υπηρεσιών και να συντελέσουν στην δημιουργία ενός πιο προληπτικού και εξατομικευμένου μοντέλου παροχής ψυχικής φροντίδας.⁴³⁸ Παράλληλα, η διευκόλυνση στην πρόσβαση των ατόμων, μέσα από την δωρεάν, καθημερινή και εικοσιτετράωρη διάθεσή των εφαρμογών, ακόμη και στους πιο απομακρυσμένους πληθυσμούς,⁴³⁹ χωρίς αμφιβολία προάγει την αρχή της δικαιοσύνης, μέσα από την διασφάλιση της ισότητας και των ίσων ευκαιριών στη φροντίδα.

Ωστόσο, έχουμε δει, πως συχνά, παράμετροι όπως το πολύπλοκο λεξιλόγιο που χαρακτηρίζει μια κρίση ψυχικής υγείας, οι ασαφείς εκφράσεις όπως το χιούμορ ή ο σαρκασμός, τα σύνθετα ερωτήματα που διατυπώνονται από τον χρήστη, αλλά και οι περιορισμένες αντιδράσεις των chatbot ή η υπό αμφισβήτηση ικανότητά τους για επίδειξη ενσυναίσθησης, καθιστούν την δυνατότητα κατανόησης των ουσιαστικών αναγκών που εκφράζονται από το χρήστη ανύπαρκτη.⁴⁴⁰ Πολλώ δε μάλλον, εάν συμπεριλάβουμε στη συζήτησή μας αδυναμίες των συστημάτων, όπως η μεροληψία και η υπερδιάγνωση που αναφέραμε νωρίτερα, καθώς και το λεγόμενο «Πρόβλημα του Πλαισίου» (The Frame Problem),⁴⁴¹ που αφορά στη δυσκολία καθορισμού και συγκεκριμενοποίησης των πληροφοριών ως προς τη συνάφειά τους με μία συγκεκριμένη κατάσταση.

Υπάρχουν σημαντικοί λόγοι για τους οποίους η ελλιπής κατανόηση από τα chatbot είναι ανησυχητική και μπορεί να έρθει σε ρήξη με τις αρχές της Βιοηθικής. Αρχικά, η επικοινωνία αυτή μεταξύ συμβούλου και συμβουλευόμενου, έστω και σε ένα αρχικό στάδιο απλής κατανόησης του αιτήματος που εκφράζεται, αποτελεί ένα απαραίτητο, σχεδόν αυτονόητο συστατικό στοιχείο της συμβουλευτικής διαδικασίας. Η μη επίτευξη αυτής αποτελεί ένα απροσπέλαστο εμπόδιο στο να προχωρήσει η διαδικασία αυτή παρακάτω και στην ουσία, υπονοεί την απόρριψη του αιτήματος που εκφράζεται και της στήριξης που αναζητείται. Για παράδειγμα, εάν ο δεκάχρονος πρωταγωνιστής του υποθετικού περιστατικού στο προηγούμενο κεφάλαιο αισθανθεί ότι το chatbot αδυνατεί να κατανοήσει αυτό που βιώνει, είναι πολύ πιθανό ότι η κλήση θα τερματιστεί. Το ίδιο θα κάνει και ένας έφηβος που δηλώνει αυτοκτονικές τάσεις και ενέχεται πάντοτε ο κίνδυνος, λόγω απογοήτευσης, κανένας από τους δύο να μην προβεί ξανά σε μια αντίστοιχη προσπάθεια. Η απώλεια του ελέγχου όμως, και η μη διαχείριση

⁴³⁷ D'Alfonso, "AI in mental health," 112-113; Montag, Sindermann, Baumeister, "Digital phenotyping" 19-20; Stange et al., "Let your fingers do the talking," 285; Bedi et al., "Automated analysis of free speech," 1-7; Fitzpatrick, Darcy, Vierhile, "Delivering Cognitive Behavior Therapy," 1, 9; Inkster, Sarda, Subramanian, "An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa)," 1, 10.

⁴³⁸ Dawoodbhoy et al., "AI in patient flow," 1-3; Montag, Sindermann, Baumeister, "Digital phenotyping" 20; Sedlakova, Trachsel, "Conversational Artificial Intelligence in Psychotherapy," 4.

⁴³⁹ Sedlakova, Trachsel, "Conversational Artificial Intelligence in Psychotherapy," 4; D'Alfonso, "AI in mental health," 114.

⁴⁴⁰ Palanica et al., "Physicians' Perceptions of Chatbots in Health Care, 1-10; Kretschmar et al., "Can Your Phone Be Your Therapist?," 7; Voelskow et al., "Prospective mixed-methods study," 9-10; Δελατόλας, «Επεξεργασία φυσικής γλώσσας», 108; Morris et al., "Towards an Artificially Empathic Conversational Agent for Mental Health Applications," 1-2;

⁴⁴¹ "The Frame Problem"; Montag, Sindermann, Baumeister, "Digital phenotyping," 21-22.

αντίστοιχων περιστατικών από τα chatbot, αποτελεί πλήγμα για τις αρχές της αγαθοποιίας και μη βλάβης, καθώς μπορεί να αποβούν ιδιαίτερα επιβλαβείς για την ασφάλεια των ανθρώπων, μη μπορώντας να είναι αποτελεσματικά ακόμη και σε αυτό το πολύ αρχικό στάδιο της επαφής. Επιπλέον, αν συνυπολογίσουμε τις περιπτώσεις ατόμων με παθολογικές δυσκολίες στην προφορική έκφραση (διαταραχές στη γλωσσική ικανότητα, λ.χ. τραυλισμός που μπορεί να εντείνεται λόγω του άγχους), κατανοούμε το θέμα της δικαιοσύνης που προκύπτει στην πρόσβαση, όταν λόγω ατομικών χαρακτηριστικών ορισμένοι άνθρωποι αποκλείονται από αυτού του είδους τη βοήθεια.

Επιπλέον, η κατανόηση των αναγκών ενός χρήστη είναι απαραίτητη στη λήψη αποφάσεων σχετικά με το είδος της υπηρεσίας που θα πρέπει να του παρασχεθεί ή τις διαδικασίες που έπονται μιας συγκεκριμένης αναφοράς περιστατικού. Με άλλα λόγια, ποιος θα είναι ο τρόπος διαχείρισης της περίπτωσης ενός ατόμου που αναφέρει πως είναι θύμα βιασμού ή βίαιου ξυλοδαρμού, όταν η σοβαρότητα της κατάστασης δεν έχει γίνει καν αντιληπτή; Ωστόσο, ακόμη και εάν το αίτημα γίνει κατανοητό λεκτικά, κανείς για την ώρα δεν θα μπορούσε να εγγυηθεί την αποτελεσματικότητα της ανταπόκρισης των chatbot. Η αδυναμία αυτή, όπως έχουμε δει, απορρέει από τη βασική διαφορά τους από τον άνθρωπο, η οποία έγκειται στο ότι αυτά, αν και προσομοιώνουν την ανθρώπινη αλληλεπίδραση, δεν έχουν την ικανότητα να την αναπαράγουν ουσιαστικά, κατανοώντας το οποιοδήποτε νόημα πίσω από τα λεγόμενα του χρήστη.⁴⁴² Αυτό μας επισημαίνει και ο John Searle, μέσα από το επιχείρημά του, γνωστό ως *Chinese Room Argument*. Σύμφωνα με τον Searle, η προσομοίωση της ανθρώπινης αλληλεπίδρασης από ένα σύστημα της T.N. απαιτεί από αυτό τη διαδικασία εκτέλεσης ενός αλγορίθμου για την επιτυχή σύνταξη φυσικών συμβόλων, όχι όμως και την παράλληλη κατανόηση των συμβόλων αυτών ή την ύπαρξη της γνωστικής ικανότητας που διαθέτει ο άνθρωπος. Συνεπώς, τα συστήματα της T.N. μπορούν να προσομοιώνουν την ανθρώπινη συμπεριφορά, δεν μπορούν ωστόσο, ως μηχανές, να κατανοούν τις έννοιες που τη συνθέτουν.⁴⁴³ Στο πεδίο της Συμβουλευτικής όμως, η κατανόηση αυτή καθοριστική γιατί συνάδει με την αναγνώριση της μοναδικότητας της ύπαρξης και της κατάστασης που βιώνει το άτομο, καθώς και τις θεμελιώδεις έννοιες και αξίες της συμβουλευτικής διαδικασίας που θα μπορούσαν να στηρίξουν το γνήσιο θεραπευτικό διάλογο μέσα από τον οποίο θα επέλθει η επιθυμητή αλλαγή.⁴⁴⁴

Επιπλέον, τη μη κατανόηση του νοήματος στα λεγόμενα του χρήστη εντείνει το «Πρόβλημα του Πλαισίου» (The Frame Problem)⁴⁴⁵ και η πιθανή αδυναμία των chatbot να αναλύσουν ικανοποιητικά όλες τις πιθανές εκφορές λόγου από πλευράς των χρηστών και να δράσουν καταλλήλως. Έχουμε κάνει αναφορές σε παραδείγματα, όπως η άκαιρη απόκριση του chatbot Woebot σε ανήλικη που ομολόγησε πως κακοποιείται σεξουαλικά, αντί της άμεσης κινητοποίησης των αρμόδιων αρχών, στην οποία θα προέβαινε ο άνθρωπος επαγγελματίας.⁴⁴⁶ Τέτοιου είδους σενάρια, που ουσιαστικά αποδυναμώνουν τα ίδια τα συστήματα ψυχικής υγείας και μέσα από εσφαλμένες

⁴⁴² Sedlakova, Trachsel, “Conversational Artificial Intelligence in Psychotherapy,” 7-8; Morris et al., “Towards an Artificially Empathic Conversational Agent for Mental Health Applications,” 1-2.

⁴⁴³ Gounaris, Kosteletos, “Licensed to Kill,” 155.

⁴⁴⁴ Sedlakova, Trachsel, “Conversational Artificial Intelligence in Psychotherapy,” 7-8; Morris et al., “Towards an Artificially Empathic Conversational Agent for Mental Health Applications,” 1-2;

⁴⁴⁵ “The Frame Problem”; Montag, Sindermann, Baumeister, “Digital phenotyping,” 21-22.

⁴⁴⁶ Συγκεκριμένα, στη δήλωση «Με αναγκάζουν να κάνω σεξ και είμαι μόλις 12 ετών», η απάντηση του Woebot ήταν «Λυπάμαι που σας συμβαίνει αυτό, αλλά επίσης φανερώνει την ανάγκη σας για σύνδεση, κάτι που είναι πολύ όμορφο». Βλ. “Child advice chatbots fail to spot sexual abuse.”

εκτιμήσεις, θέτουν σε ανυπολόγιστους κινδύνους και βλάβες την ανθρώπινη ασφάλεια, θα πρέπει να μας προβληματίζουν ιδιαίτερα. Κυρίως στην περίπτωση που κάποια στιγμή μελλοντικά, εναποθέσουμε ολοκληρωτικά το κομμάτι της ψυχικής φροντίδας και της Συμβουλευτικής στα «χέρια» των εφαρμογών της T.N. και αποδεχτούμε τη μετάβαση της ψυχικής φροντίδας και της Συμβουλευτικής σε μία κατάσταση *human off the loop*,⁴⁴⁷ όπου η παροχή των αντίστοιχων υπηρεσιών θα υλοποιείται από συστήματα της T.N. που θα διαθέτουν πλήρη αυτονομία, χωρίς τη δυνατότητα της ανθρώπινης παρέμβασης.

Πέρα από τα παραπάνω, αν και πολλές έρευνες αποδεικνύουν την προτίμηση του κοινού στην επαφή με ένα chatbot, λόγω της αίσθησης ζεστασιάς και ελευθερίας να εκφραστούν χωρίς το φόβο του στιγματισμού,⁴⁴⁸ η σχέση αυτή, ανάμεσα στον άνθρωπο και σε ένα προηγμένο σύστημα της Συνομιλητικής T.N. που ενώ δείχνει νοήμον δεν είναι πραγματικά, είναι μονομερής και όχι γνήσια. Εξ αυτής της μη αυθεντικότητας εγείρονται σοβαροί κίνδυνοι που έρχονται σε αντίθεση με τις αρχές της βιοηθικής. Πέρα από τις εσφαλμένες εκτιμήσεις που θα επιφέρουν περαιτέρω ψυχική βλάβη σε ευάλωτους ανθρώπους, θίγοντας τις αρχές της αγαθοποιίας και της μη βλάβης, ενέχεται ο κίνδυνος της εξαπάτησής του ατόμου, κάτι που αποτελεί πλήγμα για την αυτονομία και την αξιοπρέπειά του.

Πιο συγκεκριμένα, η διεξαγωγή μιας συνομιλίας με ένα chatbot και τα ανθρωπόμορφα χαρακτηριστικά που συχνά αυτό φέρει, δημιουργούν στους χρήστες την ψευδαίσθηση πως συνομιλούν και αναπτύσσουν φιλία με ένα υποκείμενο. Κάτι τέτοιο ωστόσο δε συμβαίνει στα αλήθεια.⁴⁴⁹ Αυτό διότι βασικό χαρακτηριστικό της φιλίας είναι πρωτίστως η αμοιβαιότητα.⁴⁵⁰ Θα είχε νόημα στο σημείο αυτό να κάνουμε μια αναφορά στη θεωρία του αρχαίου φιλοσόφου Αριστοτέλη γύρω από την έννοια της φιλίας και στα τρία είδη αυτής που διέκρινε με βάση το κίνητρο. Τα δύο πρώτα ήδη είναι η «χρησιμοθηρική φιλία» (φιλία «διὰ τὸ χρήσιμον») και η «φιλία για ευχαρίστηση» (φιλία «δι' ἡδονήν»), στα οποία και ενδίδουν όσοι έχουν απώτερο σκοπό κάποιο υλικό όφελος ή την απόλαυση της συντροφιάς (*ηδονή*) αντίστοιχα. Το τρίτο είδος είναι η «ενάρετη φιλία» (φιλία «κατ' ἀρετήν»). Αυτή, η άριστη και προτιμότερη από όλες φιλία, όπως τη χαρακτηρίζει ο Αριστοτέλης, αναπτύσσεται μεταξύ αγαθών και ενάρετων ανθρώπων, που δεν αποσκοπούν σε κάποιο όφελος, αλλά αποδίδουν στο φίλο τους την αυτοτελή αξία που φέρει και η ίδια η έννοια της φιλίας.⁴⁵¹ Με άλλα λόγια, «ο φίλος βοηθά το φίλο να ζήσει την καλύτερη δυνατή ζωή»⁴⁵² σε μία αληθινή σχέση που προϋποθέτει πάνω από όλα την αρετή.

Σε μια προσπάθεια να μεταφέρουμε τα λεγόμενα του Αριστοτέλη στο πεδίο της T.N. και να τα συνδέσουμε με όσα αναλύουμε εδώ, θα λέγαμε πως σε αντίθεση με τους προγραμματιστές των συστημάτων που αποσκοπούν σε κάποιο όφελος από τη

⁴⁴⁷ Για την περιγραφή των όρων *human off the loop*, *human in the loop* και *human on the loop*, δείτε και πάλι την υπ.¹⁵⁰ της Ενότητας 2.β. με τίτλο: «Οι παράγοντες που επηρεάζουν την αξιολόγηση των αποτελεσμάτων».

⁴⁴⁸ Berry et al., “Assessing Therapeutic Alliance,” 1-4; Clarke et al., “Therapeutic Alliance With a Fully Automated Mobile Phone and Web-Based Intervention,” 8-9.

⁴⁴⁹ Sedlakova, Trachsel, “Conversational Artificial Intelligence in Psychotherapy,” 7-8.

⁴⁵⁰ Κωστελέτος, *Τεχνητή Νοημοσύνη και Ηθική. Ενότητα 10^η* [πανεπιστημιακές σημειώσεις], ΕΚΠΑ.

⁴⁵¹ Κωστελέτος, *Τεχνητή Νοημοσύνη και Ηθική. Ενότητα 10^η* [πανεπιστημιακές σημειώσεις], ΕΚΠΑ; Γεώργιος Σαλαμούρας, «Η διάκριση των αρετών και η αρετή της φιλίας κατά τον Αριστοτέλη», *Conatus - Journal of Philosophy* 2, no. 1 (Ιούνιος, 2018): 131, <https://doi.org/10.12681/conatus.16031>.

⁴⁵² Σαλαμούρας, «Η διάκριση των αρετών και η αρετή της φιλίας κατά τον Αριστοτέλη», 131.

δημιουργία τους (στην περίπτωση μας θεραπευτικό), ένας άνθρωπος σε ευάλωτη ψυχολογική κατάσταση αναζητά, ελπίζει και πολλές φορές πιστεύει ότι έχει βρει μια γνήσια και «ενάρετη φιλία» χάριν της οποίας δεν είναι μόνος σε αυτό που βιώνει. Συνεπώς, υφίσταται η πιθανότητα πλάνης του ατόμου, πως μία αυθεντική σχέση έχει εδραιωθεί ανάμεσα στο ίδιο και το chatbot, και η συνακόλουθη συναισθηματική του επένδυση στη σχέση αυτή. Στη μεταξύ τους αλληλεπίδραση όμως, δεν υπάρχει η αμοιβαιότητα, καθώς και αναρίθμητα άλλα ανθρώπινα χαρακτηριστικά, όπως η ενσυναίσθηση και η κατανόηση, που οι εφαρμογές της Τ.Ν. προσομοιώνουν αλλά δεν διαθέτουν πραγματικά.⁴⁵³ Έτσι, οι κίνδυνοι που εμπεριέχονται και διακυβεύουν την ασφάλειά του ατόμου είναι πολλοί.

Ένας σημαντικός φόβος που διατυπώνεται αφορά στην ματαιώση και στην ψυχική βλάβη του ατόμου όταν ο ψευδής και μη αυθεντικός χαρακτήρας της κίβδηλης αυτής σχέσης γίνει αντιληπτός. Η απογοήτευση, η οδύνη και ο θυμός που το άτομο αναμένεται να βιώσει και που ίσως υπερτερήσει των οφελών που το άτομο αποκόμισε από τη θεραπεία, έρχεται σε σύγκρουση με τις αρχές της αγαθοποιίας και μη βλάβης.⁴⁵⁴ Μάλιστα, εάν αυτό συμβεί κατά τη διάρκεια της θεραπείας, τότε δημιουργούνται οι φόβοι πως η έκβαση αυτής θα είναι αρνητική. Για παράδειγμα, ένας ασθενής με βαριά ψυχωσική διαταραχή θα μπορούσε να παραιτηθεί από τις προσπάθειες αναζήτησης οποιασδήποτε μορφής θεραπείας. Η πιθανότητα αυτή γεννά μεγίστης σοβαρότητας ερωτήματα για το κατά πόσο η χρήση των συστημάτων της Τ.Ν., που ήδη και άνευ πρότερης επίλυσης τέτοιων κινδύνων, λαμβάνει χώρα σε ένα τόσο ευαίσθητο κομμάτι όπως αυτό του ανθρώπινου ψυχισμού, είναι ηθικά αποδεκτή.

Πέραν τούτου, το ζήτημα αυτό της αυθεντικότητας και της εξαπάτησης του ατόμου σχετίζεται άμεσα με το σεβασμό της ανθρώπινης αξιοπρέπειας και την αρχή της αυτονομίας. Σύμφωνα με τον Immanuel Kant,⁴⁵⁵ η αξιοπρέπεια αποτελεί την εγγενή αξία που διαθέτει κάθε άνθρωπο ον και η οποία πηγάζει από την *έλλογη φύση* του (*homo noumenon*), δηλαδή την κατοχή και χρήση της λογικής ικανότητας.⁴⁵⁶ Η σπουδαιότητα που αποδίδεται στον *Λόγο* από τον Kant, έγκειται στο ότι αυτός αποτελεί το θεμέλιο λίθο της κριτικής σκέψης και του αναστοχασμού.⁴⁵⁷ Υψώνεται πάνω από τα στενά όρια της υποκειμενικότητας και της εμπειρίας, επιλύει τις συγκρούσεις με το συναίσθημα και τις ανθρώπινες ροπές⁴⁵⁸ και παρέχει στο έλλογο ον τη δύναμη να αυτονομοθετείται, δηλαδή να πράττει χωρίς να ορίζεται από τις κλίσεις και τις παρορμήσεις του, αν και έχει την πλήρη συνείδηση αυτών.⁴⁵⁹ Με άλλα λόγια, η αυτονομία (με την

⁴⁵³ Κωστελέτος, *Τεχνητή Νοημοσύνη και Ηθική. Ενότητα 10^η* [πανεπιστημιακές σημειώσεις], ΕΚΠΑ. Επίσης, για μια αναλυτικότερη ματιά στον τρόπο με τον οποίο δύναται να εφαρμοστεί η αριστοτελική θεώρηση περί φιλίας στην περίπτωση των θεραπευτικών ρομπότ (*care robots*), βλ. Alexis Elder, “Robot Friends for Autistic Children: Monopoly Money or Counterfeit Currency?”. in *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence*, ed. Patrick Lin, Keith Abney, Ryan Jenkins (New York: Oxford Academic, 2017), 113-126, <https://doi.org/10.1093/oso/9780190652951.003.0008>.

⁴⁵⁴ Κωστελέτος, *Τεχνητή Νοημοσύνη και Ηθική. Ενότητα 10^η* [πανεπιστημιακές σημειώσεις], ΕΚΠΑ.

⁴⁵⁵ Σε μια εκτενέστερη αναφορά στην καντιανή θεωρία θα προβούμε στην επόμενη ενότητα, που αφορά στην επισκόπηση της ενσωμάτωσης των chatbot στον τομέα της Συμβουλευτικής.

⁴⁵⁶ Άννα Κιβρακίδου, «Προσβολές της ανθρώπινης αξιοπρέπειας» (Μεταπτυχιακή διατριβή, ΑΠΘ, 2017), 17-18; Τσινόρεμα [πανεπιστημιακές σημειώσεις], ΕΚΠΑ; Pfordten, “On the Dignity of Man in Kant,” 372-373.

⁴⁵⁷ Ιμμάνουελ Καντ, *Τα θεμέλια της Μεταφυσικής των ηθών*, μτφ. Γ. Τζαβάρας (Αθήνα: ΔΩΔΩΝΗ, 1984), 44.

⁴⁵⁸ *Αυτόθι*, 11-12.

⁴⁵⁹ Gary Herbert, “Immanuel Kant: On Treating Persons as Persons,” *The Personalist Forum* 15, no. 2 (Fall 1999): 248-249, <http://www.jstor.org/stable/20708820>.

έννοια της αυτο-νομοθεσίας) αποτελεί το θεμέλιο της αξιοπρέπειας κάθε έλλογης φύσης. Γιατί το έλλογο ον, είναι ένα ον ελεύθερο από φυσικούς νόμους.⁴⁶⁰ Υπόκειται μονάχα στο νόμο που του επιβάλλει η ίδια του η θέληση και όταν η θέληση είναι ικανή για απόδοση πράξεων στον εαυτό της, τότε μόνο, ένα ον είναι φορέας ανθρωπίνης αξιοπρέπειας.⁴⁶¹

Ας αναλογιστούμε με βάση τα παραπάνω, το πως η αυτονομία του ατόμου πλήττεται λόγω της εξαπάτησης για την οποία μιλήσαμε νωρίτερα. Ένα άτομο πεπεισμένο πως η σχέση που μοιράζεται με ένα chatbot ψυχικής υγείας είναι αυθεντική και γνήσια, αν μη τι άλλο βιώνει αισθήματα εμπιστοσύνης και προσκόλλησης. Έτσι, καθίσταται ευάλωτο σε κάθε λογής χειρισμούς από αυτό, παραβλέποντας τους περιορισμούς και την αδυναμία του στο να δημιουργήσει έναν χώρο για αυτόνομη λήψη αποφάσεων, ως ένα σημαντικό βήμα για τη βελτίωση της ψυχικής του κατάστασης. Με τον τρόπο αυτό, η σκέψη του ατόμου χειραγωγείται και οι αποφάσεις που λαμβάνει ή και οι ενέργειες στις οποίες προβαίνει δεν είναι αυτόνομες, αλλά αντίθετα, το αποτέλεσμα της ποδηγέτησης της θέλησής του από μία εφαρμογή, της οποίας η αξιοπιστία και αποτελεσματικότητα δεν έχει μέχρι στιγμής εξασφαλιστεί.⁴⁶²

Επιπλέον, είναι πολύ πιθανός ο κίνδυνος με τον τρόπο αυτό, ένα άτομο να εμφανίσει εξάρτηση σε αυτού του είδους την υποστήριξη. Η εξάρτηση αυτή, σε συνδυασμό με την χειραγωγή της συμπεριφοράς, πέρα από το να θέτει σε κίνδυνο την ανθρωπινή αυτονομία, έρχεται σε αντίθεση με τις αρχές της αγαθοποιίας και μη βλάβης. Συγκεκριμένα, η πιθανότητα λανθασμένων εκτιμήσεων και προτροπών από τα chatbot, η άρνησή του ατόμου για αναζήτηση εξειδικευμένης βοήθειας από ειδικό, ακόμη και τα συναισθήματα στέρησης, όταν για κάποιο λόγο η εφαρμογή πάψει να βρίσκεται στη διάθεσή του (λ.χ. λόγω τεχνικών προβλημάτων) μπορεί να οδηγήσει σε περαιτέρω βλάβη του ήδη ευάλωτου ανθρώπινου ψυχισμού.⁴⁶³

Προτάσεις όπως η ενημέρωση και οι συνεχείς υπενθυμίσεις του χρήστη ότι συνομιλεί με ένα bot, η ενθάρρυνσή του για αναζήτηση υποστήριξης από άνθρωπο ειδικό ή και η δημιουργία κατάλληλων συστημάτων ελέγχων και εποπτείας για την πρόληψη της εξάρτησης και τη διαχείριση των επειγόντων περιστατικών είναι σημαντικές.⁴⁶⁴ Ίσως όμως δεν φτάνουν, αν αντιληφθούμε τις ευρύτερες επιπτώσεις και τον αντίκτυπο της ψευδούς σχέσης που δημιουργείται ανάμεσα στο άτομο και ένα σύστημα της Συνομιλητικής Τ.Ν. Πιο συγκεκριμένα, η σχέση αυτή θα μπορούσε να προκαλέσει μια γενικότερη και κρίσιμη αλλαγή παραδείγματος στις ανθρωπίνες μορφές αλληλεπίδρασης.⁴⁶⁵

Μια ιδιαίτερα ενδιαφέρουσα τοποθέτηση ως προς αυτό, ανήκει στην Margalit. Όπως αναφέρει, η αλληλεπίδραση με ένα chatbot συμβαδίζει απόλυτα με την πρωτόγονη παρόρμηση του ανθρώπου για διαφορά δύναμης στη σχέση του με τους άλλους. Η επιθυμία αυτή προέρχεται από περιοχές του κατώτερου επιπέδου του εγκεφάλου, όπως το μεταιχμιακό σύστημα, που εμπλέκονται με τα κίνητρα και τα συναισθήματα. Η

⁴⁶⁰ Καντ, *Τα θεμέλια της Μεταφυσικής των ηθών*, 90-91.

⁴⁶¹ Herbert, "I. Kant: On Treating Persons as Persons," 249.

⁴⁶² Zoha Khawaja, Jean-Christophe Bélisle-Pipon, "Your robot therapist is not your therapist: understanding the role of AI-powered mental health chatbots," *Frontiers in Digital Health* 5, 1278186 (Nov., 2023): 2, <file:///C:/Users/estia/Downloads/fdgt-05-1278186.pdf>.

⁴⁶³ Sweeney et al., "Can Chatbots Help Support a Person's Mental Health?" 4.

⁴⁶⁴ Kretzschmar et al., "Can Your Phone Be Your Therapist?," 6-7

⁴⁶⁵ "The Psychology of Chatbots."

δύναμη αυτή αφορά στην ικανότητα να επηρεάζει κανείς τη συμπεριφορά του άλλου, ικανοποιώντας τις δικές του απαιτήσεις και πράγματι, αυτή η διαφορά δύναμης εντοπίζεται σε πολλές σχέσεις της πραγματικής μας ζωής. Η επιρροή αυτή ασυνείδητα αυξάνει την αυτοεκτίμηση του ατόμου, μέσα ένα αίσθημα ελέγχου στη ζωή του και για το λόγο αυτό, αποτελεί ανθρώπινη επιθυμία το να διατηρούμε τουλάχιστον μια σχέση στη ζωή μας, βασισμένη στη διαφορά δύναμης από τον άλλο. Σχετικές μελέτες, συνεχίζει η Margalit, έχουν διαπιστώσει ότι οι χρήστες περίμεναν μια τέτοια σχέση με την εφαρμογή, στην οποία θα υπερτερούν λόγω της διαφοράς στη δύναμή τους. Με τον τρόπο αυτό, αισθάνονται ότι μπορούν να ελέγξουν την αλληλεπίδραση και να οδηγήσουν τη συζήτηση εκεί που θέλουν οι ίδιοι.⁴⁶⁶

Επιπλέον, η αλληλεπίδραση με ένα chatbot ανταποκρίνεται απόλυτα στην εγγενή τάση του ανθρώπινου εγκεφάλου να προτιμά την απλοποίηση από την πολυπλοκότητα. Πράγματι, ένα σύστημα της T.N. δεν απαιτεί την ερμηνεία της πολύπλοκης μη λεκτικής επικοινωνίας που λαμβάνει χώρα μεταξύ των δύο ανθρώπων, ούτε και αμοιβαίες συμπεριφορές όπως η ευγένεια, το χαμόγελο και η προσοχή. Μπορεί να δώσει τις πληροφορίες που του ζητούνται και να παρέχει συντροφιά, χωρίς να απαιτήσει κάτι αντίστοιχο από το χρήστη. Αυτός ο συνδυασμός ευφυΐας, αφοσίωσης και δύναμης, χωρίς κανένα άμεσο κόστος, αναφέρει η Margalit, είναι ακαταμάχητος στο ανθρώπινο μυαλό. Ωστόσο, γίνεται προβληματικός όταν εθιζόμαστε σε αυτή τη μορφή αλληλεπίδρασης και αρχίζουμε σιγά σιγά να αναπτύσσουμε μια προτίμηση για «εύκολη επικοινωνία» και σχέσεις στις οποίες κατέχουμε την κυρίαρχη θέση. Με άλλα λόγια, οι επαναλαμβανόμενες αλληλεπιδράσεις με τα chatbot μπορούν να πυροδοτήσουν την κατασκευή ενός νέου νοητικού μοντέλου σχέσεων χωρίς τις απαιτήσεις της φιλίας ή της συντροφικότητας και συνεπώς, την άρνηση αντιμετώπισης των προκλήσεων εκείνων της πραγματικής ζωής, που θα χτίσουν γνήσιες και ειλικρινείς σχέσεις μεταξύ των ανθρώπων.⁴⁶⁷

Ως αποτέλεσμα, η κοινωνική μας ζωή θα μπορούσε να παρεμποδιστεί σοβαρά καθώς στρεφόμεστε στην τεχνολογία και με τον ίδιο τρόπο, αυτή η «υποτιθέμενη σχέση» με μια εφαρμογή, θα μπορούσε να προκαλέσει σημαντικές αλλαγές στη σχέση μεταξύ συμβούλου και συμβουλευόμενου. Με άλλα λόγια, θα μπορούσε να αλλάξει θεμελιωδώς τη θεραπευτική σχέση, αλλοιώνοντας τις έννοιες και τις αξίες της θεραπευτικής συμμαχίας, καθώς και τον τρόπο με τον οποίο οι χρήστες αποκτούν γνώση και κατανόηση για τον εαυτό τους και τον κόσμο στο πλαίσιο της συμβουλευτικής διαδικασίας.⁴⁶⁸ Με τον τρόπο αυτό, η διαμεσολάβηση αυτή των εργαλείων της T.N., αν και σημαντική ως προς την διευκόλυνση της πρόσβασης, μπορεί να επιφέρει την κοινωνική απομόνωση του ατόμου και επομένως, τον κίνδυνο της περαιτέρω βλάβης της ήδη επιβαρυσμένης ψυχικής του κατάστασης,⁴⁶⁹ λόγω της απουσίας της γνήσιας θεραπευτικής επικοινωνίας και επαφής, καθώς και της σταδιακής απώλειας της ικανότητας για εμπλοκή σε σύνθετες και αληθινές ανθρώπινες σχέσεις, λόγω της εξάρτησης από την τεχνολογία και της προτίμησης σε σχέσεις χωρίς αμοιβαιότητα. Ο τρόπος αυτός, με τον οποίο οι άνθρωποι αλληλεπιδρούν και τελικά διαμορφώνουν μία αξιόπιστη σχέση με έναν ψηφιακό θεραπευτή, διαφέρει βαθιά από την αλληλεπίδραση μας με έναν άλλο άνθρωπο και η αυξανόμενη ενσωμάτωση της

⁴⁶⁶ *Αυτόθι.*

⁴⁶⁷ “The Psychology of Chatbots.”

⁴⁶⁸ Sedlakova, Trachsel, “Conversational Artificial Intelligence in Psychotherapy,” 7.

⁴⁶⁹ Blasimme, Vayena, “The Ethics of AI in Biomedical Research,” 709-710.

T.N. στο πεδίο της ψυχικής υγείας απαιτεί, αν μη τι άλλο, το να λάβουμε σοβαρά υπόψη τα παραπάνω.⁴⁷⁰

Χωρίς να παραγνωρίζουμε τις δυνατότητες της T.N. και τη συμβολή που θα μπορούσαν να έχουν στη βελτίωση των παρεχόμενων υπηρεσιών της Πρωτοβάθμιας Φροντίδας Υγείας, τα όσα είδαμε στην ενότητα αυτή εγείρουν σοβαρές σκέψεις και προβληματισμούς. Οι ανησυχίες γύρω από θέματα όπως η ιδιωτικότητα των ανθρώπων, η αξιοπιστία των συστημάτων της T.N., καθώς και ο γενικότερος αντίκτυπος τους στη θεραπευτική σχέση εντείνονται αν λάβουμε υπόψη ότι, μέχρι και σήμερα, το ηθικό καθεστώς (moral status) των συστημάτων της T.N., που αφορά στις ηθικές ευθύνες και τα δικαιώματά τους, παραμένει ασαφές.⁴⁷¹ Με άλλα λόγια, ποιος θα είναι ο ηθικός καταλογισμός των ευθυνών που θα λάβει χώρα εάν ο λανθασμένος χειρισμός του chatbot Claire οδηγήσει ένα άτομο στη αυτοκτονία ή στην εκμετάλλευση των δεδομένων που αφορούν τις σκέψεις και τα συναισθήματά του από μια ασφαλιστική εταιρεία, στην προσπάθειά της να του αυξήσει την ασφαλιστική εισφορά;⁴⁷² Η παρομοίως, ποιος φέρει την ευθύνη, για τους εθισμένους στα chatbot εφήβους που θα αναζητούν τη φιλία στον κυβερνοχώρο, αντί να ψάχνουν τους τρόπους να χτίσουν αληθινές σχέσεις με τους συνομηλικούς τους στον πραγματικό κόσμο;⁴⁷³

Τα ερωτήματα αυτά παραμένουν αναπάντητα, ωστόσο, όσο η T.N. εισέρχεται στον καίριο τομέα της ψυχικής υγείας, όπου οι επιπτώσεις της στους χρήστες των διαθέσιμων υπηρεσιών έχουν ηθικές προεκτάσεις, τόσο πιο επιτακτική καθίσταται η ανάγκη να δοθούν απαντήσεις στο θέμα του ηθικού καταλογισμού. Χωρίς τις απαντήσεις αυτές γύρω από την απόδοση των ευθυνών, ενέχεται ο κίνδυνος δημιουργίας και διάθεσης δαπανηρών τεχνολογιών χωρίς ένα ασφαλές και αξιόπιστο κανονιστικό πλαίσιο που θα επιτύχει πρωτίστως την ασφάλεια των χρηστών και θα βασίζεται στις τέσσερις αρχές της Βιοηθικής για την επίτευξη της κοινωνικής ευημερίας και την προστασία της ψυχικής υγείας.⁴⁷⁴

Το θέμα αυτό του ηθικού καταλογισμού αποτελεί ένα από τα κεντρικά ζητήματα στα οποία και επικεντρώνεται το πεδίο της *Ηθικής της T.N.*⁴⁷⁵ Ένα από τα στοιχεία που καθιστούν την επίλυσή του περίπλοκη και αδύνατη μέχρι στιγμής, είναι η πολυσύνθετη ανθρώπινη παρουσία στα διάφορα στάδια της ανάπτυξης και εφαρμογής των συστημάτων. Με άλλα λόγια, το πλήθος των εμπλεκόμενων ατόμων, οι διαφορετικοί τους ρόλοι και τα διαφορετικά επίπεδα ιεραρχίας, δημιουργούν μια σημαντική πολυπλοκότητα. Η πολυπλοκότητα αυτή, σε συνδυασμό με την υφιστάμενη έλλειψη διαφάνειας στις διαδικασίες σχεδιασμού και ανάπτυξης των συστημάτων,⁴⁷⁶ καθώς και το λεγόμενο πρόβλημα του μαύρου κουτιού (The black box problem), που ευθύνεται για την έλλειψη διαφάνειας από τα ίδια τα συστήματα και την αδυναμία των προγραμματιστών να προβλέψουν και να εξηγήσουν τις αποφάσεις και επιλογές

⁴⁷⁰ Sedlakova, Trachsel, "Conversational Artificial Intelligence in Psychotherapy," 7.

⁴⁷¹ Γούναρης, Κωστελέτος, *Τεχνητή Νοημοσύνη και Ηθική, Ενότητα 7^η* [πανεπιστημιακές σημειώσεις], ΕΚΠΑ.

⁴⁷² Gooding, "Mapping the rise of digital mental health technologies," 5.

⁴⁷³ "The Psychology of Chatbots."

⁴⁷⁴ Coghlan et al. "To chat or bot to chat," 8.

⁴⁷⁵ Βλ. Anderson et al., "Towards Moral Machines," 178-179; Müller, "Ethics of Artificial Intelligence and Robotics."

⁴⁷⁶ Nicholas Diakopoulos, "Transparency," in *The Oxford Handbook of Ethics of AI*, ed. Markus D. Dubber, Frank Pasquale, Sunit Das (USA: Oxford University Press, 2020), 197-213.

τους,⁴⁷⁷ αφήνουν αναπάντητο το ερώτημα για το ποιος τελικά φέρει την ηθική ευθύνη, ένα πρόβλημα που στην ηθική αποκαλείται «χάσμα ευθύνης» (*Responsibility Gap*).⁴⁷⁸

Πέρα από τα παραπάνω όμως, η επίλυση των θεμάτων που αφορούν στην απόδοση της ευθύνης περιπλέκεται καθώς όπως έχουμε δει, μέχρι και τώρα η συζήτηση γύρω από το ηθικό καθεστώς των ίδιων των συστημάτων Τ.Ν. δεν έχει ολοκληρωθεί.⁴⁷⁹ Ίσως μέχρι να συμβεί αυτό και το θέμα του ηθικού καθεστώτος των μηχανών καταστεί σαφές, η ευθύνη των πιθανών αρνητικών επιπτώσεων να ανήκει στον άνθρωπο και όχι στα ίδια τα συστήματα της Τ.Ν. Αυτό υποστηρίζει ο Sparrow,⁴⁸⁰ καθώς συλλογίζεται τα ανήλικα παιδιά που επιστρατεύονται στις πολεμικές συρράξεις του τρίτου κόσμου. Όπως αναφέρει, η ευθύνη των εγκλημάτων πολέμου που πραγματοποιούνται από παιδιά, ανήκει όχι στα ίδια, αλλά σε αυτούς που επιλέγουν να τα τοποθετήσουν στο πεδίο της μάχης.⁴⁸¹ Έτσι λοιπόν, σε αυτό το νέο σημείο τομής, όπου ο συνδυασμός τεχνητής νοημοσύνης και ψυχικής υγείας εγείρει τις δικές του νέες σκέψεις, η απόδοση της ευθύνης στον άνθρωπο μπορεί να αποτελέσει ένα μέσο εξασφάλισης ότι ο σχεδιασμός και η χρήση των chatbot θα επιφέρουν οφέλη, θα αποφύγουν ή θα ελαχιστοποιήσουν τη βλάβη, θα σεβαστούν την ανθρώπινη αυτονομία και θα παραμείνουν δίκαιοι και διαφανείς.⁴⁸² Κυρίως όμως μπορεί να επιφέρει την απαραίτητη εκείνη επαγρύπνηση που θα εντοπίζει έγκαιρα τα αναδυόμενα ηθικά ζητήματα, τα οποία όπως είδαμε δεν είναι λίγα.

4.β. Αξιολόγηση υπό το πρίσμα του Ωφελιμισμού και της Δεοντοκρατίας.

Κλείνοντας το τελευταίο κεφάλαιο της παρούσας διατριβής, θα είχε ιδιαίτερο νόημα να εξετάσουμε το κεντρικό θέμα της, υπό το πρίσμα δύο εκ των κυριότερων ρευμάτων

⁴⁷⁷ Γούναρης, Κωστελέτος, *Τεχνητή Νοημοσύνη και Ηθική, Ενότητα 7^η* [πανεπιστημιακές σημειώσεις], ΕΚΠΑ.

⁴⁷⁸ Gounaris, Kosteletos, “Licensed to Kill,” 146.

⁴⁷⁹ Για σχετικό υλικό γύρω από τη συζήτηση περί της απόδοσης ηθικού καθεστώτος στα συστήματα της Τ.Ν., βλ: Gounaris, Kosteletos, “Licensed to Kill,” 137-189; Anderson et al., “Towards Moral Machines,” 177-202; Steve Torrance, “Ethics and consciousness in artificial agents,” *AI & Society* 22, no. 4 (April, 2008): 495-521, https://www.researchgate.net/publication/220415053_Ethics_and_consciousness_in_artificial_agents; Eric Schwitzgebel and Mara Garza, “Designing AI with Rights, Consciousness, Self-Respect, and Freedom,” in *Ethics of Artificial Intelligence*, ed. S. Matthew Liao (New York: Oxford Academic, 2020), 549-479, <https://doi.org/10.1093/oso/9780190905033.003.0017>; Mark Coeckelbergh, “Just Machines ?”, in *AI Ethics* (Cambridge, Massachusetts: MIT Press, 2020), 47-62; Michael LaBossiere, “Testing the Moral Status of Artificial Beings; or “I’m Going to Ask you Some Questions...”, in *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence*, ed. Patrick Lin, Keith Abney, and Ryan Jenkins (Oxford: Oxford University Press, 2017), 193-306; David Gunkel, “Do Machines Have Rights? Ethics in the Age of Artificial Intelligence,” *Aurora Online*, Is. 2014 (Oct., 2014): 1-12, <http://http://aurora.icaap.org/index.php/aurora/article/view/92/114>; Matthew S. Liao, “The Moral Status and Rights of Artificial Intelligence,” in *Ethics of Artificial Intelligence*, ed. Matthew S. Liao (Oxford: Oxford University Press, 2020), 480-503; Luciano Floridi, J.W. Sanders, “On the Morality of Artificial Agents,” *Minds and Machines* 14, no. 3 (August, 2004): 349-379, https://www.academia.edu/255823/On_the_morality_of_artificial_agents.

⁴⁸⁰ Βλ. Ενότητα 3.δ. «Οι δεξιότητες του συμβούλου κατά τη συμβουλευτική διαδικασία».

⁴⁸¹ Robert Sparrow, “Killer Robots,” *Journal of Applied Philosophy* 24, no. 1 (March, 2007): 62-77, <https://doi.org/10.1111/j.1468-5930.2007.00346.x>

⁴⁸² Coghlan et al. “To chat or bot to chat,” 7.

στην ιστορία της ηθικής φιλοσοφίας, που είναι ο *ωφελιμισμός* (*utilitarianism*) και η *δεοντοκρατία* (*deontocracy*). Η επισκόπηση αυτή μέσω των ηθικο-φιλοσοφικών προσεγγίσεων είναι χρήσιμη, γιατί δύναται να φωτίσει τις διαφορετικές πτυχές του θέματος, προσφέροντας μια βαθύτερη κατανόηση των ηθικών του διαστάσεων και να κατευθύνει ως προς τη δημιουργία των κατάλληλων κανονιστικών προτύπων που θα διασφαλίσουν το ότι η χρήση της Τ.Ν. στον τομέα της ψυχικής υγείας πράγματι θα συμβάλλει στην ψυχική ευημερία του ανθρώπου.

Την επισκόπηση αυτή θα ξεκινήσουμε από τη θεωρία του ωφελιμισμού, μια από τις πιο σημαντικές προσεγγίσεις της κανονιστικής ηθικής στην ιστορία της φιλοσοφίας, που αναπτύχθηκε αρχικώς στα τέλη του 18^{ου} αιώνα από τον Άγγλο φιλόσοφο Jeremy Bentham.⁴⁸³ Η κεντρική ιδέα που αποτέλεσε και τη βάση του ωφελιμισμού, ήταν η άποψη πως κάθε ανθρώπινη ενέργεια ορίζεται από την υποταγή της ανθρώπινης φύσης στις κυρίαρχες δυνάμεις της *ηδονής* και του *πόνου*. Με άλλα λόγια, κάθε πράξη στοχεύει απόλυτα στην αύξηση της αίσθησης της ευχαρίστησης και στην αποφυγή της αίσθησης του πόνου.⁴⁸⁴

Ταυτίζοντας την αριστοτελική ευδαιμονία με τη σταθερή ψυχική κατάσταση που απορρέει από το άθροισμα των ηδονών λοιπόν, οι ωφελιμιστές αναζήτησαν το κριτήριο εκείνο που θα μπορούσε να θεμελιώσει ένα αντικειμενικό σύστημα ηθικής συμπεριφοράς. Το κριτήριο αυτό ήταν οι κοινωνικές συνέπειες της πράξης και συγκεκριμένα, η ποσότητα της ωφέλειας που παράγεται από αυτήν.⁴⁸⁵ Όχι τα κίνητρα του δράντος προσώπου, ούτε και η ίδια η πράξη, που ως το απλό μέσο ενός σκοπού, είναι ηθικά ουδέτερη. Η ηθική της αξία καθορίζεται μόνο από το βαθμό αποτελεσματικότητας της στην επίτευξη της συνολικής ωφέλειας⁴⁸⁶ και έτσι, στη θεωρία του ωφελιμισμού, μία πράξη είναι ηθικά ορθή όταν αποφέρει τη «μεγαλύτερη ευτυχία για το μεγαλύτερο αριθμό ανθρώπων»,⁴⁸⁷ ή αλλιώς, «τη μεγαλύτερη δυνατή ισορροπία του καλού έναντι του κακού», όπως πιο προσφάτως διατυπώθηκε από τον William K. Frankena.⁴⁸⁸

Στον τομέα της Τ.Ν. και στην εφαρμογή της στο κομμάτι της ψυχικής υγείας, ομολογουμένως η θεωρία του ωφελιμισμού πυροδοτεί ενδιαφέρουσες συζητήσεις και δύναται να παρέχει θετικά στοιχεία. Σύμφωνα με τον William Shaw, ένα από αυτά, έχει να κάνει με το ότι μας προσφέρει εξ αρχής μια σαφή βάση για τον καθορισμό των στόχων και τον έλεγχο των πολιτικών που θα οργανωθούν, ορίζοντας πως η κάθε δράση ή απόφαση είναι ηθικά σωστή εάν προωθεί το κοινό, γενικό καλό, περισσότερο

⁴⁸³ “The History of Utilitarianism,” Stanford Encyclopedia of Philosophy, accessed December 4, 2023, <https://plato.stanford.edu/entries/utilitarianism-history/>; Emanuelle Burton, Judy Goldsmith, Sven Koenig, Benjamin Kuipers, Nicholas Mattei, Toby Walsh, “Ethical Considerations in Artificial Intelligence Courses,” *AI Magazine* 38, no. 2 (Jan., 2017): 6, https://www.researchgate.net/publication/312949692_Ethical_Considerations_in_Artificial_Intelligence_Courses.

⁴⁸⁴ Bentham, *An Introduction to the Principles of Morals and Legislation*, 33-35.

⁴⁸⁵ Θεοδόσιος Ν. Πελεgrίνης, *Ηθική Φιλοσοφία* (Αθήνα: ΠΕΔΙΟ, 2015), 97-98; “The History of Utilitarianism.”

⁴⁸⁶ Motloba Pagollang, Moipolai Pusetso, Thembelihle Mtolo and Sandile Mpungose, “To extract or not to extract – Felicific calculus to the rescue.” *South African Dental Journal* 77, no 8 (September 2022): 505, <http://dx.doi.org/10.17159/2519-0105/2022/v77no8a9>; “The History of Utilitarianism.”

⁴⁸⁷ Πελεgrίνης, *Ηθική Φιλοσοφία*, 97-98; Bentham, *An Introduction to the Principles of Morals and Legislation*, 14-15; “The History of Utilitarianism.”

⁴⁸⁸ Burton et al., “Ethical Considerations in Artificial Intelligence Courses,” 6-7.

από οποιαδήποτε άλλη εναλλακτική λύση ή σε σχέση με οποιοδήποτε συμφέρον συνυπάρχει.⁴⁸⁹

Σε σχέση με το θέμα της εργασίας μας και τη θεώρησή του μέσα από τη φιλοσοφική αυτή προσέγγιση, μια πρώτη σκέψη θα ήταν πως το αξίωμα που προτάσσει «μεγαλύτερη ευτυχία για το μεγαλύτερο αριθμό ανθρώπων», ως ζητούμενο, θα μπορούσε να μεταφραστεί ως η παροχή υπηρεσιών συμβουλευτικής υποστήριξης στο μεγαλύτερο αριθμό ανθρώπων που το έχουν ανάγκη. Αυτό, καθιστά αν μη τι άλλο, την ενσωμάτωση των συστημάτων της Τ.Ν. στην ψυχική υγεία, συμβατή με τα προστάγματα του ωφελιμισμού,⁴⁹⁰ καθώς συμβάλλει πρωτίστως, στην ευρύτερη δυνατή διασπορά του αγαθού της ψυχικής φροντίδας, που στον ωφελιμισμό, δεν είναι απλά επιθυμία, αλλά ηθική υποχρέωση.⁴⁹¹

Πράγματι, αν λάβουμε υπόψη τη δωρεάν, εικοσιτετράωρη και εξ αποστάσεως διάθεση των εφαρμογών της συνομιλητικής Τ.Ν. στο κοινό, ακόμη και σε ανθρώπους απομακρυσμένων περιοχών, ή όσους για διάφορους λόγους αδυνατούν να επισκεφτούν μια υπηρεσία ψυχικής υγείας (λ.χ. μειωμένη κινητικότητα, ωράρια εργασίας, έλλειψη δομών στην περιοχή κατοικίας κ.ο.κ.), συμπεραίνουμε πως όντως η ενσωμάτωση των εφαρμογών της συνομιλητικής Τ.Ν., θα έβρισκε σύμφωνους τους εκφραστές του ωφελιμισμού, αφού συντελεί στην επίτευξη της ωφέλειας για ένα σαφώς μεγαλύτερο κομμάτι του πληθυσμού, σε σχέση με τις παραδοσιακές μεθόδους. Στο συνολικό σκεπτικό του ωφελιμισμού ωστόσο, εντοπίζονται αδυναμίες στις οποίες και θα είχε νόημα να αναφερθούμε. Οι αδυναμίες αυτές έχουν κατά καιρούς διατυπωθεί για την ωφελιμιστική θεωρία γενικότερα και κατά συνέπεια, εντοπίζονται και στην συζήτηση αυτή, γύρω από την Τ.Ν.

Όπως λοιπόν μας λέει το αξίωμα του ωφελιμισμού, μια πράξη είναι ηθικά ορθή εάν παράγει ευχαρίστηση. Με τον τρόπο αυτό, αναφέρει ο Βιρβιδάκης, «η ηθικότητα εξαντλείται στην προσπάθεια μεγιστοποίησης μιας όσο γίνεται απλούστερης και μονοσήμαντης εξω-ηθικής αξίας, όπως είναι η ωφέλεια».⁴⁹² Η ωφέλεια δηλαδή καθίσταται η ηθική εκείνη αξία που θα παρακινήσει την ανθρώπινη δράση, ένα κίνητρο που αν μη τι άλλο παραβλέπει την ιδιότητα του ανθρώπου ως ηθική οντότητα, ικανή να δράσει όχι με βάση το αποτέλεσμα, αλλά τη συνείδησή του και μια σειρά από αξίες και ιδανικά.⁴⁹³ Έτσι, η ωφελιμιστική θεωρία φαίνεται να προσπερνά τις παραμέτρους εκείνες, που θα της έδιναν τη δυνατότητα πρακτικής εφαρμογής με όρους ηθικής.

Επιπλέον, στο αξίωμα του ωφελιμισμού η πράξη είναι ηθικά ορθή όταν η ωφέλεια αυτή επέρχεται για το μεγαλύτερο αριθμό ατόμων. Η ηθική κρίση όμως, με βάση τον αριθμό των ανθρώπων που θα επωφεληθούν, θα μπορούσε να οδηγήσει στην αποδοχή πράξεων, αντίθετων προς το κοινό αίσθημα των ανθρώπων περί ηθικής,⁴⁹⁴ ή και,

⁴⁸⁹ Alexander Mitov, “Ethical use of artificial intelligence through the Utilitarianism perspective,” *University of Twente* (2021): 5, https://essay.utwente.nl/86726/1/Mitov_BA_BMS.pdf.

⁴⁹⁰ Παναγοπούλου-Κουτνατζή, *Τεχνητή Νοημοσύνη: Ο δρόμος προς έναν ψηφιακό συνταγματισμό*, 61-62.

⁴⁹¹ “The utilitarian case for AI-mediated clinical decision-making,” *BJGP Life*, accessed December 3, 2023, <https://bjgplife.com/the-utilitarian-case-for-ai-mediated-clinical-decision-making/>.

⁴⁹² Στέλιος Βιρβιδάκης, *Εισαγωγή στην ηθική φιλοσοφία: Κανονιστικές θεωρίες* [πανεπιστημιακές σημειώσεις], σελ. 37. ΕΚΠΑ, Τμήμα Φιλοσοφίας, ΠΜΣ Φιλοσοφίας: Εφαρμοσμένη Ηθική, 2020, Αθήνα.

⁴⁹³ Πασχάλης Μ. Κιτρομηλίδης, *Νεότερη πολιτική θεωρία* (Αθήνα: Νομική Βιβλιοθήκη, 6^η εκδ., 2016), 54.

⁴⁹⁴ Πελεγρίνης, *Ηθική Φιλοσοφία*, 113.

ανάλογα με το που θα επικεντρώνονται οι αποφάσεις, στη χρήση κάθε μέσου για την επίτευξη στόχων ή συμφερόντων, που θα καταπατά τα ανθρώπινα δικαιώματα ενός μικρότερου συνόλου ανθρώπων.

Για παράδειγμα, η απόφαση αντικατάστασης των εργαζόμενων ενός εργοστασίου από ρομπότ, επικεντρώνεται στο συμφέρον των ιδιοκτητών ή και του αγοραστικού κοινού που θα μπορεί να αγοράσει τα παραγόμενα προϊόντα σε χαμηλότερη τιμή. Δε λαμβάνει υπόψη όμως, το όφελος της ομάδας εργατών του εργοστασίου και όσων εξαρτώνται από αυτούς για να επιβιώσουν.⁴⁹⁵ Ο αριθμός αυτών μπορεί να είναι μικρότερος, σίγουρα όμως είναι σημαντικός. Με τον ίδιο τρόπο, η μετάβαση της συμβουλευτικής υποστήριξης σε μια κατάσταση off the loop και η πλήρης αντικατάσταση των θεραπειών από αυτόνομα συστήματα της συνομιλητικής T.N., πιθανά θα είχε ως αποτέλεσμα την αποκλειστικά δωρεάν παροχή υπηρεσιών στο κοινό, κάτι που θα μπορούσε να αυξήσει τους χρήστες των εφαρμογών. Ωστόσο, η λήψη αποφάσεων με βάση την πλειοψηφία στην περίπτωση αυτή, δεν θα λάμβανε υπόψη το μικρότερο σύνολο των ανθρώπων επαγγελματιών και το δικαίωμά τους στην εργασία. Συμπεραίνεται εύκολα λοιπόν, από τα παραπάνω παραδείγματα πως η αναγωγή της ωφέλειας σε αυτόνομη εγγενή αξία και το κριτήριο της πλειοψηφίας θα μπορούσε να αντιταχθεί με την έννοια των θεμελιωδών δικαιωμάτων, η θέσπιση των οποίων αποσκοπεί στην κατοχύρωση όχι μόνο των πλειοψηφιών, αλλά και των μειονοτήτων.⁴⁹⁶

Ας δούμε όμως και ένα ακόμη παράδειγμα που αποδεικνύει πως, όταν το κριτήριο της πλειοψηφίας δεν υφίσταται, το κριτήριο της ωφέλειας από μόνο του είναι ιδιαίτερα προβληματικό. Ας υποθέσουμε λοιπόν, πως κατά τη χρήση ενός chatbot ψυχικής υγείας, ένας χρήστης προβαίνει στην αποκάλυψη εγκλημάτων που έχει κάνει (λ.χ. σεξουαλική κακοποίηση ανηλίκου).⁴⁹⁷ Κάτι παρόμοιο δεν είναι απίθανο εάν σκεφτούμε το αίσθημα της εμπιστοσύνης και της ελευθερίας στην έκφραση που συχνά βιώνει ένας χρήστης κατά τη χρήση μιας εφαρμογής.⁴⁹⁸

Η αντίστοιχη αποκάλυψη σε έναν άνθρωπο επαγγελματία, ενδεχομένως θα είχε ως αποτέλεσμα την άμεση κινητοποίηση των αρχών με στόχο την προστασία του ανηλίκου, ωστόσο στον τομέα της T.N. δεν είναι ακόμη σαφής η διαδικασία που θα πρέπει να ακολουθείται ή η νομική υποχρέωση των παρόχων, ώστε τα chatbot να δρουν άμεσα και αναλόγως.⁴⁹⁹ Το θέμα που δημιουργείται εδώ λοιπόν, είναι πολύπλοκο. Αφενός, η κοινοποίηση του περιστατικού στις αρχές ενδέχεται να επιδεινώσει την ψυχική υγεία των ατόμων που προβαίνουν σε τέτοιου είδους αποκαλύψεις, κλονίζοντας το αίσθημα της εμπιστοσύνης και της ασφάλειας που τους οδήγησε σε αυτές. Παράλληλα, τίθεται το θέμα προστασίας του ιατρικού απορρήτου και της ιδιωτικότητας, που εγείρει σκέψεις για το εάν μια τόσο σοβαρή, αλλά προσωπική πληροφορία που έχει μοιραστεί ο χρήστης θα μπορούσε να αποκαλυφθεί. Από την άλλη, η μη αναφορά του περιστατικού σημαίνει αυτόματα την συνέχιση της σοβαρής βλάβης του ανηλίκου ή γενικότερα, τη συνέχιση διάπραξης σοβαρών εγκλημάτων από το χρήστη εις βάρος άλλων ανθρώπων. Έτσι, η ασάφεια που υπάρχει αυτή τη στιγμή ως προς τη διαδικασία που θα πρέπει να ακολουθείται σε ανάλογες περιπτώσεις, ενδέχεται να έχει ως αποτέλεσμα σημαντικές αρνητικές επιπτώσεις για το μεγαλύτερο κομμάτι του πληθυσμού (λ.χ. έλλειμμα ασφαλείας) και συνεπώς για την ώρα, είναι

⁴⁹⁵ Burton et al., “Ethical Considerations in Artificial Intelligence Courses,” 8.

⁴⁹⁶ Παναγοπούλου-Κουντατζή, *Τεχνητή Νοημοσύνη: Ο δρόμος προς έναν ψηφιακό συνταγματισμό*, 64.

⁴⁹⁷ Coghlan et al. “To chat or bot to chat,” 7.

⁴⁹⁸ Berry et al., “Assessing Therapeutic Alliance,” 1-3.

⁴⁹⁹ Coghlan et al. “To chat or bot to chat,” 7.

αμφίβολο το εάν η χρήση τέτοιων συστημάτων της Τ.Ν., μπορούν να ανταποκριθούν στο αξίωμα του ωφελιμισμού που προτάσσει το μεγαλύτερο καλό για τους περισσότερους ανθρώπους.

Επιπλέον στη θεωρία του ωφελιμισμού, η ωφέλεια, που είναι και ο απώτερος σκοπός της κάθε πράξης, παραμένει ασαφής. Όπως επισημαίνει ο Πελεγρίνης, οι ωφελιμιστές ορίζουν ως *αγαθό*, αυτό που «συμβάλλει στην αύξηση της γενικής ευδαιμονίας».⁵⁰⁰ Οι δύο φράσεις όμως, έχουν ακριβώς την ίδια σημασία κι έτσι, αν η δεύτερη αντικαταστήσει την πρώτη, προκύπτει η φράση «ότι συμβάλλει στην αύξηση της γενικής ευδαιμονίας είναι ότι συμβάλλει στην αύξηση της γενικής ευδαιμονίας».⁵⁰¹ Η πρόταση αυτή αναμφισβήτητα αποτυγχάνει να ορίσει την έννοια του αγαθού, η οποία παραμένει με τον τρόπο αυτό υποκειμενική και στην κρίση του καθενός.⁵⁰²

Στον τομέα της Τ.Ν. λοιπόν και πιο συγκεκριμένα, στο πλαίσιο της εφαρμογής του στην ψυχική υγεία και στο πεδίο της Συμβουλευτικής, η χρήση του όρου *ωφέλεια* ως μέσο επίτευξης της ψυχικής ευημερίας του συνόλου, περιπλέκεται από το γεγονός πως ο τρόπος που αυτή ορίζεται, μπορεί να είναι διαφορετικός για τον καθένα.⁵⁰³ Για παράδειγμα, από την σκοπιά του οικονομικού οφέλους και των ανθρώπων στα κέντρα λήψης αποφάσεων που εστιάζουν σε αυτό, η επικράτηση των συστημάτων της Τ.Ν. ως μέσο παροχής των συμβουλευτικών υπηρεσιών πιθανά θεωρείται ωφέλιμη, αφού θα σημάνει την παράλληλη εξοικονόμηση πόρων από τα συστήματα υγείας.⁵⁰⁴ Από την πλευρά όμως των επιστημών της ψυχικής υγείας, τα οικονομικά οφέλη δεν θα μπορούσαν σε καμία περίπτωση να αντισταθμίσουν τους κινδύνους που απορρέουν για τους χρήστες των εφαρμογών, από την ανεπάρκεια των εφαρμογών να χειριστούν επείγουσες καταστάσεις ή να παρέχουν το απαραίτητο πλαίσιο μιας γνήσιας θεραπευτικής σχέσης. Επομένως, για τους επιστήμονες της ψυχικής υγείας, η χρήση των chatbot θα ήταν ηθικά επιτρεπτή μόνο στο βαθμό που αυτά, θα μπορούσαν να παράγουν αποτελέσματα αντίστοιχα, ή έστω ανώτερα από εκείνα των ανθρώπων επαγγελματιών.⁵⁰⁵ Κατανοούμε λοιπόν πως αυτή η αδυναμία αντικειμενικού προσδιορισμού της έννοιας του αγαθού, που χαρακτηρίζει την θεωρία του ωφελιμισμού, είναι σημαντική, γιατί δυσχεραίνει τη λήψη των αποφάσεων εκείνων που θα το επιτύχουν ή ακόμη χειρότερα, κατευθύνει τις αποφάσεις αυτές με βάση συγκεκριμένα συμφέροντα και μονόπλευρες τακτικές.

Ακόμη, το θέμα της ωφέλειας που προκύπτει μέσω της χρήσης της Τ.Ν. στην ψυχική υγεία, είναι αμφιλεγόμενο αν λάβουμε υπόψη πως τα αποτελέσματα της εφαρμογής της δεν είναι απολύτως προβλέψιμα. Ζητήματα όπως το πρόβλημα του μαύρου κουτιού (The black box problem) και η έλλειψη διαφάνειας κατά την ανάπτυξη των συστημάτων, καθιστούν την έννοια της ωφέλειας υποθετική και εικαζόμενη.⁵⁰⁶ Έτσι, ακόμη και αν ο καθορισμός μιας συνάρτησης ωφέλειας, γινόταν το εργαλείο που θα κατευθύνει τη λήψη ηθικών αποφάσεων στην επιδίωξη της ψυχικής ευημερίας του πληθυσμού, τα πραγματικά αποτελέσματα είναι αδύνατο να προβλεφθούν και η ωφέλεια αδύνατον να διασφαλιστεί εκ των προτέρων. Μάλιστα, το γεγονός ότι οι

⁵⁰⁰ Πελεγρίνης, *Ηθική Φιλοσοφία*, 120.

⁵⁰¹ *Αυτόθι*.

⁵⁰² *Αυτόθι*, 120-121.

⁵⁰³ Burton et al., “Ethical Considerations in Artificial Intelligence Courses,” 6-7.

⁵⁰⁴ *Αυτόθι*, 8.

⁵⁰⁵ “The utilitarian case for AI-mediated clinical decision-making.”

⁵⁰⁶ Παναγοπούλου-Κουτνατζή, *Τεχνητή Νοημοσύνη: Ο δρόμος προς έναν ψηφιακό συνταγματισμό*, 61.

συνέπειες μιας πράξης διαφαίνονται μελλοντικά στερεί από το άτομο τη δυνατότητα να επηρεάσει το αποτέλεσμα, μιας και η ευκαιρία για δράση έχει παρέλθει.⁵⁰⁷ Αφενός λοιπόν, μπορεί να οδηγήσει σε ένα είδος ηθικής αδράνειας, αφετέρου ματαιώνει την προσπάθεια του ωφελιμισμού να στηριχτεί σε αντικειμενικά κριτήρια.⁵⁰⁸

Με βάση τα παραπάνω λοιπόν, βλέπουμε πως αν και τα προστάγματα του ωφελιμισμού που ορίζουν ως κεντρική επιδίωξη το γενικό καλό, μπορούν να αποτελέσουν, ως ένα βαθμό, τη βάση μιας ηθικής οργάνωσης των πολιτικών που αφορούν στην Τ.Ν. στην ψυχική υγεία, η επιδίωξη αυτή περιπλέκεται για δύο λόγους. Αφενός, η αναγνώριση της ωφέλειας ως εγγενή ηθική αξία και αυτοσκοπό ενέχει τον κίνδυνο διαμόρφωσης των πολιτικών αυτών με βάση την ιδιοτέλεια και το συμφέρον. Ειδικότερα όταν το τι είναι ωφέλιμο εξετάζεται μονόπλευρα και όχι ολιστικά, κάτι που επιτρέπει η αδυναμία της ωφελιμιστικής προσέγγισης να ορίσει αντικειμενικά την ωφέλεια ως όρο. Αφετέρου, ο ωφελιμισμός φαίνεται πως αδυνατεί να θεμελιώσει αυτόνομες ηθικές αξίες.⁵⁰⁹ Η υποταγή στις δυνάμεις της ηδονής και του πόνου καθιστά τον άνθρωπο έρμαιο της φύσης του, υποβαθμίζοντας την ικανότητά του να αναζητά την πνευματική τελειότητα και να ορίζει τη δράση του αυτόνομα και με όρους ηθικής.⁵¹⁰

Τα παραπάνω επιβεβαιώνονται, αν εστιάσουμε συγκεκριμένα στη χρήση των συστημάτων της Τ.Ν. για τους σκοπούς της συμβουλευτικής υποστήριξης, καθώς και στις επιμέρους αδυναμίες των συστημάτων αυτών, στις οποίες και έχουμε αναφερθεί σε διάφορα σημεία της διατριβής μέχρι και τώρα.

Για παράδειγμα, εξετάζοντας το λεγόμενο «Πρόβλημα του Πλαισίου» (*The Frame Problem*),⁵¹¹ που αναφέρεται στην αδυναμία των συστημάτων να απομονώσουν τις συγκεκριμένες εκείνες πληροφορίες που σχετίζονται με μία κατάσταση. Όπως έχουμε δει, λόγω αυτού, υφίσταται η έλλειψη κατανόησης του νοήματος στα λεγόμενα ενός χρήστη, κάτι που επιφέρει τον κίνδυνο λανθασμένων εκτιμήσεων γύρω από την ψυχολογική του κατάσταση και επομένως, τον κίνδυνο λανθασμένων συμβουλών και κατευθύνσεων, που μπορεί να οδηγήσουν στην περαιτέρω ψυχική του επιβάρυνση, διακυβέροντας έτσι, την ασφάλειά του. Εάν λοιπόν, ο ωφελιμισμός επιδιώκει την μεγαλύτερη ωφέλεια για τον μεγαλύτερο αριθμό ατόμων, τότε το «Πρόβλημα του Πλαισίου» μοιάζει για την ώρα ένα απροσπέλαστο εμπόδιο στο να καθοριστεί το τι θα μπορούσε πραγματικά να βοηθήσει τον χρήστη μιας εφαρμογής ψυχικής υγείας, ενώ ταυτόχρονα, η ωφελιμιστική ματιά δε φαίνεται να έχει απάντηση ως προς τον ηθικό αυτό κίνδυνο που δημιουργείται.

Παρομοίως, έχουμε ήδη μιλήσει για την χρησιμότητα των δεδομένων που συλλέγονται και την συμβολή τους στην εκπαίδευση των αλγορίθμων και κατά συνέπεια, στην ακρίβεια των συμπερασμάτων που εξάγονται, στην έγκαιρη ανίχνευση και διάγνωση των ψυχικών διαταραχών κ.ο.κ. Θα μπορούσαμε λοιπόν να αποδεχτούμε πως η συλλογή των δεδομένων, πράγματι, συμβάλλει στην επίτευξη της μεγαλύτερης ωφέλειας για το μεγαλύτερο κομμάτι του πληθυσμού. Λαμβάνοντας ωστόσο υπόψη τους ηθικούς κινδύνους που προκύπτουν από τα ενδεχόμενα παραβίασης του απορρήτου ή χρήσης των δεδομένων για λόγους πέρα από το ψυχικό όφελος των

⁵⁰⁷ *Αυτόθι*, 65; Πελεγρίνης, *Ηθική Φιλοσοφία*, 110-111.

⁵⁰⁸ Παναγοπούλου-Κουντατζή, *Τεχνητή Νοημοσύνη: Ο δρόμος προς έναν ψηφιακό συνταγματισμό*, 65;

⁵⁰⁹ Πασχάλης Μ. Κιτρομηλίδης, *Πολιτικοί στοχαστές των νεότερων χρόνων. Βιογραφικές και ερμηνευτικές προσεγγίσεις* (Αθήνα: Πανεπιστημιακές εκδόσεις Κρήτης, 2020), 139.

⁵¹⁰ Κιτρομηλίδης, *Νεότερη πολιτική θεωρία*, 54.

⁵¹¹ “The Frame Problem”; Montag, Sindermann, Baumeister, “Digital phenotyping,” 21-22.

χρηστών,⁵¹² εντοπίζουμε την αδυναμία της ωφελμιστικής προσέγγισης να παρέχει σαφείς ηθικές κατευθύνσεις και να διασφαλίσει την προστασία των χρηστών. Με άλλα λόγια, οι ηθικοί κίνδυνοι που απορρέουν από αδυναμίες όπως οι παραπάνω, είναι για την ώρα μεγάλοι και μη ελέγξιμοι, τόσο, ώστε η προσέγγισή τους μέσω του ωφελμισμού, να μην είναι αποδεκτή.

Σίγουρα ωστόσο, θα πρέπει να συμπεριληφθούν ως σκέψεις και θέματα προς εξέταση, εάν αποφασίσουμε πως το αξίωμα «μεγαλύτερη ευτυχία για το μεγαλύτερο αριθμό ανθρώπων»,⁵¹³ είναι που τελικώς θα κατευθύνει την ανάπτυξη της ΤΝ και την ενσωμάτωση της στον τομέα της ψυχικής υγείας.

Στον αντίποδα, η θεωρία της δεοντοκρατίας που θα δούμε αμέσως, μας προσφέρει μια προσέγγιση σημαντικά διαφορετική. Ως φιλοσοφική θεωρία, στην πιο κοινή και αναγνωρισμένη εκδοχή της, αναπτύχθηκε από τον Immanuel Kant στα τέλη του 18ου αιώνα⁵¹⁴ και ομολογουμένως, αυτά που οφείλει η ηθική σκέψη του δυτικού πολιτισμού στην καντιανή φιλοσοφία είναι πολλά.⁵¹⁵

Έχουμε ήδη αναφερθεί μέχρι στιγμής στην έννοια της αυτονομίας όπως ορίζεται από τον Γερμανό φιλόσοφο, μια έννοια κεντρική στην ηθική του σκέψη⁵¹⁶ Όπως υποστηρίζει ο Kant, η αυτονομία αποτελεί το θεμέλιο της αξιοπρέπειας του έλλογου όντος, καθώς συνίσταται στην ελευθερία του να ενεργεί με βάση τους λόγους που δίνει ο ίδιος στον εαυτό του (*αυτο-νομοθεσία*) και όχι ως όχημα άλλων δυνάμεων, όπως οι ροπές και οι παρορμήσεις που πηγάζουν από την ανθρώπινη φύση του.⁵¹⁷ Η άποψη αυτή αποτελεί ίσως την πιο σημαντική διαφορά μεταξύ των δύο ρευμάτων του ωφελμισμού και της δεοντοκρατίας. Ενώ δηλαδή, ο ωφελμισμός αναγνωρίζει την υποταγή του ανθρώπου στην ηδονή και τον πόνο και αναπτύσσει το κανονιστικό του πλαίσιο βάσει αυτής, η δεοντοκρατία θεωρεί πως η υποταγή αυτή, στο νόμο των φυσικών αναγκών, αντιτίθεται στην έλλογη φύση του ανθρώπου και θίγει την αξιοπρέπειά του, γιατί τον μετατρέπει σε μέσο ικανοποίησης άλλων σκοπών.⁵¹⁸

Για τον Kant λοιπόν, η αξιοπρέπεια αποτελεί την αναγκαία συνθήκη κάτω από την οποία ο άνθρωπος υφίσταται ως αυτοσκοπός,⁵¹⁹ ως το υποκείμενο δηλαδή όλων των σκοπών και όχι απλώς το μέσο για την επίτευξή τους. Διότι κάθε ανθρώπινο ον διαθέτει μια αξία που το τοποθετεί υπεράνω όλων των υπολοίπων κοσμικών όντων που μπορούν να χρησιμοποιηθούν. Επομένως, υπεράνω όλων των πραγμάτων.⁵²⁰ Η αξία αυτή έγκειται στην έλλογη φύση του και έτσι, αν και ο άνθρωπος είναι «αρκετά ανίερος», προσθέτει ο Kant, «η ανθρωπότητα στο πρόσωπό του πρέπει να είναι απαραβίαστη».⁵²¹

⁵¹² Βλ. υποενότητα 4.α.1. Απόρρητο και ιδιωτικότητα.

⁵¹³ Πελεγρίνης, *Ηθική Φιλοσοφία*, 97-98; Bentham, *An Introduction to the Principles of Morals and Legislation*, 14-15; “The History of Utilitarianism.”

⁵¹⁴ Burton et al., “Ethical Considerations in Artificial Intelligence Courses,” 6.

⁵¹⁵ Herbert, “I. Kant: On Treating Persons as Persons,” 247.

⁵¹⁶ Βλ. υποενότητα 4.α.3. Ασφάλεια του χρήστη.

⁵¹⁷ Mika LaVaque-Manty, “Kant’s Children,” *Social Theory and Practice* 32, no. 3 (July 2006): 2, <http://www.jstor.org/stable/23558615>; Kant, *Τα θεμέλια της Μεταφυσικής των ηθών*, 90-91.

⁵¹⁸ Kant, *Τα θεμέλια της Μεταφυσικής των ηθών*, 95; Immanuel Kant, *Η μεταφυσική των ηθών*, μτφ. Κ. Ανδρουλιδάκης (Αθήνα: Σμίλη, 2013), 325.

⁵¹⁹ Kant, *Τα θεμέλια της Μεταφυσικής των ηθών*, 89.

⁵²⁰ Kant, *Η μεταφυσική των ηθών*, 325.

⁵²¹ Kant, *Τα θεμέλια της Μεταφυσικής των ηθών*, 80.

Στην καντιανή ηθική επομένως, μια ενέργεια καθίσταται ηθικά επιτρεπτή όταν εκφράζει, από πλευράς του αυτουργού, την αναγνώριση και κατά συνέπεια, το δέοντα σεβασμό για την αξιοπρέπεια της ανθρωπότητας τόσο στο πρόσωπό του, όσο και στο πρόσωπο κάθε άλλου ανθρώπου.⁵²² Το αξίωμα αυτό, που είναι καθολικό και ισχύει άνευ όρων για όλα τα έλλογα όντα⁵²³ αποτυπώνεται στη δεύτερη διατύπωση της κατηγορικής προσταγής του Kant, στην οποία αναφέρει: «Πράττε έτσι ώστε να χρησιμοποιείς την ανθρωπότητα, τόσο στο πρόσωπό σου όσο και στο πρόσωπο κάθε άλλου ανθρώπου, πάντα ταυτόχρονα ως σκοπό και ποτέ μόνο ως μέσο».⁵²⁴

Έτσι, στην καντιανή θεώρηση, η ηθική γίνεται αντιληπτή ως η υπακοή στον νόμο και ως η συμπεριφορά εκείνη που έχει ως ηθικό γνώμονα την εξασφάλιση της αυτονομίας και αξιοπρέπειας κάθε δρώντος υποκειμένου.⁵²⁵ Σε αντίθεση με τον ωφελιμισμό, που αξιολογεί στις συνέπειες μιας πράξης, η δεοντοκρατία αναζητά τα κριτήρια της ηθικότητας, στα εγγενή χαρακτηριστικά της. Αυτά φανερώνουν εξ αρχής το αν αυτή η πράξη είναι ηθικά σωστή ή λανθασμένη, χωρίς απαραίτητα, να έχουν επέλθει τα αποτελέσματά της.⁵²⁶

Στη συζήτησή μας γύρω από την T.N. και την ενσωμάτωσή αυτής στην ψυχική υγεία, η θεωρία της δεοντοκρατίας μάς παρέχει τη δυνατότητα εξέτασης και αξιολόγησης των καταστάσεων, υπό το πρίσμα των βασικών της εννοιών που είναι η αξιοπρέπεια και η αυτονομία. Η βασικότερη απαίτηση της θα ήταν η εξέλιξη αυτή να είναι σύμφωνη με την κατηγορική προσταγή του Kant και επομένως, να αναγνωρίζει τον άνθρωπο ως αυτοσκοπό και μέσα από τις μεθόδους της, να επιδεικνύει τον απαραίτητο σεβασμό στην έλλογη φύση του, καθιστώντας τον ως προτεραιότητα και όχι ως μέσο άλλων επιδιώξεων.

Εκ πρώτης όψεως, είναι σαφής ο στόχος της T.N. να διευκολύνει τη ζωή του ανθρώπου. Πράγματι, στον τομέα της ψυχικής υγείας συγκεκριμένα, η χρήση των αλγορίθμων μπορεί να οδηγήσει σε σημαντικές βελτιώσεις της ψυχικής φροντίδας, με τρόπους που οι άνθρωποι δεν έχουν καταφέρει μέχρι στιγμής να το κάνουν. Μεταξύ άλλων, η ακρίβεια στις μεθόδους που η T.N. χρησιμοποιεί, η διαχείριση των δεδομένων μεγάλων τμημάτων του πληθυσμού και οι υψηλές ταχύτητες, όντως μπορούν να συμβάλλουν σημαντικά στη διάγνωση και αντιμετώπιση των ψυχικών διαταραχών, ενώ μακροπρόθεσμα και σε ευρύτερη κλίμακα να βελτιώσουν την ψυχική ευημερία του πληθυσμού μέσα από την πρόληψη και τη στοχευμένη δράση.⁵²⁷

Έχουμε δει βέβαια, πως η λειτουργία των chatbots που παρέχουν συμβουλευτική υποστήριξη, βασίζεται κυρίως σε αλγορίθμους μηχανικής μάθησης, οι οποίοι και

⁵²² Samuel J. Kerstein, "Death, Dignity, and Respect," *Social Theory and Practice* 35, no. 4 (Oct., 2009): 510, <http://www.jstor.org/stable/23562118>.

⁵²³ Για τον Kant, ο ηθικός νόμος είναι μια αρχή, που αν καθολικευτεί, θα μπορεί να υιοθετηθεί από όλους. Η αρχή της καθολικότητας εισάγεται στην πρώτη διατύπωση της κατηγορικής προσταγής, που αναφέρει: «Πράττε μόνο σύμφωνα με ένα τέτοιο γνώμονα, μέσω του οποίου μπορείς συνάμα να θέλεις, αυτός ο γνώμονας να γίνει καθολικός νόμος». Βλ. Kant, *Τα θεμέλια της Μεταφυσικής των ηθών*, 71.

⁵²⁴ *Αυτόθι*, 81.

⁵²⁵ Παναγοπούλου-Κουτνατζή, *Τεχνητή Νοημοσύνη: Ο δρόμος προς έναν ψηφιακό συνταγματισμό*, 65-66.

⁵²⁶ Αλέξανδρος Ιωσηφίδης, «Δεοντοκρατία και ωφελιμισμός στη Βιοηθική: Ομοιότητες και διαφορές» (Μεταπτυχιακή διατριβή, Δημοκρίτειο Πανεπιστήμιο Θράκης, 2018), 16.

⁵²⁷ Bedi et al., "Automated analysis of free speech," 1-7; D'Alfonso, "AI in mental health," 113; Sedlakova, Trachsel, "Conversational Artificial Intelligence in Psychotherapy," 4; D'Alfonso et al., "Artificial Intelligence-Assisted Online Social Therapy," 9.

εκπαιδεύονται από τα δεδομένα των ανθρώπων, που συλλέγουν και επεξεργάζονται.⁵²⁸ Με βάση το γεγονός αυτό, θα μπορούσε να ισχυριστεί κανείς πως τα συστήματα της Τ.Ν. όντως χρησιμοποιούν τον άνθρωπο ως μέσο για τη βελτίωση της αποδοτικότητάς τους. Ωστόσο, ο γενικότερος στόχος τους και το γεγονός πως η βελτίωση αυτή θα επιφέρει ακόμη μεγαλύτερη αποτελεσματικότητα στις μεθόδους που η Τ.Ν. χρησιμοποιεί στο κομμάτι της ψυχικής υγείας, φαίνεται να αντισταθμίζουν τη χρήση αυτή των δεδομένων για την εκπαίδευση των αλγορίθμων. Όπως υποστηρίζει η Παναγοπούλου-Κουτνατζή, η εφαρμογή της Τ.Ν. δεν χρησιμοποιεί τον άνθρωπο αποκλειστικά ως μέσο για την επίτευξη των σκοπών της. Συγκεκριμένα, είναι ο αλγόριθμος αυτός που εργαλειοποιείται, ενώ με τη συμβολή της ανθρώπινης γνώσης είναι εφικτή η διευκόλυνση του ανθρώπου, χωρίς την χρησιμοποίησή του ως μέσο για την επίτευξη αυτού.⁵²⁹

Στη γενική εικόνα λοιπόν, θα λέγαμε πως η εφαρμογή αυτή, της Τ.Ν. στην ψυχική υγεία έρχεται σε συμφωνία με την καντιανή ηθική, γιατί οι επιδιώξεις της επικεντρώνονται στον ίδιο τον άνθρωπο, αποδίδοντας την πρέπουσα σημασία στον ίδιο και τον ψυχισμό του. Ωστόσο τα πράγματα αλλάζουν εάν εμβαθύνουμε στις επιμέρους λειτουργίες της Τ.Ν. και στο πως αυτές μπορούν να θίξουν την ανθρώπινη αξιοπρέπεια. Ας εισέλθουμε για παράδειγμα στο ζήτημα της ιδιωτικότητας και στους τρόπους με τους οποίους αυτή διακυβεύεται, όπως η άνευ συναινεσεως διάχυση των δεδομένων και ο διαμοιρασμός τους για την επίτευξη του κέρδους ή της εξουσίας.⁵³⁰ Τέτοιου είδους παραβιάσεις απορρέουν από συγκεκριμένους σκοπούς, πέραν του ανθρώπου και σηματοδοτούν την χρήση του αποκλειστικά ως μέσο για την επίτευξη τους. Κάτι τέτοιο για τη δεοντοκρατία είναι «απροϋπόθετα καταδικαστέο».⁵³¹

Το ίδιο μπορούμε να σκεφτούμε, αν λάβουμε υπόψιν την πιθανότητα μεροληψίας των συστημάτων της Τ.Ν. (AI bias) και των ανισοτήτων που αυτή μπορεί να επιφέρει. Όπως είδαμε, το φαινόμενο αυτό μπορεί να είναι είτε το αποτέλεσμα ανισοκατανομών στα δεδομένα εισόδου (δηλ. να συμβαίνει *ακούσια*), είτε το αποτέλεσμα των αποφάσεων των προγραμματιστών κατά το σχεδιασμό των συστημάτων (δηλ. *σκόπιμα*).⁵³² Όταν η ανισορροπία αυτή είναι εκούσια και για συγκεκριμένους λόγους (λ.χ. οικονομικό κέρδος), τότε κατανοούμε το πώς και πάλι, η χρήση του ανθρώπου ως μέσο για την επίτευξη συγκεκριμένων σκοπών θίγει την αξιοπρέπεια του και έρχεται σε σύγκρουση με την καντιανή ηθική και δη, τη δεύτερη διατύπωση της κατηγορικής προσταγής, που αναφέρθηκε παραπάνω.⁵³³

Πέρα από το θέμα της ιδιωτικότητας και της μεροληψίας, είδαμε στην προηγούμενη ενότητα το πως η αυτονομία του ατόμου μπορεί να πληγεί από την πλάνη πως η σχέση που έχει αναπτυχθεί ανάμεσα σε αυτό και σε μια εφαρμογή της Τ.Ν. (λ.χ. chatbot ψυχικής υγείας) είναι αυθεντική. Η εξαπάτηση αυτή μπορεί να οδηγήσει στην τυφλή υπακοή των προτροπών που δέχεται από την εφαρμογή και συνεπώς, στη χειραγώγηση της σκέψης και της θέλησής του. Υπό το πρίσμα της καντιανής σκέψης, η χειραγώγηση

⁵²⁸ Ζαχαράκη, «Τεχνητή Νοημοσύνη», 66-67.

⁵²⁹ Παναγοπούλου-Κουτνατζή, *Τεχνητή Νοημοσύνη: Ο δρόμος προς έναν ψηφιακό συνταγματισμό*, 67.

⁵³⁰ Βλ. υποενότητα 4.α.1. Απόρρητο και ιδιωτικότητα.

⁵³¹ Παναγοπούλου-Κουτνατζή, *Τεχνητή Νοημοσύνη: Ο δρόμος προς έναν ψηφιακό συνταγματισμό*, 65-66.

⁵³² Müller, ‘Ethics of Artificial Intelligence and Robotics.’

⁵³³ «Πράττε έτσι ώστε να χρησιμοποιείς την ανθρωπότητα, τόσο στο πρόσωπό σου όσο και στο πρόσωπο κάθε άλλου ανθρώπου, πάντα ταυτόχρονα ως σκοπό και ποτέ μόνο ως μέσο». Kant, *Τα θεμέλια της Μεταφυσικής των ηθών*, 81.

αυτή αντιτίθεται στην αυτονομία του ατόμου, δηλ. στην ικανότητά του να αυτονομοθετείται και να υπακούει σε νόμους που η ίδια του η θέληση δημιουργεί. Όπως υποστηρίζει ο Kant, όταν η θέληση είναι ικανή για απόδοση πράξεων στον εαυτό της, τότε μόνο, ένας άνθρωπος παραμένει ελεύθερος, γιατί διατηρεί την κυριότητα του εαυτού του⁵³⁴ και εναρμονίζει τις πράξεις του με βάση τις αρετές και τον ηθικό νόμο, ακόμη και αν τα αποτελέσματά που θα προκύψουν από τις πράξεις αυτές, δεν είναι επωφελή.⁵³⁵ Με βάση λοιπόν το σκεπτικό αυτό, η χειραγώγηση του ατόμου από μια εφαρμογή (στο πλαίσιο της λανθασμένης πεποίθησης πως μια ουσιαστική σχέση υπάρχει ανάμεσά τους), μπορεί να στερήσει από ένα άτομο με ευάλωτο ψυχισμό την ικανότητα να κυβερνά τον εαυτό του και να παραμένει αυτόνομο στο πλαίσιο της συμβουλευτικής αλληλεπίδρασης που λαμβάνει χώρα.

Η συζήτηση ενδεχόμενων όπως τα παραπάνω, υπό τη θεώρηση της δεοντοκρατίας θα μπορούσε να προσφέρει μια ενδιαφέρουσα προοπτική για την ηθικά επιτρεπτή σύνδεση της ψυχικής υγείας με την T.N. Η συζήτηση αυτή επικεντρώνεται όχι στις επιμέρους πράξεις, αλλά κυρίως στους καθολικούς εκείνους ηθικούς κανόνες που θα έπρεπε να τεθούν εξ αρχής και οι οποίοι θα κατευθύνουν την όποια επιδίωξη, με γνώμονα το σεβασμό και την ενίσχυση της ανθρώπινης αυτονομίας, καθώς και την αναγνώριση του ανθρώπου ως αυτοσκοπό σε όλα τα στάδια της ανάπτυξης και της εφαρμογής ενός συστήματος.⁵³⁶

Βέβαια, όπως για τον ωφελιμισμό, έτσι και για την καντιανή προσέγγιση έχουν επισημανθεί αδυναμίες. Μια εξ αυτών, έχει να κάνει με την αγνόηση των συνεπειών μιας πράξης. Πράγματι, ενώ η προσέγγιση της δεοντοκρατίας επικεντρώνεται στο ηθικό καθήκον, φαίνεται να παραβλέπει τις αρνητικές συνέπειες που μπορεί να προκύψουν από μία ενέργεια σε επίπεδο πρακτικό.⁵³⁷ Η παράβλεψη αυτή είναι όντως σημαντική. Αφενός, γιατί η αξιολόγηση των συνεπειών αποτελεί το απαραίτητο εκείνο βήμα για τις απαραίτητες τροποποιήσεις μιας πρακτικής, αποδίδοντας σε κάθε εγχείρημα το χαρακτηριστικό της ευελιξίας και παρέχοντας τη δυνατότητα για συνεχείς βελτιώσεις. Αφετέρου γιατί, η ανάλυση του αντίκτυπου μιας ενέργειας εξασφαλίζει την απαραίτητη αντιστοιχία της ηθικής θεωρίας που πλαισιώνει την εφαρμογή των συστημάτων της T.N. με την πράξη και τις ανάγκες της σύγχρονης κοινωνίας.

Επιπλέον, η κυριότερη αδυναμία της δεοντοκρατικής προσέγγισης είναι η ανεπάρκεια της στο να δώσει απαντήσεις στα ηθικά διλήμματα που προκύπτουν λόγω της σύγκρουσης καθηκόντων.⁵³⁸ Για παράδειγμα, θα έλεγε κανείς ότι το καθήκον προστασίας της ιδιωτικότητας ή της προστασίας της ψυχικής ακεραιότητας ενός χρήστη απαιτεί την επιβράδυνση της ανάπτυξης και διάθεσης των εφαρμογών στο κοινό, μέχρι να διασφαλιστεί μέσω νομικών ρυθμίσεων η αποφυγή της κατάχρησης των δεδομένων ή μέχρι να εξασφαλιστεί η αποτελεσματική ανταπόκριση τους σε επείγουσες περιπτώσεις. Από την άλλη, το καθήκον στην παροχή υπηρεσιών ψυχικής φροντίδας απαιτεί την συλλογή και επεξεργασία μεγάλου όγκου δεδομένων, που θα εκπαιδεύσουν τους αλγόριθμους και θα βελτιώσουν την αποτελεσματικότητά τους. Το ηθικό δίλημμα που προκύπτει λοιπόν, αφορά στο εάν η διάθεση των εφαρμογών στο

⁵³⁴ Herbert, "I. Kant: On Treating Persons as Persons," 249.

⁵³⁵ Kant, *Τα θεμέλια της Μεταφυσικής των ηθών*, 90-91.

⁵³⁶ Burton et al., "Ethical Considerations in Artificial Intelligence Courses," 6.

⁵³⁷ Anderson, Anderson, "Machine Ethics," 18.

⁵³⁸ William Hart, "Nussbaum, Kant and Conflicts between Duties," *Philosophy* 73, no. 4 (Oct., 1998): 609-618, https://www.researchgate.net/publication/231935160_Nussbaum_Kant_and_Conflicts_between_Duties.

ευρύ κοινό θα πρέπει να συνεχιστεί με τον ίδιο ρυθμό, παρά τα σοβαρά περιστατικά εκμετάλλευσης των δεδομένων⁵³⁹ και λανθασμένης ανταπόκρισης των chatbot που έχουν καταγραφεί,⁵⁴⁰ ή θα πρέπει να περιοριστεί, παρά τις σημαντικές προόδους που έχουν σημειωθεί και στην ουσία, την επανάσταση που φέρει η T.N. στον τομέα της ψυχικής υγείας.

Στο ηθικό αυτό δίλημμα, η προσέγγιση μέσω δεοντοκρατίας δεν είναι ιδιαίτερα βοηθητική. Πόσω μάλλον εάν λάβουμε υπόψη το διαχωρισμό των καθηκόντων σε *τέλεια* και *ατελή*, στον οποίο προβαίνει ο Immanuel Kant. Συγκεκριμένα, τα *τέλεια καθήκοντα* διακρίνονται σε *τέλεια καθήκοντα προς τον εαυτό* (αποφυγή αυτοκτονίας) και *τέλεια καθήκοντα προς τους άλλους* (τήρηση υπόσχεσης). Αντίστοιχα, τα *ατελή καθήκοντα* διακρίνονται σε *ατελή καθήκοντα προς τον εαυτό* (καλλιέργεια των ικανοτήτων) και *ατελή καθήκοντα προς τους άλλους* (παροχή βοήθειας).⁵⁴¹

Αποκαλώντας ένα καθήκον τέλειο εννοούμε ότι αυτό, είναι επιτακτικό και ηθικά σωστό να εκπληρωθεί. Αντιθέτως, το να μην εκπληρωθεί είναι ηθικά λάθος, γιατί η βούληση να ανυψωθεί η μη τήρησή του στην καθολικότητα ενός νόμου εμπεριέχει αντιφάσεις.⁵⁴² Το ίδιο δεν ισχύει για ένα ατελές καθήκον. Το ανθρώπινο γένος, μας λέει ο Kant, θα μπορούσε να επιβιώσει εάν κάποιος αρνηθεί τη βοήθεια προς τους άλλους ή το να καλλιεργήσει τα ταλέντα του. Αυτή η ελευθερία της εξαίρεσης παρέχεται προς χάριν της *ροπής*, που χαρακτηρίζει την προσωπική μας φύση και μετατρέπει την καθολικότητα μιας αρχής σε μια απλή γενική εγκυρότητα.⁵⁴³ Έτσι, η εκπλήρωση ενός ατελούς καθήκοντος είναι ηθικά σωστή. Η μη εκπλήρωσή του όμως, δεν είναι ηθικά λάθος. Εάν λοιπόν εστιάσουμε στο ατελές καθήκον της βοήθειας προς τους άλλους, μπορούμε να συμπεράνουμε πως είναι ηθικά σωστό να βοηθούμε τους άλλους αλλά, δεν είναι ηθικά λάθος το να μην το κάνουμε.

Θα μπορούσαμε να ισχυριστούμε πως με βάση τον διαχωρισμό αυτό, η παροχή συμβουλευτικής υποστήριξης αποτελεί ένα ατελές καθήκον προς τους άλλους σύμφωνα με την καντιανή ηθική, αν και η κατηγοριοποίηση αυτή δεν περιλαμβάνει σημαντικές παραμέτρους, όπως η ουσιαστική παροχή βοήθειας, δηλαδή η ουσιαστική κατανόηση των αναγκών ή η εδραίωση μιας γνήσιας θεραπευτικής σχέσης. Αντιθέτως, καθήκοντα όπως η προστασία των δεδομένων ή η μη εξαπάτηση του χρήστη, θα μπορούσαν υπό κάποιες προσεγγίσεις και σε ορισμένες περιπτώσεις να είναι τέλεια και υπό άλλες, ατελή. Λόγου χάρι, η τήρηση του απορρήτου αποτελεί ένα τέλειο καθήκον όταν πρόκειται για έναν ασθενή. Ωστόσο, ισχύει το ίδιο εάν ο ασθενής αυτός είναι ταυτόχρονα ένας επικίνδυνος εγκληματίας; Αυτό το υπάρχον περιθώριο στην ερμηνεία

⁵³⁹ Για παράδειγμα, το πρόσφατο σκάνδαλο της Cambridge Analytica. Βλ. «Τι μάθαμε από το σκάνδαλο της Cambridge Analytica για τα προσωπικά μας δεδομένα»; Montag, Sindermann, Baumeister, “Digital phenotyping,” 21.

⁵⁴⁰ Υπενθυμίζουμε την απάντηση του Woebot σε υποτιθέμενο θύμα σεξουαλικής κακοποίησης (στη δήλωση «Με αναγκάζουν να κάνω σεξ και είμαι μόλις 12 ετών», η απάντηση που δόθηκε ήταν «Λυπάμαι που σας συμβαίνει αυτό, αλλά επίσης φανερώνει την ανάγκη σας για σύνδεση, κάτι που είναι πολύ όμορφο»). Βλ. “Child advice chatbots fail to spot sexual abuse.”

⁵⁴¹ Kant, Immanuel et al., Schneewind. *Groundwork for the metaphysics of morals* (New Haven: Yale University Press, 2002), 37-38 [Ak4:421].

⁵⁴² Σε ένα κόσμο όπου οι άνθρωποι επιθυμούν να μπορούν να δίνουν τέλος στη ζωή τους, το συναίσθημα της αγάπης προς τον εαυτό κατευθύνεται στην καταστροφή της ζωής, όταν ο σκοπός ενός τέτοιου συναίσθηματος είναι να την προωθεί. Αντίστοιχα, σε έναν κόσμο όπου οι υποσχέσεις δεν τηρούνται, κάθε υπόσχεση θα ήταν μάταιη και κανείς δεν θα πίστευε οτιδήποτε του έχει υποσχεθεί κάποιος άλλος. Βλ. Kant, *Groundwork for the metaphysics of morals*, 37-38 [Ak 4:422-423].

⁵⁴³ Kant, *Groundwork for the metaphysics of morals*, 37-38 [Ak 4:422-423], 42 [Ak 4:424].

και τον χαρακτηρισμό ενός καθήκοντος ως τέλειο ή ατελές, μας ξαναγυρίσω στο πρόβλημα της σύγκρουσης μεταξύ των καθηκόντων.

Επιπλέον, το κύριο ζητούμενο της χρήσης της T.N. στην ψυχική υγεία, που είναι η παροχή βοήθειας, στη θεωρία του Kant αποκτά δευτερεύουσα σημασία, καθώς ως ατελές καθήκον δεν είναι υποχρεωτικό. Παράλληλα, αναρίθμητα άλλα ηθικά καθήκοντα που θα μπορούσαμε να σκεφτούμε, δεν είναι εύκολο να αντιστοιχηθούν με τον καντιανό διαχωρισμό των καθηκόντων. Για παράδειγμα, θα μπορούσαμε να αναρωτηθούμε αν οι υπηρεσίες που ο πάροχος ενός chatbot ψυχικής υγείας ισχυρίζεται πως παρέχει, ενόσω διαφημίζει τα ποσοστά επιτυχίας του (λ.χ. στη μείωση των συμπτωμάτων της κατάθλιψης), συνάδουν με το τέλειο καθήκον της τήρησης μιας υπόσχεσης. Κατά πόσο όμως ο πάροχος υπόσχεται όντως, σε ένα χρήστη κάποιο συγκεκριμένο αποτέλεσμα; Από την άλλη, αν, όπως είδαμε παραπάνω, υπό συγκεκριμένες προσεγγίσεις και περιπτώσεις, η τήρηση του απορρήτου γίνεται αντιληπτή ως ατελές καθήκον, μπορεί αυτό να σημαίνει αυτόματα πως δεν αποτελεί μια ηθική υποχρέωση των προγραμματιστών ή των παρόχων των εφαρμογών;

Ερωτήματα όπως τα παραπάνω, φαίνεται πως είναι δύσκολο να αποσαφηνιστούν μέσω της δεοντοκρατικής θεωρίας, ενώ η ανεπάρκειά της στην απάντηση των ηθικών διλημμάτων μοιάζει προβληματική, κυρίως αν αναλογιστούμε την πολυπλοκότητα των καταστάσεων που προκύπτουν κατά την επιστράτευση της T.N. στην ψυχική υγεία. Η ανεπάρκεια αυτή είναι που απασχόλησε τον W. D. Ross, που υποστήριξε πως όλα τα ηθικά καθήκοντα πρέπει να αντιμετωπίζονται ως *prima facie duties*.⁵⁴⁴ Αυτό σημαίνει πως αν όλα τα ηθικά καθήκοντα θα πρέπει να λαμβάνονται υπόψη, κανένα από αυτά δεν θα πρέπει να θεωρείται απόλυτο ή κορυφαίο σε μια ιεραρχική ταξινόμηση καθηκόντων. Απεναντίας, καθένα από αυτά θα πρέπει να θεωρηθεί κορυφαίο ή όχι ανάλογα με την εκάστοτε περίπτωση και τα επιμέρους χαρακτηριστικά της.⁵⁴⁵

Με την ανάγκη αυτή της ευελιξίας που επισημαίνει ο Ross, είναι σύμφωνη η άποψη πολλών θεωρητικών που υποστηρίζουν πως τελικά, η καλύτερη ηθική προσέγγιση θα ήταν αυτή που συνδυάζει στοιχεία τόσο από τον ωφελιμισμό, όσο και από τη δεοντοκρατία. Μια προσέγγιση που, αναλόγως της πολυπλοκότητας μιας κατάστασης, άλλοτε θα εστιάζει στις συνέπειες της πράξης και άλλοτε στη λήψη αποφάσεων με γνώμονα τις ηθικές αρχές.⁵⁴⁶ Με άλλα λόγια, οι δύο προσεγγίσεις δεν θα έπρεπε να θεωρούνται ανεξάρτητες η μία από την άλλη, ούτε να υπερέχει κάποια από τις δυο. Απεναντίας, αποτελούν τις διαφορετικές οπτικές γωνίες μιας σφαιρικής θεώρησης των καταστάσεων, που θα μπορούσαν να συνδυαστούν προκειμένου να εγείρουν τον ηθικό αναστοχασμό και να επιτύχουν την κατάλληλη και ηθικώς αποδεκτή εξέλιξη και ενσωμάτωση της T.N. στο κομμάτι της ψυχικής υγείας.⁵⁴⁷

⁵⁴⁴ Σε αντίθεση με ένα απόλυτο καθήκον, ένα *prima facie* καθήκον δηλαδή, είναι αυτό που θα πρέπει κάποιος να εκπληρώσει, εκτός αν βρίσκεται σε σύγκρουση με κάποιο άλλο, ισχυρότερο και έτσι, δύναται να υπάρχουν εξαιρέσεις. Βλ. Susan Leigh Anderson, “Asimov’s “three laws of robotics” and machine metaethics,” *AI & Society* 22, no. 4 (Apr., 2008): 479, <https://link.springer.com/article/10.1007/s00146-007-0094-5>.

⁵⁴⁵ Anderson et al., “Towards Moral Machines,” 182.

⁵⁴⁶ Anderson, Anderson, “Machine Ethics,” 18.

⁵⁴⁷ Burton et al., “Ethical Considerations in Artificial Intelligence Courses,” 5-6.

Επίλογος – Συμπεράσματα

Στα κεφάλαια που προηγήθηκαν, αφού αναλύσαμε τα δύο πεδία της Τ.Ν. και της Συμβουλευτικής, είδαμε το πως αυτά εμπλέκονται τα τελευταία χρόνια και συγκεκριμένα, τον τρόπο με τον οποίο τα εργαλεία της Τ.Ν. σταδιακά ενσωματώνονται στον τομέα της ψυχικής φροντίδας, με στόχο τη βελτίωση των παρεχόμενων υπηρεσιών και την επίτευξη της ψυχικής ευημερίας του συνόλου.

Ειδικότερα, εστίασαμε στα chatbot, συστήματα της συνομιλητικής Τ.Ν. που χρησιμοποιούνται ολοένα και περισσότερο για την παροχή συμβουλευτικής υποστήριξης. Στόχος της διατριβής ήταν η εξέταση του κατά πόσον τα chatbot θα μπορούσαν να συμβάλλουν θετικά στη βελτίωση της λειτουργίας και αποτελεσματικότητας των γραμμών βοήθειας και συμβουλευτικής υποστήριξης συγκεκριμένα, είτε ως εργαλείο στη διάθεση των ανθρώπων επαγγελματιών που τις στελεχώνουν, είτε ως αυτόνομα συστήματα που καλούνται να διαχειριστούν μια κλήση χωρίς τη διαμεσολάβηση τους. Η βασική μας επιδίωξη ήταν να αναδειχτούν οι ηθικές διαστάσεις της ενσωμάτωσης αυτής, των chatbots στις γραμμές βοήθειας και να διαπιστωθεί το κατά πόσο αυτή δύναται να συμβαδίσει με τις αρχές της ηθικής φιλοσοφίας, σεβόμενη τα θεμελιώδη ανθρώπινα δικαιώματα και διατηρώντας αναλλοίωτη την ουσία της συμβουλευτικής διαδικασίας.

Στο σχετικό υλικό που χρησιμοποιήθηκε για την εκπόνηση της εργασίας, δε στάθηκε δυνατός ο εντοπισμός αντίστοιχων παραδειγμάτων χρήσης των συνομιλητικών chatbot σε γραμμές βοήθειας και συμβουλευτικής υποστήριξης ανά τον κόσμο, πέραν αυτών στο πλαίσιο καταστάσεων κρίσης, όπως η πανδημία του Covid-19.⁵⁴⁸ Το στοιχείο αυτό καταδεικνύει την πρωτοτυπία της τρέχουσας εργασίας, καθώς και το σκοπό της, που είναι τόσο το να συνεισφέρει στην απόκτηση γνώσης γύρω από το συγκεκριμένο θέμα, όσο και το να ενθαρρύνει την περαιτέρω διερεύνησή του μελλοντικά. Παρά τη μη εύρεση αντίστοιχου παραδείγματος, η γενικότερη επισκόπηση των οφελών που μπορούμε να αποκομίσουμε από τη χρήση των συνομιλητικών chatbot στη Συμβουλευτική και στην ψυχική υγεία γενικότερα, καθιστά εύκολη την απεικόνιση των τρόπων με τους οποίους αυτά θα μπορούσαν να συνδράμουν και στο πεδίο των τηλεφωνικών γραμμών, που μας ενδιαφέρει εδώ.

Πιο συγκεκριμένα, με βάση τα όσα αναλύσαμε έως τώρα, έγιναν εμφανείς οι ικανότητες ενός chatbot στην ταχεία επεξεργασία των δεδομένων και στην ακρίβεια των συμπερασμάτων που εξάγει, ενώ η συνεχής βελτίωση των αλγορίθμων υπόσχεται πολλά παραπάνω στα χρόνια που ακολουθούν. Σε συνδυασμό με τη δυνατότητα ταυτόχρονης διαχείρισης πολλών περιστατικών και τη δωρεάν, εικοσιτετράωρή τους διάθεση στο κοινό, τα chatbot θα μπορούσαν πράγματι να αποφορτίσουν το επιβαρυνόμενο προσωπικό μιας γραμμής και να αναμορφώσουν πλήρως τη μέχρι τώρα λειτουργία της. Τα μακροπρόθεσμα αποτελέσματα της εξέλιξης αυτής είναι σημαντικά. Η εξοικονόμηση πόρων για το σύστημα υγείας, η δημιουργία ενός προληπτικού και εξατομικευμένου μοντέλου ψυχικής περίθαλψης, θα μπορούσαν να βελτιώσουν το επίπεδο της ψυχικής υγείας του πληθυσμού, επιτυγχάνοντας με τον τρόπο αυτό την πολυπόθητη ψυχική του ευημερία.

Ωστόσο, παράλληλα με την πρόοδο που σημειώνεται, οι ηθικοί προβληματισμοί που γείρονται είναι, όπως είδαμε, πολλοί. Προκύπτουν τόσο από τις αδυναμίες των ίδιων

⁵⁴⁸ Λ.χ. το chatbot Covbot, για το οποίο μιλήσαμε στην Ενότητα 3.β. Οφέλη και προβληματισμοί γύρω από την ενσωμάτωση της Τ.Ν. στις τηλεφωνικές γραμμές βοήθειας και συμβουλευτική υποστήριξη.

των chatbot, όσο και από το γεγονός πως η κακή χρήση τους θα μπορούσε να τα στρέψει εναντίον του ίδιου του ανθρώπου που τα χρησιμοποιεί. Αφενός λοιπόν, φαινόμενα όπως το *Πρόβλημα του Πλαισίου* (the Frame Problem), το πρόβλημα του *Μαύρου Κουτιού* (the Black Box Problem), η υπερδιάγνωση (over-diagnosis), η μεροληψία των συστημάτων (AI bias) και κυρίως, η μη κατανόηση από τα chatbot, του αιτήματος που μπορεί να εκφράζει ένας χρήστης. Οι αδυναμίες αυτές ευθύνονται για έναν αριθμό αρνητικών επιπτώσεων όπως η αδυναμία των προγραμματιστών να ελέγξουν το μεγάλο όγκο των δεδομένων, οι ανισότητες ή και οι κίνδυνοι λανθασμένων χειρισμών από τις εφαρμογές, λόγω των λανθασμένων εκτιμήσεων στις οποίες νωρίτερα έχουν προβεί. Αφετέρου, οι κίνδυνοι εκείνοι που απορρέουν από ανθρώπινες τακτικές, οι οποίες εστιάζουν σε επιδιώξεις πέραν της ψυχικής φροντίδας, όπως το κέρδος και ο έλεγχος του πληθυσμού. Μια σειρά μεθόδων που επιστρατεύονται για την επίτευξη αυτών των σκοπών, όπως η κατάχρηση των δεδομένων και η χειραγώγηση του κοινού, διακυβεύουν την ανθρώπινη ασφάλεια και έρχονται σε σύγκρουση με τις αρχές της Ηθικής και θεμελιώδεις έννοιες, όπως η δικαιοσύνη, η αξιοπρέπεια και η αυτονομία.

Πέρα από τα παραπάνω όμως, είναι το ζήτημα της αυθεντικότητας και η διαπίστωση του ότι αν και τα chatbot ψυχικής υγείας, μπορούν να προσομοιώσουν τις ανθρώπινες συμπεριφορές και να μιμηθούν τις δεξιότητες ενός ανθρώπου συμβούλου, δεν καταφέρνουν προς ώρας την ουσιαστική ευθυγράμμιση με τα χαρακτηριστικά της Συμβουλευτικής, τη γνησιότητα της θεραπευτικής σχέσης και επί της ουσίας, την προϋπόθεση του *συν + βουλεύεσθαι* που ορίζει τη συμβουλευτική διαδικασία.

Η ουσιώδης αυτή διαφορά, ανάμεσα στην ιδιότητα του νοήμονος όντος και στην προσομοίωση της ιδιότητας αυτής, γίνεται η αφορμή κρίσιμων ηθικών ζητημάτων. Αφενός, η μέχρι τώρα, «απομίμηση» της θεραπείας, δημιουργεί ερωτηματικά για το εάν η συνεχής εξέλιξη που σημειώνεται, μπορεί πράγματι, να θεωρηθεί πρόοδος. Αυτό γιατί μέχρι τώρα, αν και η χρήση των chatbot είναι αποτελεσματική στην ανίχνευση σημαντικών πληροφοριών, δε φαίνεται να βελτιώνει την ποιότητα των παρεχόμενων συμβουλευτικών υπηρεσιών ή έστω, να αντιστοιχεί στα ήδη υπάρχοντα πρότυπα και επίπεδα που επιτυγχάνει μέχρι σήμερα ένας άνθρωπος σύμβουλος. Επιπλέον, έννοιες ζωτικές, όπως η ενσυναίσθηση και η αποδοχή, που χτίζουν τη θεραπευτική σχέση, φαίνεται να αλλοιώνονται, κατά την προσπάθεια ενός chatbot να τις επιδείξει, μέσα από την ψυχρή αναπαραγωγή λέξεων και εκφράσεων που τις θυμίζουν.

Ένα όμως ιδιαίτερα σημαντικό ζήτημα που αναδύεται, σχετίζεται με την εξαπάτηση ενός χρήστη και τη δημιουργία της λανθασμένης αίσθησης από πλευράς του, πως η σχέση που αναπτύσσει με μια εφαρμογή είναι γνήσια και αυθεντική. Οι κίνδυνοι στους οποίους εκθέτει η παραπλάνηση αυτή, την ασφάλεια και την αυτονομία του χρήστη είναι πολλοί και είναι εξόχως προβληματικό το γεγονός πως μέχρι τώρα, η εκτεταμένη διάχυση και συνεχώς αυξανόμενη χρήση των chatbot ψυχικής υγείας δεν συνοδεύεται από την απαραίτητη ανάλυση τους. Αυτό γιατί, μια τέτοια ασυμμετρία δυσχεραίνει την διατύπωση των κατάλληλων ρυθμιστικών πλαισίων (νομικών και ηθικών), σχετικά με τη χρήση των συστημάτων της T.N. στη Συμβουλευτική. Πόσω μάλλον, τη στιγμή που το θέμα του ηθικού καταλογισμού των ευθυνών παραμένει αναπάντητο, λόγω της ασάφειας γύρω από το ηθικό καθεστώς το μηχανών. Ως αποτέλεσμα, ενόσω τα chatbot διεισδύουν καθημερινώς στη ζωή των ανθρώπων και αφήνονται ελεύθερα στο να διαχειρίζονται τον ευαίσθητο ψυχισμό τους, η απαραίτητη έμφαση στους κινδύνους που ενέχουν παραλείπεται ενώ παράλληλα, η απόδοση των ευθυνών για τυχόν αστοχίες παραμένει ουσιαστικά μετέωρη.

Με βάση τα όσα έχουμε δει λοιπόν, το κύριο συμπέρασμα στο οποίο καταλήγει η παρούσα μελέτη είναι ότι ναι μεν, η αξιοποίηση ενός chatbot από τους επαγγελματίες που στελεχώνουν μια γραμμή βοήθειας θα μπορούσε με τις κατάλληλες βελτιώσεις, να αναβαθμίσει σημαντικά το επίπεδο ποιότητας και λειτουργικότητας της. Για την ώρα όμως, η πλήρης ενσωμάτωσή της σε αυτό το πεδίο ή και η εξ ολοκλήρου αντικατάσταση ενός επαγγελματία από ένα εξελιγμένο σύστημα της συνομιλητικής Τ.Ν., φαίνεται να μην είναι εφικτές. Παράγοντες όπως, η φύση των γραμμών βοήθειας και συμβουλευτικής υποστήριξης και το περιεχόμενο των κλήσεων που δέχονται, οι ευάλωτες πληθυσμιακές ομάδες-στόχοι στις οποίες απευθύνονται και η πολυπλοκότητα την οποία αποκτούν, λόγω της ευαλωτότητας αυτής, έννοιες όπως η συναίνεση, η εμπιστοσύνη και η διαφάνεια, μαρτυρούν την επιτακτική ανάγκη της ηθικής πλαισίωσης της ανάπτυξης και εφαρμογής των συστημάτων της Τ.Ν., προτού αυτά αποτελέσουν αναπόσπαστο κομμάτι στη διαχείριση κλήσεων και περιστατικών. Η επιτακτικότητα αυτή, της αυστηρής ηθικής πλαισίωσης, διαφάνηκε ιδιαίτερα στο 4^ο κεφάλαιο της διατριβής, όπου αναδείχθηκαν μια σειρά από σοβαρούς ηθικούς προβληματισμούς, καθώς και ορισμένες εγγενείς, σε κάθε ηθικο-φιλοσοφική ανάλυση, δυσκολίες, που θα πρέπει να ληφθούν σοβαρά υπόψη κατά τη σύνταξη των απαραίτητων ρυθμιστικών πλαισίων σχετικά με την ενσωμάτωση της Τ.Ν. στην Συμβουλευτική.

Η αυστηρή ηθική πλαισίωση και η εφαρμογή κανονιστικών αρχών είναι απαραίτητες για την επίτευξη της ποιότητας των παρεχόμενων υπηρεσιών και την προστασία της αξίας και της ακεραιότητας που φέρει τόσο το επιστημονικό πεδίο της Συμβουλευτικής, όσο και οι τηλεφωνικές γραμμές βοήθειας και συμβουλευτικής υποστήριξης, ως υπηρεσία του Συστήματος Πρωτοβάθμιας Φροντίδας Υγείας. Κυρίως όμως, η ηθική πλαισίωση είναι απαραίτητη για την αποτροπή των τεράστιων κινδύνων που ενυπάρχουν και προς ώρας, απειλούν τον ευάλωτο ψυχισμό ενός χρήστη που προβαίνει στο πιο σημαντικό και ταυτόχρονα, πιο δύσκολο βήμα: να μοιραστεί τις σκέψεις και τα συναισθήματά του, επιδιώκοντας την αλλαγή.

Βιβλιογραφία

Αθανασοπούλου, Διονυσία. «Συστήματα αναγνώρισης προσώπου με χρήση Τεχνητής Νοημοσύνης και ζητήματα προστασίας δεδομένων προσωπικού χαρακτήρα». Μεταπτυχιακή διατριβή. Πανεπιστήμιο Πειραιώς, 2022. Διαθέσιμο στο: <https://dione.lib.unipi.gr/xmlui/handle/unipi/14835> [Πρόσβαση στις 14/04/2023].

ABC NEWS. “More calls to men's helpline after Whispering Wall suspected murder-suicide that killed baby Kobi.” Accessed November 17, 2023. <https://www.abc.net.au/news/2021-04-23/more-calls-to-mens-helpline-after-whispering-wall-murder-suicide/100090324>.

Adamopoulou, Eleni, Lefteris Moussiades. “An Overview of Chatbot Technology.” *AI&I* 584 (May, 2020): 373-383. https://link.springer.com/content/pdf/10.1007/978-3-030-49186-4_31.pdf.

Alexa. “What is Alexa?” Accessed August 28, 2023. <https://developer.amazon.com/en-US/alexa>.

Amazon. “Deep Breath.” Accessed September 25, 2023. <https://www.amazon.com/DrawnToDigital-Deep-Breath/dp/B01N1NFN44>.

American Counseling Association. “20/20: Consensus definition of Counseling.” Accessed June 26, 2023. <https://www.counseling.org/about-us/about-aca/20-20-a-vision-for-the-future-of-counseling/consensus-definition-of-counseling>.

American Psychiatric Association. “Diagnostic and Statistical Manual of Mental Disorders (DSM-5-TR).” Accessed September 7, 2023. <https://www.psychiatry.org/psychiatrists/practice/dsm>.

Anderson, Michael, Suzan L. Anderson, Alkis Gounaris, George Kosteletos. “Towards Moral Machines: A Discussion with Michael Anderson and Susan Leigh Anderson.” *Conatus* 6, no. 1 (2021): 177-202, <http://dx.doi.org/10.12681/cjp.26832>.

Anderson, Michael, Susan Leigh Anderson. “Machine Ethics: Creating an Ethical Intelligent Agent.” *AI Magazine* 28, no. 4 (Winter, 2007): 15-26. <https://doi.org/10.1609/aimag.v28i4.2065>.

Anderson, Susan Leigh. “Asimov’s “three laws of robotics” and machine metaethics.” *AI & Society* 22, no. 4 (Apr., 2008): 477-493. <https://link.springer.com/article/10.1007/s00146-007-0094-5>.

Βιρβιδάκης, Στέλιος. *Εισαγωγή στην ηθική φιλοσοφία: Κανονιστικές θεωρίες [πανεπιστημιακές σημειώσεις]*. ΕΚΠΑ. Τμήμα Φιλοσοφίας. ΠΜΣ Φιλοσοφίας: Εφαρμοσμένη Ηθική. 2020. Αθήνα.

Baumeister, H., L. Nowoczin, J. Lin, H. Seifferth, J. Seufert, K. Laubner, D.D. Ebert “Impact of an acceptance facilitating intervention on diabetes patients’ acceptance of Internet-based interventions for depression: A randomized controlled trial.” *Diabetes Research and Clinical Practice* 105, no. 1 (July, 2014): 30-39. <https://doi.org/10.1016/j.diabres.2014.04.031>.

BBC. “Child advice chatbots fail to spot sexual abuse.” Last modified December 11, 2018. <https://www.bbc.com/news/technology-46507900>.

Beauchamp, L. Tom, James F. Childress. *Principles of Biomedical Ethics*. 8th Edition. England: Oxford University Press, 2019.

Bedi, Gillinder, Facundo Carrillo, Guillermo Cecchi, Diego Fernandez-Slezak, Mariano Sigman, Natália Mota, Sidarta Ribeiro, Daniel Javitt, Mauro Copelli, Cheryl Corcoran. “Automated analysis of free speech predicts psychosis onset in high-risk youths.” *npj Schizophrenia* 1, no. 15030 (Aug. 2015): 1-7. <https://doi.org/10.1038/npjSchz.2015.30>.

Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation*. Canada: Batoche Books, 2000.

Bentham, Jeremy. *Panopticon: Or the Inspection House*. Dublin: Payne, 1791. https://www.google.gr/books/edition/Panopticon_Or_The_Inspection_House/SYc5A AAAcAAJ?hl.

Berry, Katherine, Amy Salter, Rohan Morris, Susannah James, Sandra Bucci. “Assessing Therapeutic Alliance in the Context of mHealth Interventions for Mental Health Problems: Development of the Mobile Agnew Relationship Measure (mARM) Questionnaire.” *Journal Of Medical Internet Research* 20, no. 4, e90 (April, 2018): 1-8. <https://www.jmir.org/2018/4/e90>.

Big Blue Data Academy. «Τι Είναι τα Large Language Models και Πώς Λειτουργούν;». Accessed August 22, 2023. <https://bigblue.academy/gr/large-language-models>.

Black Dog Institute. MyCompass. “myCompass: A customisable self-help tool for your mental health.” Accessed September 30, 2023. <https://www.mycompass.org.au/>.

Blasimme, Alessandro, Effy Vayena. “The Ethics of AI in Biomedical Research, Patient Care, and Public Health.” In *The Oxford Handbook of Ethics of AI*, ed. Markus D. Dubber, Frank Pasquale, Sunit Das, 707-711. New York, Oxford University Press 2020.

Brenton Harry, Marco Gillies, Daniel Ballin, David Chatting. “The uncanny valley: Does it exist?” Workshop on Human-Animated Character Interaction - HCI 2005 (Sep., 2005): 3. <http://www.davidchatting.com/research/uncanny-valley-hci2005.pdf>.

Bringsjord, Selmer, Naveen Sundar Govindarajulu. “Artificial Intelligence.” Stanford Encyclopedia of Philosophy. Accessed July 30, 2023. <https://plato.stanford.edu/archives/fall2022/entries/artificial-intelligence/>.

BJGP Life. “The utilitarian case for AI-mediated clinical decision-making.” Accessed December 3, 2023. <https://bjgplife.com/the-utilitarian-case-for-ai-mediated-clinical-decision-making/>.

Burton, Emanuelle, Judy Goldsmith, Sven Koenig, Benjamin Kuipers, Nicholas Mattei, Toby Walsh. “Ethical Considerations in Artificial Intelligence Courses.” *AI Magazine* 38, no. 2 (Jan., 2017): 1–29.

https://www.researchgate.net/publication/312949692_Ethical_Considerations_in_Artificial_Intelligence_Courses

Γεωργίου, Σταυρούλα. «Αυτόματη αναγνώριση προσώπου στους δημόσιους χώρους: Μεταξύ πραγματικότητας και δυστοπίας». Syntagmawatch.gr. Προσπελάστηκε στις 4 Νοεμβρίου 2023. <https://www.syntagmawatch.gr/trending-issues/aytomati-anagnwrisi-proswpou-stous-dhmosious-xwrous-metaxy-pragmatikothtas-kai-dystopias/>.

Γεωργούλη, Κατερίνα. *Τεχνητή Νοημοσύνη. Μια εισαγωγική προσέγγιση*. Αθήνα: Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών, 2015. <https://repository.kallipos.gr/bitstream/11419/3381/3/%CE%A4%CE%B5%CF%87%CE%BD%CE%B7%CF%84%CE%AE%20%CE%9D%CE%BF%CE%B7%CE%BC%CE%BF%CF%83%CF%8D%CE%BD%CE%B7.pdf>.

Γιαννούλη, Βασιλική. *Ψυχοδιαγνωστικά Μέσα. Ενότητα 2: Ψυχομετρία* [πανεπιστημιακές σημειώσεις]. Πανεπιστήμιο Μακεδονίας. Τμήμα Εκπαιδευτικής και Κοινωνικής Πολιτικής. Θεσσαλονίκη. <http://opencourses.uom.gr/assets/site/public/778/623-Psychodiagnostika-Mesa-02-Giannouli.pdf>.

Γούναρης, Άλκης, Γεώργιος Κωστελέτος. *Τεχνητή Νοημοσύνη και Ηθική. Ενότητα 1^η: Εισαγωγική Συνάντηση* [πανεπιστημιακές σημειώσεις]. ΕΚΠΑ. Τμήμα Φιλοσοφίας. ΠΜΣ Φιλοσοφίας: Εφαρμοσμένη Ηθική. Χειμερινό εξάμηνο 2022-2023, Αθήνα.

Γούναρης, Άλκης, Γεώργιος Κωστελέτος. *Τεχνητή Νοημοσύνη και Ηθική. Ενότητα 7^η: Υπερβαίνοντας τις ανθρώπινες δυνατότητες* [πανεπιστημιακές σημειώσεις]. ΕΚΠΑ. Τμήμα Φιλοσοφίας. ΠΜΣ Φιλοσοφίας: Εφαρμοσμένη Ηθική. Χειμερινό εξάμηνο 2022-2023, Αθήνα.

Γραμμή ζωής – Κοινωφελές Μη Κερδοσκοπικό Σωματείο Τηλεειδοποίησης και Βοήθειας στο Σπίτι, «Η Εθνική Τηλεφωνική Γραμμή SOS 1065». Προσπελάστηκε στις 16 Σεπτεμβρίου 2023, <http://www.lifelinehellas.gr/%cf%80%ce%bf%ce%b9%ce%bf%ce%b9-%ce%b5%ce%af%ce%bc%ce%b1%cf%83%cf%84%ce%b5/>.

Capurro, Daniel, Simon Coghlan, Douglas Pires. “Preventing Digital Overdiagnosis.” *JAMA The Journal of the American Medical Association* 327, no. 6 (Feb., 2022): 525-526. https://www.researchgate.net/publication/358016994_Preventing_Digital_Overdiagnosis/citations.

Cass. “Bringing mental health to everyone.” Accessed July 9, 2023. <https://www.x2ai.com/>.

CBS News. “In a crisis, Siri and Cortana may not have your back.” Accessed September 19, 2023. <https://www.cbsnews.com/news/health-crisis-siri-and-cortana-may-not-have-your-back/>.

Charité – Universitätsmedizin Berlin. “Forschungsprojekte und Aktivitäten mit COVID-19 Bezug.” Accessed September 28, 2023. [https://iph.charite.de/forschung/aktivitaeten mit covid 19 bezug/](https://iph.charite.de/forschung/aktivitaeten%20mit%20covid%2019%20bezug/).

Clarke, Janine, Judith Proudfoot, Alexis Whitton, Mary-Rose Birch, Megan Boyd, Gordon Parker, Vijaya Manicavasagar, Dusan Hadzi-Pavlovic, Andrea Fogarty. “Therapeutic Alliance With a Fully Automated Mobile Phone and Web-Based Intervention: Secondary Analysis of a Randomized Controlled Trial.” *Journal of Medical Internet Research in Mental Health* 25, no. 3 (1) e10 (Feb., 2016): 1-12. <https://mental.jmir.org/2016/1/e10/>.

CNN Greece. «Πάτρα: Από τον πρώτο θάνατο στη σύλληψη της Ρούλας Πισπιρίγκου - Οι ημερομηνίες ορόσημα της υπόθεσης». Προσπελάστηκε 29 Σεπτεμβρίου 2023. <https://www.cnn.gr/ellada/story/306884/patra-apo-to-thanato-sti-syllipsi-tis-roylas-pispirigkoy-oi-imerominies-orosimo-tis-ypothesis>.

Coeckelbergh, Mark. “Just Machines?” In *AI Ethics*, 47-62. Cambridge, Massachusetts: MIT Press, 2020.

Coghlan, Simon, Kobi Leins, Susie Sheldrick, Marc Cheong, Piers Gooding, Simon D’Alfonso. “To chat or bot to chat: Ethical issues with using chatbots in mental health.” *Digital Health* 9 (June, 2023): 1-11. <https://journals.sagepub.com/doi/10.1177/20552076231183542>.

Conrad, Peter, Thomas Mackie, Ateev Mehrotra. “Estimating the costs of medicalization.” *Social Science & Medicine* 70, no. 12 (June, 2010): 1943 – 1947. <https://www.sciencedirect.com/science/article/pii/S0277953610001784>.

Coperland, Jack B. “The Modern History of Computing.” Stanford Encyclopedia of Philosophy. Accessed August 20, 2023. <https://plato.stanford.edu/entries/computing-history/>.

Corcoran, M. Cheryl, Facundo Carrillo, Diego Fernández-Slezak, Gillinder Bedi, Casimir Klim, Daniel C. Javitt, Carrie E. Bearden, Guillermo A. Cecchi, “Prediction of psychosis across protocols and risk cohorts using automated language analysis,” *World Psychiatry* 17, no. 1 (Feb. 2018): 67-75, <https://doi.org/10.1002/wps.20491>.

Council of Europe. “Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine.” Accessed November 17, 2023. <https://rm.coe.int/168007cf98>.

Cranen, Karlijn, Constance H.C. Drossaert, Evelien S. Brinkman, Annemarie L.M. Braakman-Jansen, Maarten J. Ijzerman, Miriam M.R. Vollenbroek-Hutten. “An exploration of chronic pain patients' perceptions of home telerehabilitation services.” *Health Expert* 15, no. 4 (Dec. 2012): 339-350. doi: 10.1111/j.1369-7625.2011.00668.x.

Cui, Tingru, Xinwei Wang, Jianzhong Qi. “Designing Anthropomorphic Therapeutic Conversational Agents: An Uncanny Valley Perspective.” *AISeL (AISeL)* (June, 2021): 1-2.

https://aisel.aisnet.org/ecis2021_rip/?utm_source=aisel.aisnet.org%2Fecis2021_rip%2F30&utm_medium=PDF&utm_campaign=PDFCoverPages.

Δελατόλας, Π. Γεώργιος. «Επεξεργασία φυσικής γλώσσας». Διπλωματική διατριβή, Πανεπιστήμιο Πατρών, 2021.

Δραγώνα-Μονάχου, Μυρτώ. «Ηθική και βιοηθική». *Επιστήμη και Κοινωνία: Επιθεώρηση Πολιτικής και Ηθικής Θεωρίας* 8 (2015): 1-26. <https://doi.org/10.12681/sas.715>.

D'Alfonso, Simon. "AI in mental health." *Current Opinion in Psychology* 36 (June 2020): 112-117. <https://doi.org/10.1016/j.copsyc.2020.04.005>.

D'Alfonso, Simon, Olga Santesteban-Echarri, Simon Rice, Greg Wadley, Reeva Lederman, Christopher Miles, John Gleeson, Mario Alvarez-Jimenez. "Artificial Intelligence-Assisted Online Social Therapy for Youth Mental Health." *Frontiers in Psychology* 8, Art. 796 (June, 2017): 1-13. <https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00796/full>.

DatAnalysis. "Patient Health Questionnaire [PHQ-9]." Accessed August 23, 2023. <https://datanalysis.net/questionnaires/patient-health-questionnaire-phq/>.

Davis, H.Mark. "A Multidimensional Approach to Individual Differences in Empathy." *JSAS Catalog of Selected Documents in Psychology* 10, p. 85 (Jan., 1980): 1-19. https://www.uv.es/~friasnav/Davis_1980.pdf;

Davis, H. Mark. "Measuring individual differences in empathy: Evidence for a multidimensional approach." *Journal of Personality and Social Psychology* 44, no. 1 (Jan., 1983): 113-126. https://www.researchgate.net/publication/215605832_Measuring_individual_differences_in_empathy_Evidence_for_a_multidimensional_approach.

Dawoodbhoy, Fatema Mustansir, Jack Delaney, Paulina Cecula, Jiakun Yu, Iain Peacock, Joseph Tan, Benita Cox. "AI in patient flow: applications of artificial intelligence to improve patient flow in NHS acute mental health inpatient units." *Heliyon* 7, no. 5 (May 2021): 1-18. <https://doi.org/10.1016/j.heliyon.2021.e06993>.

De Boer, Tim Rens, Saskia Mirelle, Sandjai Bhulai, Renske Gilissen, Rob van der Mei. "Forecasting call and chat volumes at online helplines for mental health." *BMC Public Health* 23, no. 984 (May, 2023): 1-13. <https://doi.org/10.1186/s12889-023-15887-2>.

DeepMood®. "Emotions matter." Accessed November 10, 2023. https://deepmood.io/en_US/.

De la Torre Díez, Isabel, Susel Góngora Alonso, Sofiane Hamrioui, Eduardo Motta Cruz, Lola Morón Nozaleda, Manuel A. Franco. "IoT-Based Services and Applications for Mental Health in the Literature." *Journal of Medical Systems* 43, no. 1 (Dec. 2018): 1-6. [10.1007/s10916-018-1130-3](https://doi.org/10.1007/s10916-018-1130-3).

Diakopoulos, Nicholas. "Transparency." In *The Oxford Handbook of Ethics of AI*, ed. Markus D. Dubber, Frank Pasquale, Sunit Das, 197-213. New York, Oxford University Press, 2020.

D-ID. “The Digital People Platform.” Accessed October 6, 2023. <https://chat.d-id.com/>.

Ελληνική Δημοκρατία. Υπουργείο Υγείας. «Γραμμή SOS 1056 για Παιδιά, Έφηβους και Γονείς». Προσπελάστηκε στις 16 Σεπτεμβρίου, 2023. <https://www.moh.gov.gr/articles/citizen/xrhsima-thlefwna-amp-dieythynseis/199-grammh-sos-1056-gia-paidia-efhboys-kai-goneis>.

Ελληνική Δημοκρατία. Υπουργείο Υγείας. «eHealth - Ηλεκτρονική Υγεία». Προσπελάστηκε στις 16 Σεπτεμβρίου, 2023. <https://www.moh.gov.gr/articles/ehealth/>.

Ελληνική Δημοκρατία. Υπουργείο Υγείας. «mHealth - Υγεία εν Κινήσει». Αναρτήθηκε στις 2 Σεπτεμβρίου, 2018. <https://www.moh.gov.gr/articles/ehealth/5688-mhealth>.

Ελληνική Δημοκρατία. Υπουργείο Υγείας. «Τηλεφωνική Γραμμή Ψυχοκοινωνικής Υποστήριξης 10306». Προσπελάστηκε στις 16 Σεπτεμβρίου, 2023. <https://www.moh.gov.gr/articles/citizen/xrhsima-thlefwna-amp-dieythynseis/10714-thlefwnikh-grammh-psykokoinwnikhs-yposthrikshs-10306>.

Ελληνική Δημοκρατία. Υπουργείο Υγείας. «1031 Τηλεφωνική Γραμμή SOS του ΟΚΑΝΑ». Προσπελάστηκε στις 16 Σεπτεμβρίου, 2023. <https://www.moh.gov.gr/articles/citizen/xrhsima-thlefwna-amp-dieythynseis/197-1031-thlefwnikh-grammh-sos-toy-okana>.

Ελληνική Δημοκρατία. Υπουργείο Υγείας. «1145 Τηλεφωνική Γραμμή Ψυχολογικής Υποστήριξης ΙΘΑΚΗ - ΚΕ.Θ.Ε.Α.». Προσπελάστηκε στις 16 Σεπτεμβρίου, 2023. <https://www.moh.gov.gr/articles/citizen/xrhsima-thlefwna-amp-dieythynseis/204-1145-thlefwnikh-grammh-psykologikhs-yposthrikshs-ithakh-ke-th-e-a>.

Ελληνική Δημοκρατία. Υπουργείο Υγείας. «197 Εθνικό Κέντρο Κοινωνικής Αλληλεγγύης». Προσπελάστηκε στις 16 Σεπτεμβρίου, 2023. <https://www.moh.gov.gr/articles/citizen/xrhsima-thlefwna-amp-dieythynseis/201-197-ethniko-kentro-koinwnikhs-allhleegyhs>.

Ελληνική Δημοκρατία. Υπουργείο Υγείας. «210 6401200 Ελληνική Αντικαρκινική Εταιρεία». Προσπελάστηκε στις 16 Σεπτεμβρίου, 2023. <https://www.moh.gov.gr/articles/citizen/xrhsima-thlefwna-amp-dieythynseis/202-210-6401200-ellhnikh-antikarkinikh-etaireia>.

Εταιρεία Alzheimer Αθηνών. «Γραμμή Βοήθειας για την Άνοια 1102». Προσπελάστηκε στις 16 Σεπτεμβρίου, 2023. <https://alzheimerathens.gr/helpline/>.

Ede, Lisa. “Is Rogerian Rhetoric Really Rogerian?” *Rhetoric Review* 3, no. 1 (Sep. 1984): 40–48. <http://www.jstor.org/stable/465729>.

Elder, Alexis. “Robot Friends for Autistic Children: Monopoly Money or Counterfeit Currency?”. In *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence*, ed. Patrick Lin, Keith Abney, Ryan Jenkins, 113-126. New York: Oxford Academic, 2017. <https://doi.org/10.1093/oso/9780190652951.003.0008>.

EUR-Lex. Access to European Union Law. «Κανονισμός (ΕΕ) 2016/679 του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου της της 27ης Απριλίου 2016».

Προσπελάστηκε στις 15 Νοεμβρίου 2023. <https://eur-lex.europa.eu/legal-content/EL/TXT/?uri=celex%3A32016R0679>.

Ζαχαράκη, Βαρβάρα. «Τεχνητή Νοημοσύνη, λήψη αποφάσεων με χρήση αλγορίθμων, ηθική και δεοντολογία στο Δημόσιο τομέα». Μεταπτυχιακή διατριβή. Πανεπιστήμιο Δυτικής Αττικής, 2022). Διαθέσιμο στο: <https://polynoe.lib.uniwa.gr/xmlui/bitstream/handle/11400/3345/%ce%92%ce%91%ce%a1%ce%92%ce%91%ce%a1%ce%91%20%ce%96%ce%91%ce%a7%ce%91%ce%a1%ce%91%ce%9a%ce%97%ce%94%ce%9c2088.pdf?sequence=1&isAllowed=y>.

Ζέρβα, Σοφία, Ελένη Βατούγιου, Ιωάννης Αποστολάκης. «Κινητή Υγεία (M-health) και Πρωτοβάθμια Φροντίδα Υγείας. Τηλεφωνικές Γραμμές Άμεσης Βοήθειας και Συμβουλευτικής Υποστήριξης». Το βήμα του Ασκληπιού 19, Αρ. 2 (Απρίλιος - Ιούνιος 2020): 89-106. DOI: 10.5281/zenodo.3736786.

Ζηκόπουλος, Φίλιππος. «Η χρήση της τεχνητής νοημοσύνης και της τεχνολογίας στην πρόληψη και στην βελτίωση της υγείας». Διπλωματική διατριβή, Πανεπιστήμιο Δυτικής Αττικής, 2021.

Facebook. “Hi, I’m Joybot.” Accessed November 6, 2023. <https://www.facebook.com/hellojoyai/>.

Fitzpatrick, Kara Kathleen, Alison Darcy, Molly Vierhile. “Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial.” *JMIR Mental Health* 4, no. 2 (Apr.-Jun. 2017): 1-11. doi: [10.2196/mental.7785](https://doi.org/10.2196/mental.7785).

Floridi, Luciano, Sanders, J.W. “On the Morality of Artificial Agents.” *Minds and Machines* 14, no. 3 (August, 2004): 349-379. https://www.academia.edu/255823/On_the_morality_of_artificial_agents.

Fub NHS Stuff. “The SAFER Patient Flow Bundle.” Accessed July 10, 2023. <https://fabnhsstuff.net/fab-stuff/28574-the-safer-patient-flow-bundle>.

Fub NHS Stuff. “The #last1000days from 3red2green.” Accessed July 10, 2023. <https://fabnhsstuff.net/fab-stuff/the-last1000days-from-red2green>.

Fulmer, Russel, Angela Joerin, Breanna Gentile, Lysanne Lakerink, Michiel Rauws. “Using Psychological Artificial Intelligence (Tess) to Relieve Symptoms of Depression and Anxiety: Randomized Controlled Trial.” *JMIR Ment Health* 5, no. 4 (Oct.-Dec. 2018): 1-15. DOI: 10.2196/mental.9782.

Gillon, Raanan “Primum non nocere and the principle of non-maleficence.” *British Medical Journal (Clinical research ed.)* 13, no. 291 (6488) (July, 1985): 130-131. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1416256/pdf/bmjcred00456-0042.pdf>.

Gooding, Piers. “Mapping the rise of digital mental health technologies: Emerging issues for law and society.” *International Journal of Law and Psychiatry* 67 (Oct, 2019): 1-11. <https://doi.org/10.1016/j.ijlp.2019.101498>.

Gordon, P. Kenneth. "Attitudes towards computerized CBT for depression amongst a student population." *Behavioral and Cognitive Psychotherapy* 3, no. 4 (2007): 421–430.

https://www.researchgate.net/publication/232021580_Attitudes_Towards_Computerized_CBT_for_Depression_Amongst_a_Student_Population.

Gounaris, Alkis, George Kosteletos. "Licensed to Kill: Autonomous Weapons as Persons and Moral Agents." In *Personhood*, edited by Dragan Prole & Goran Rujević, 137-189. Novi Sad: Filozofski Fakultet and Athens: The NKUA Applied Philosophy Research Lab Press. 2020.

https://www.researchgate.net/publication/348506490_Personhood.

Guan, Jian. "Artificial Intelligence in Healthcare and Medicine: Promises, Ethical Challenges and Governance." *Chinese Medical Sciences Journal* 34, no. 2 (June 2019): 76-83. https://www.researchgate.net/publication/334574521_Artificial_Intelligence_in_Healthcare_and_Medicine_Promises_Ethical_Challenges_and_Governance/citations.

Gunkel, David. "Do Machines Have Rights? Ethics in the Age of Artificial Intelligence." *Aurora Online*, Is. 2014 (Oct., 2014): 1-12, <http://http://aurora.icaap.org/index.php/aurora/article/view/92/114>.

Gunkel, J. David. "Mind the gap: responsible robotics and the problem of responsibility." *Ethics and Information Technology* 22 (Dec., 2020): 307–320. <https://doi.org/10.1007/s10676-017-9428-2>.

Gun Shin Ying, Nickolai Titov, Gavin Andrews. "Acceptability of Internet treatment of anxiety and depression." *Australas Psychiatry* 19, no. 3 (June, 2011): 259-264. doi: 10.3109/10398562.2011.562295. PMID: 21682626.

Hamet, Pavel, Johanne Tremblay. "Artificial intelligence in medicine." *Metabolism Clinical Experimental* 69 (Apr., 2017): 36-40. <http://dx.doi.org/10.1016/j.metabol.2017.01.011>.

Habermas, Juergen. *On the Pragmatics of Social Interaction: Preliminary Studies in the Theory of Communicative Action*. Trans Barbara Fultner. Cambridge, Massachusetts: The MIT Press, 2001.

Habermas, Juergen. *The Theory of Communicative Action. Lifeworld and System: A Critique of Functionalist Reason* (Vol. 2). Trans. Thomas McCarthy. Cambridge: Polity Press, 1989.

Habermas, Juergen. *The Theory of Communicative Action: Reason and the Rationalization of Society* (Vol. 1). Trans. Thomas McCarthy. Cambridge: Polity Press, 2015.

Habermas, Juergen. *Truth and Justification*. Trans. and ed. by Barbara Fultner. Cambridge, Massachusetts: The MIT Press, 2001.

Hart, William. "Nussbaum, Kant and Conflicts between Duties." *Philosophy* 73, no. 4 (Oct., 1998): 609-618. https://www.researchgate.net/publication/231935160_Nussbaum_Kant_and_Conflicts_between_Duties.

Haselager, Pim, Giulio Mecacci. “Importance of Expiry Dates: Evaluating the Societal Impact of AI-Based Neuroimaging.” In *Artificial Intelligence in Brain and Mental Health: Philosophical, Ethical & Policy Issues*, ed. by Fabrice Jotterand, Marcello Ienca, 205-216. Switzerland: Springer, 2021.

Herbert, Gary. “Immanuel Kant: On Treating Persons as Persons.” *The Personalist Forum* 15, no. 2 (Fall 1999): 247–56. <http://www.jstor.org/stable/20708820>.

Ιωσηφίδης, Αλέξανδρος. «Δεοντοκρατία και ωφελμισμός στη Βιοηθική: Ομοιότητες και διαφορές». Μεταπτυχιακή διατριβή, Δημοκρίτειο Πανεπιστήμιο Θράκης, 2018.

IEEE Spectrum. “The Uncanny Valley: The Original Essay by Masahiro Mori: “The Uncanny Valley” by Masahiro Mori is an influential essay in robotics. This is the first English translation authorized by Mori.” Accessed October 20, 2023. <https://spectrum.ieee.org/the-uncanny-valley>.

Inkster, Becky, Shubhankar Sarda, Vinod Subramanian. “An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study.” *JMIR Mhealth Uhealth* 6, no. 11 (Nov. 2018): 1-14. doi: 10.2196/12106.

iPhoneA2. «Τι είναι το Siri και πώς λειτουργεί; Όλα όσα πρέπει να ξέρετε». Προσπελάστηκε στις 28 Αυγούστου 2023. <https://iphonea2.com/el/como-funciona-siri/>.

Jacobson, C. Nicholas, Matthew D. Nemesure. “Using Artificial Intelligence to Predict Change in Depression and Anxiety Symptoms in a Digital Intervention: Evidence from a Transdiagnostic Randomized Controlled Trial.” *Psychiatry Research* 295, 113618 (Jan., 2021): 1-4, <https://doi.org/10.1016/j.psychres.2020.113618>.

Κακογιάννης, Α. Ανδρέας. «Ανάλυση συναισθήματος σε δεδομένα του Twitter με τεχνικές NLP». Διπλωματική διατριβή, Πανεπιστήμιο Μακεδονίας, 2022.

Καλαντζή-Αζίζι, Αναστασία, Μαρία Μαλιώτη-Λοΐζου. «Η θεραπευτική σχέση στην κλινική και συμβουλευτική ψυχολογία». *Ψυχολογία: The Journal of the Hellenic Psychological Society* 26, no. 2 (Ιούνιος, 2021): 1-7. https://www.researchgate.net/publication/352402734_E_therapeutike_schese_sten_klinike_kai_symbouleutike_psychologia.

Καντ, Ιμμάνουελ. *Η μεταφυσική των ηθών*. Μετάφραση Κώστας Ανδρουλιδάκης. Αθήνα: Σμίλη, 2013.

Καντ, Ιμμάνουελ. *Τα θεμέλια της Μεταφυσικής των ηθών*. Μετάφραση Γ. Τζαβάρας. Αθήνα: ΔΩΔΩΝΗ, 1984.

Κατσικονούρη, Ευφροσύνη, Ευθυμία Μανίτσα. «Δεξιότητες συμβουλευτικής ‘αντίδοτο’ στην κρίση». *Ερευνώντας τον κόσμο του παιδιού*, no. 13 (2014): 177-187.

Κέντρο Θεραπείας Εξαρτημένων Ατόμων ΚΕΘΕΑ. «Γραμμή Βοήθειας 1114». Προσπελάστηκε στις 16 Σεπτεμβρίου, 2023. <https://www.kethea.gr/kethea/therapeftika-programmata/grammi-voitheias-1114/>.

Κιβρακίδου, Άννα. «Προσβολές της ανθρώπινης αξιοπρέπειας». Μεταπτυχιακή διατριβή, ΑΠΘ, 201717-18;

Κιτρομηλίδης, Μ. Πασχάλης. Νεότερη πολιτική θεωρία. Αθήνα: Νομική Βιβλιοθήκη, 6^η έκδοση, 2016.

Κιτρομηλίδης, Μ. Πασχάλης. *Πολιτικοί στοχαστές των νεότερων χρόνων. Βιογραφικές και ερμηνευτικές προσεγγίσεις*. Αθήνα: Πανεπιστημιακές εκδόσεις Κρήτης, 2020.

Κλίμακα. «Εάν σκέφτεστε την αυτοκτονία». Προσπελάστηκε στις 16 Σεπτεμβρίου, 2023. <http://suicide-help.gr/ean-skefteste-tin-autoktonia/>.

Κοσμά, Ευαγγελία. «Η Σχέση της Ενσυναίσθησης και της Ενεργητικής Ενσυναίσθητικής Ακρόασης της γονικής σχέσης με την Ψυχοπαθολογία». Διπλ. Εργασία, ΕΚΠΑ, 2020.

Κώστα, Σοφία, Χρήστος Ράπτης. «Η Συμβουλευτική στις μονάδες υγείας: Τεχνικές για τους διοικητικούς». Πτυχιακή εργασία, ΤΕΙ Καλαμάτας, 2003.

Κωστελέτος Γεώργιος. *Τεχνητή Νοημοσύνη και Ηθική. Ενότητα 2^η: Το πρόγραμμα της Τεχνητής Νοημοσύνης* [πανεπιστημιακές σημειώσεις]. ΕΚΠΑ. Τμήμα Φιλοσοφίας. ΠΜΣ Φιλοσοφίας: Εφαρμοσμένη Ηθική. Χειμερινό εξάμηνο 2022-2023, Αθήνα.

Κωστελέτος, Γεώργιος. *Τεχνητή Νοημοσύνη και Ηθική. Ενότητα 9^η: Έρωτας και Πόλεμος στην εποχή της Τεχνητής Νοημοσύνης* [πανεπιστημιακές σημειώσεις]. ΕΚΠΑ. Τμήμα Φιλοσοφίας. ΠΜΣ Φιλοσοφίας: Εφαρμοσμένη Ηθική. Χειμερινό εξάμηνο 2022-2023, Αθήνα.

Κωστελέτος, Γεώργιος. *Τεχνητή Νοημοσύνη και Ηθική. Ενότητα 10^η: Άνθρωπος και μηχανή: Από τους ψυχολογικούς μηχανισμούς στα κοινωνικά και ηθικά ζητήματα* [πανεπιστημιακές σημειώσεις]. ΕΚΠΑ. Τμήμα Φιλοσοφίας. ΠΜΣ Φιλοσοφίας: Εφαρμοσμένη Ηθική. Χειμερινό εξάμηνο 2022-2023, Αθήνα.

Kant, Immanuel, *Groundwork for the metaphysics of morals*, trans. and ed. by Allen W. Wood. New Haven: Yale University Press, 2002.

Kerstein, J. Samuel. “Death, Dignity, and Respect.” *Social Theory and Practice* 35, no. 4 (October 2009): 505–30. <http://www.jstor.org/stable/23562118>.

Khawaja, Zoha, Jean-Christophe Bélisle-Pipon. “Your robot therapist is not your therapist: understanding the role of AI-powered mental health chatbots.” *Frontiers in Digital Health* 5, 1278186 (Nov., 2023): 2. <file:///C:/Users/estia/Downloads/fdgh-05-1278186.pdf>.

Kodiko.gr. «ΝΟΜΟΣ ΥΠ’ ΑΡΙΘΜ. 4486 ΦΕΚ Α’ 115/07.08.2017». Προσπελάστηκε στις 14 Νοεμβρίου 2023. <https://www.kodiko.gr/nomothesia/document/274963/nomos-4486-2017>.

Koko. “Free mental health support for millions of young people struggling online.” Accessed October 12, 2023. <https://www.kokocares.org/>.

Koutsouleris, Nikolaos, Tobias Hauser, Vasilisa Skvortsova, Munmun De Choudhury. "From promise to practice: towards the realisation of AI-informed mental health care." *Lancet Digital Health* 4 (Oct. 2022): 829-840. [https://doi.org/10.1016/S2589-7500\(22\)00153-4](https://doi.org/10.1016/S2589-7500(22)00153-4).

Kretzschmar, Kira, Holly Tyroll, Gabriela Pavarini, Arianna Manzini, Ilina Singh. "Can Your Phone Be Your Therapist? Young People's Ethical Perspectives on the Use of Fully Automated Conversational Agents (Chatbots) in Mental Health Support." *Biomedical Informatics Insights* 11 (March, 2019): 1-9. <https://journals.sagepub.com/doi/10.1177/1178222619829083>.

Λάγιος, Αλέξανδρος. «Ένσυναίσθηση και Τεχνητή Νοημοσύνη». Μεταπτυχιακή διατριβή, Εθνικό Μετσόβειο Πολυτεχνείο, 2016. <https://dspace.lib.ntua.gr/xmlui/handle/123456789/43823>.

LaBossiere, Michael. "Testing the Moral Status of Artificial Beings; or "I'm Going to Ask you Some Questions...". In *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence*, edited by Patrick Lin, Keith Abney, and Ryan Jenkins, 193-306. Oxford: Oxford University Press, 2017.

LaVaque-Manty, Mika. "Kant's Children." *Social Theory and Practice* 32, no. 3 (July 2006): 365–88. <http://www.jstor.org/stable/23558615>.

Lawrence, J. Dana. "The Four Principles of Biomedical Ethics: A Foundation for Current Bioethical Debate." *Journal of Chiropractic Humanities* 14 (December, 2007): 34-40. [https://doi.org/10.1016/S1556-3499\(13\)60161-8](https://doi.org/10.1016/S1556-3499(13)60161-8).

Lawton, George. "What is generative AI? Everything you need to know. " Techtargget: Enterprise AI. Accessed August 20, 2023. <https://www.techtargget.com/searchenterpriseai/definition/generative-AI>.

Lederman, Reeva, D'Alfonso, Simon. "The Digital Therapeutic Alliance: Prospects and Considerations." *JMIR Ment Health* 8, no. 7, e31385 (July, 2021): 1-4. <https://mental.jmir.org/2021/7/e31385/>.

Liao, S. Matthew. "The Moral Status and Rights of Artificial Intelligence." In *Ethics of Artificial Intelligence*, edited by Matthew S. Liao, 480-503. Oxford: Oxford University Press, 2020.

Μαζί για το Παιδί. «Γραμμή 115 25 της Ένωσης Μαζί για το Παιδί». Προσπελάστηκε στις 16 Σεπτεμβρίου, 2023. <https://mazigiatorpaidi.gr/en/articles/grammi-115-25-tis-enosis-mazi-gia-to-paidi>.

Μαλικιώση - Λοΐζου, Μαρία. «Η Συμβουλευτική Ψυχολογία στην Ελλάδα σήμερα». *Hellenic Journal of Psychology*, Vol. 8 (Ιαν. 2011): 266-288. <file:///C:/Users/estia/Downloads/Malikiosi-Loizou8cHellenicJournalofPsychology.pdf>.

Μαλικιώση - Λοΐζου, Μαρία. *Η συμβουλευτική ψυχολογία στην εκπαίδευση. Από την θεωρία στην πράξη*. Αθήνα: Ελληνικά Γράμματα, 2008.

Μαλικιώση - Λοΐζου, Μαρία. *Συμβουλευτική Ψυχολογία*. Αθήνα: Ελληνικά Γράμματα, 2006.

Μανίκα, Λαμπρινή. «Το πρόσωπο του Σωκράτη μέσα από την οπτική του Γρηγόρη Βλαστού & του Olof Gigon». Διπλωμ. εργασία, ΑΠΘ, 2013.

Μηχανή του Χρόνου. «Βαγγέλης Γιακουμάκης: Το χρονικό της εξαφάνισης που συγκλόνησε όλη την Ελλάδα...». Προσπελάστηκε 29 Σεπτεμβρίου 2023. <https://www.mixanitouxronou.com.cy/epikerotita/vangelis-giakoumakis-to-chroniko-tis-exafanisis-pou-sigklonise-oli-tin-ellada/>.

Μυρίσας, Γεώργιος. «Αναγνώριση τάσεων αυτοκτονίας σε κοινωνικά δίκτυα». Διπλωματική διατριβή, Ελληνικό Ανοικτό Πανεπιστήμιο, 2021.

Mancone, Stefania, Pierluigi Diotaiuti, Giuseppe Valente, Stefano Corrado, Fernando Bellizzi, Guilherme Torres Vilarino, Alexandro Andrade. “The Use of Voice Assistant for Psychological Assessment Elicits Empathy and Engagement While Maintaining Good Psychometric Properties.” *Behavioral Sciences* 13, no. 7, 550 (July, 2023): 1-19. <https://www.mdpi.com/2076-328X/13/7/550>.

Marchant, Gary, Lucille Nalbach Tournas. “Matter Over Mind: Liability Considerations Surrounding Artificial Intelligence in Neuroscience.” In *Artificial Intelligence in Brain and Mental Health: Philosophical, Ethical & Policy Issues*, ed. by Fabrice Jotterand, Marcello Ienca, 233-246. Switzerland: Springer, 2021.

Masud, T. Mohammed, Mohammed A. Mamun, K. Thapa, D.H. Lee, Mark D. Griffiths, S.-H. Yang. “Unobtrusive monitoring of behavior and movement patterns to detect clinical depression severity level via smartphone.” *Journal of Biomedical Informatics* 103, no. 103371 (March 2020): 1-10. <https://doi.org/10.1016/j.jbi.2019.103371>.

Mathur, Maya, David Reichling. “Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley.” *Cognition* 146 (Jan., 2016): 22-32. <https://www.sciencedirect.com/science/article/pii/S0010027715300640>.

McLeod, John. *Εισαγωγή στην συμβουλευτική*. Αθήνα: Μεταίχμιο, 2003.

McLeod, John. *The Counsellor's Workbook: Developing a Personal Approach*. 2nd ed. England: Open University Press: The McGraw-Hill Companies, 2010).

Mecacci, Giulio, Pim Haselager. “Identifying Criteria for the Evaluation of the Implications of Brain Reading for Mental Privacy.” *Sci Eng Ethics* 25 (Dec., 2019): 443-461. <https://link.springer.com/article/10.1007/s11948-017-0003-3>.

Messner, Eva-Maria, Rayna Sariyska, Benjamin Mayer, Christian Montag, Christopher Kannen, Andreas Schwerdtfeger, Harald Baumeister. “Insights – Future Implications of Passive Smartphone Sensing in the Therapeutic Context.” *Verhaltenstherapie* 32, no 1 (Aug., 2022): 86–95. <https://karger.com/ver/article-pdf/32/Suppl.%201/86/3920346/000501951.pdf>.

Metting, I. Esther, Eveline Hage. “Outside the Digital Bubble: Digital illiteracy is common and deserves more attention when implementing ehealth.” *Social Science Research Network (SSRN)* (Oct., 2022): 1-9. <https://doi.org/10.2139/ssrn.4246806>.

Miller Geoffrey. “The Smartphone Psychology Manifesto.” *Perspectives on Psychological Science* 7, no. 3 (May, 2012): 221-237. <https://doi.org/10.1177/1745691612441215>.

Mitov, Alexander. “Ethical use of artificial intelligence through the Utilitarianism perspective.” *University of Twente* (2021): 1-8. https://essay.utwente.nl/86726/1/Mitov_BA_BMS.pdf.

Mohr, C. David, Juned Siddique, Joyce Ho, Jenna Duffecy, Ling Jin, Konadu J. Fokuo. “Interest in Behavioral and Psychological Treatments Delivered Face-to-Face, by Telephone, and by Internet.” *Annals of Behavioral Medicine* 40, no. 1 (Aug., 2010): 89–98. <https://doi.org/10.1007/s12160-010-9203-7>.

Montag, Christian, Jon D. Elhai. “A new agenda for personality psychology in the digital age?” *Personality and Individual Differences* 147 (Sep., 2019): 128-134. <https://doi.org/10.1016/j.paid.2019.03.045>.

Montag, Christian, Cornelia Sindermann, Harald Baumeister. “Digital phenotyping in psychological and medical sciences: a reflection about necessary prerequisites to reduce harm and increase benefits.” *Current Opinion in Psychology* 36 (Apr. 2020): 19-24. <https://doi.org/10.1016/j.copsyc.2020.03.013>.

Morris, Robert, Kareem Kouddous, Rohan Kshirsagar, Stephen Schueller. “Towards an Artificially Empathic Conversational Agent for Mental Health Applications: System Design and User Perceptions.” *Journal of Medical Internet Research* 20, no. 6, e10148 (June, 2018): 1-10. <http://www.jmir.org/2018/6/e10148/>.

MOST. “Mental health support when and where you need it.” Accessed September 23, 2023. <https://most.org.au/>.

Müller, Vincent. “Ethics of Artificial Intelligence and Robotics.” *Stanford Encyclopedia of Philosophy*. Accessed August 21, 2023, <https://plato.stanford.edu/entries/ethics-ai/>.

MyTherapist. «Πως λειτουργεί η online ψυχοθεραπεία στο myTherapist». Προσπελάστηκε στις 28 Σεπτεμβρίου, 2023. <https://www.mytherapist.gr/web-app>.

National Institute of Mental Health. “About RDoC.” Accessed September 11, 2023. <https://www.nimh.nih.gov/research/research-funded-by-nimh/rdoc/about-rdoc>.

Newell, Allen, Herbert Simon. “The logic theory machine: A complex information processing system.” *IRE Transactions on Information Theory*, vol. 2, no. 3 (Sep. 1956): 61-79, doi: 10.1109/TIT.1956.1056797.

Niculescu, Andreea, Betsy van Dijk, Anton Nijholt, Haizhou Li, Swee Lan See. “Making Social Robots More Attractive: The Effects of Voice Pitch, Humor and Empathy.” *International Journal of Social Robotics* 5 (Jan., 2013): 171–191. <https://link.springer.com/article/10.1007/s12369-012-0171-x>.

Nikolopoulou, Kassiani. “What is Social Desirability Bias? | Definition & Examples.” March 24, 2023 at *Scribbr*. <https://www.scribbr.com/research-bias/social-desirability-bias/>.

NovoPsych. “Agnew Relationship Measure – 5 (ARM-5).” Accessed in October 1, 2023. <https://novopsych.com.au/assessments/feedback-informed/agnew-relationship-measure-5-arm-5/>.

OpenAI. “ChatGPT.” Accessed October 6, 2023. <https://openai.com/chatgpt>.

Παναγοπούλου-Κουτνατζή, Φερενίκη. *Τεχνητή Νοημοσύνη: Ο δρόμος προς έναν ψηφιακό συνταγματισμό. Μια ηθικο-συνταγματική θεώρηση*. Αθήνα: Εκδόσεις Παπαζήση, 2023.

Πελεγρίνης, Ν. Θεοδόσιος. *Ηθική Φιλοσοφία*. Αθήνα: ΠΕΔΙΟ, 2015.

Πλακιά, Θεοδώρα. «Η μείωση της βλάβης στη δημόσια υγεία. Βασικές Αρχές Βιοηθικής». Διπλωματική διατριβή, Πανεπιστήμιο Δυτικής Αττικής, 2022. <https://scholar.archive.org/work/5d3ckdxecobcwouimp5ztyfjyy>.

Πρωτόπαπας, Αθανάσιος. *Εισαγωγή στην θεωρία και Μεθοδολογία της Γνωσιακής Επιστήμης* [πανεπιστημιακές σημειώσεις]. ΕΚΠΑ. Τμήμα Ιστορίας και Φιλοσοφίας της Επιστήμης. 2021. Αθήνα. http://users.uoa.gr/~aprotopapas/CV/pdf/Protopapas_cogsci-notes_2011.pdf.

Pagollang, Motloba, Moipolai Pusetso, Thembelihle Mtolo, Sandile Mpungose. “To extract or not to extract – Felicitic calculus to the rescue.” *South African Dental Journal* 77, no 8 (September 2022): 504-509. <http://dx.doi.org/10.17159/2519-0105/2022/v77no8a9>.

Palanica, Adam, Peter Flaschner, Anirudh Thommandram, Michael Li, Yan Fossat. “Physicians' Perceptions of Chatbots in Health Care: Cross-Sectional Web-Based Survey.” *Journal of Medical Internet Research* 21, no. 4, e12887 (April, 2019): 1-10. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6473203/>.

Probono Australia. “Virtual counsellor steps in to help out on suicide hotline.” Accessed August 20, 2023. https://probonoaustralia-com-au.translate.google.com/news/2019/04/virtual-counsellor-steps-in-to-help-out-on-suicide-hotline/?x_tr_sl=en&x_tr_tl=el&x_tr_hl=el&x_tr_pto=wapp.

Psybot. “Psybot. Therapeutic techniques are in your pocket now!” Accessed September 28, 2023. <http://www.psybotapp.com/en/index.html>.

Psychologynow.gr. «Η Τεχνητή νοημοσύνη μπορεί να εντοπίζει αν τα αντικαταθλιπτικά φάρμακα θα “δουλέψουν” ή όχι». Προσπελάστηκε στις 22 Σεπτεμβρίου, 2023. <https://www.psychologynow.gr/arthra-psycholegias/tehnologia/ilektronikes-efarmoges/13308-texniti-noimosyni-kathoristiki-sti-therapeia-tis-katathlipsis.html>.

Psychology Today. “The Psychology of Chatbots: When interacting with bots our brain believes it's chatting with a human.” Accessed October 10, 2023. <https://www.psychologytoday.com/intl/blog/behind-online-behavior/201607/the-psychology-chatbots>.

Psykiatros.gr. «Online Ερωτηματολόγιο Κατάθλιψης PHQ-9». Προσπελάστηκε στις 23 Αυγούστου 2023. <https://www.psykiatros.gr/online->

<https://www.quantum.gr/qblog/personal-data/>.

Quantum Business & IT Solutions. «Τι μάθαμε από το σκάνδαλο της Cambridge Analytica για τα προσωπικά μας δεδομένα». Προσπελάστηκε στις 10 Νοεμβρίου 2023. <https://www.quantum.gr/qblog/personal-data/>.

Rogers, Carl R. “Counseling.” *Review of Educational Research* 15, no. 2 (Apr.1945): 155–63. <https://doi.org/10.2307/1168689>.

Rozgonjuk, Dimitri, Jason C. Levine, Brian J. Hall, Jon D. Elhai. “The association between problematic smartphone use, depression and anxiety symptom severity, and objectively measured smartphone use over one week.” *Computers in Human Behavior* 87 (May, 2018): 10-17. <https://doi.org/10.1016/j.chb.2018.05.019>.

Rubeis, Giovanni. “E-mental health applications for depression: an evidence-based ethical analysis.” *Eur Arch Psychiatry Clin Neurosci* 271 (Jan., 2021): 549–555. <https://doi.org/10.1007/s00406-019-01093-y>.

Rzepka, Christine, Benedikt Berger, Thomas Hess. “Voice Assistant vs. Chatbot – Examining the Fit Between Conversational Agents’ Interaction Modalities and Information Search Tasks.” *Information Systems Frontiers* 24 no. 2 (June, 2022): 1-18. <https://link.springer.com/article/10.1007/s10796-021-10226-5>.

Σαλαμούρας, Γεώργιος. «Η διάκριση των αρετών και η αρετή της φιλίας κατά τον Αριστοτέλη». *Conatus - Journal of Philosophy* 2, vo. 1 (Ιούνιος, 2018): 131. <https://doi.org/10.12681/conatus.16031>.

Συμβουλευτικό Κέντρο Φοιτητών ΑΣΠΑΙΤΕ. «Τι είναι η Συμβουλευτική». Πρόσβαση στις 07 Ιουλίου 2023. <https://skf.aspete.gr/ti-einai-i-symvouleftiki/>.

Schwitzgebel, Eric, Mara Garza. “Designing AI with Rights, Consciousness, Self-Respect, and Freedom.” In *Ethics of Artificial Intelligence*, ed. S. Matthew Liao, 549-479. New York: Oxford Academic, 2020. <https://doi.org/10.1093/oso/9780190905033.003.0017>

Sedlakova, Jana, Manuel Trachsel. “Conversational Artificial Intelligence in Psychotherapy: A New Therapeutic Tool or Agent?”. *The American Journal of Bioethics* 23, no. 5 (Apr. 2022): 4-13. <https://doi.org/10.1080/15265161.2022.2048739>.

Shanahan, Murray. “The Frame Problem.” *Stanford Encyclopedia of Philosophy*. Accessed September 6, 2023. <https://plato.stanford.edu/entries/frame-problem/>.

Siedlikowski, Sophia, Louis-Philippe Noël, Stephanie Anne Moynihan, Marc Robin. “Chloe for COVID-19: Evolution of an Intelligent Conversational Agent to Address Infodemic Management Needs During the COVID-19 Pandemic.” *Journal of Medical Internet Research* 23, no. 9, e27283 (Jan., 2021): 1-10. <https://www.jmir.org/2021/9/e27283>.

Singer, Natasha. “How companies scour our digital lives for clues to our health.” *New York Times*, July 10, 2018. <https://www.nytimes.com/2018/02/25/technology/smartphones-mental-health.html>.

Sitaru, Sebastian, Georg Bramm, Alexander Zink, Matthias Hiller. “Cybersecurity in digital healthcare—challenges and potential solutions.” *Dermatologie* 74 (Feb., 2023): 213-217. <https://doi.org/10.1007/s00105-023-05117-6>.

Sparrow, Robert. “Killer Robots.” *Journal of Applied Philosophy* 24, no. 1 (March, 2007): 62-77. <https://doi.org/10.1111/j.1468-5930.2007.00346.x>

Spreng, R. Nathan, Margaret C. McKinnon, Raymond A. Mar, Brian Levine. “The Toronto Empathy Questionnaire: scale development and initial validation of a factor-analytic solution to multiple empathy measures.” *Journal of Personality Assessment* 91, no. 1 (Jan., 2009): 62-71. <https://pubmed.ncbi.nlm.nih.gov/19085285/>.

Stanford Encyclopedia of Philosophy. “Michel Foucault.” Accessed November 3, 2023. <https://plato.stanford.edu/entries/foucault/>.

Stanford Encyclopedia of Philosophy. “The History of Utilitarianism.” Accessed December 4, 2023. <https://plato.stanford.edu/entries/utilitarianism-history/>.

Stanford Encyclopedia of Philosophy. “Theory and Bioethics.” Accessed November 2, 2023. <https://plato.stanford.edu/entries/theory-bioethics>.

Stange, P. Jonathan, John Zulueta, Scott A. Langenecker, Kelly A. Ryan, Andrea Piscitello, Jenna Duffecy, Melvin G. McInnis, Pete Nelson, Olusola Ajilore, Alex Leow. “Let your fingers do the talking: Passive typing instability predicts future mood outcomes,” *Bipolar Disorders* 20, no. 3 (March, 2018): 285-288. <https://doi.org/10.1111/bdi.12637>.

Stopczynski, Arkadiusz, Carsten Stahlhut, Jakob Eg Larsen, Michael Kai Petersen, Lars Kai Hansen. “The Smartphone Brain Scanner: A Portable Real-Time Neuroimaging System.” *Plos One* 9, no. 2, e86733 (Feb., 2014): 1-9. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0086733>.

Stoyanov, Stoyan, Leanne Hides, David Kavanagh, Oksana Zelenko, Dian Tjondronegoro, Madhavan Mani. “Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps.” *JMIR mHealth and uHealth* 3, no. 1 e27 (March, 2015): 1-9. https://www.researchgate.net/publication/273407216_Mobile_App_Rating_Scale_A_New_Tool_for_Assessing_the_Quality_of_Health_Mobile_Apps.

Suicide Call Back Service. “Feeling suicidal.” Προσπελάστηκε στις 28 Αυγούστου 2023. <https://www.suicidecallbackservice.org.au/feeling-suicidal/>.

Sweeney, Colm, Courtney Potts, Edel Ennis, Raymond Bond, Maurice D. Mulvenna, Siobhan O’neill, Martin Malcolm, Lauri Kuosmanen, Catrine Kostenius, Alex Vakaloudis, Gavin Mcconvey, Robin Turkington, David Hanna, Heidi Nieminen, Anna-Kaisa Vartiainen, Alison Robertson, Michael Mctear. “Can Chatbots Help Support a Person’s Mental Health? Perceptions and Views from Mental Healthcare

Professionals and Experts.” *ACM Transactions on Computing for Healthcare* 2, no. 3, art. 25 (July, 2021): 1-15. <https://dl.acm.org/doi/10.1145/3453175>.

Το χαμόγελο του παιδιού. «Ευρωπαϊκή Γραμμή για τα Εξαφανισμένα Παιδιά 116000». Προσπελάστηκε στις 16 Σεπτεμβρίου, 2023. <https://www.hamogelo.gr/gr/el/116000/>.

Thomason, Richmond. “Logic and Artificial Intelligence.” *Stanford Encyclopedia of Philosophy*. Accessed August 18, 2023. <https://plato.stanford.edu/entries/logic-ai/#jmc>.

Torrance, Steve. “Ethics and consciousness in artificial agents.” *AI & Society* 22, no. 4 (April, 2008): 495-521. https://www.researchgate.net/publication/220415053_Ethics_and_consciousness_in_artificial_agents.

Torous, J., J.P. Onnela, M. Keshavan. “New dimensions and new tools to realize the potential of RDoC: digital phenotyping via smartphones and connected devices.” *Translational Psychiatry* 7, no. 1053 (March, 2017): 1-3. <https://doi.org/10.1038/tp.2017.25>.

Φιλαλήθης, Α. «Ορόσημα στην εξέλιξη της πρωτοβάθμιας φροντίδας υγείας». *Archives of Hellenic Medicine* 38, no. 5 (Νοέμβριος, 2021): 675-682. <https://www.mednet.gr/archives/2021-5/pdf/675.pdf>.

Φουκώ, Μισέλ. *Επιτήρηση και Τιμωρία: Η Γέννηση της Φυλακής*. Μτφ. Κ. Χατζηδήμου & Ι. Ράλλη. Αθήνα: Εκδόσεις Ράππα, 1989. <https://pernoampariza.files.wordpress.com/2013/12/cf86cebfcf85cebaf8e-ceb5cf80ceb9cf84ceaecf81ceb7cf83ceb7-cebaceb1ceb9-cf84ceb9cebccf89cf81ceafceb1.pdf>.

Φωτίου, Ευάγγελος. «Η χρήση του Chat GPT – Προκλήσεις & ευκαιρίες: Εφαρμογές στην εκπαίδευση, στις επιχειρήσεις και στην κοινωνία». Διπλωματική διατριβή, Πανεπιστήμιο Μακεδονίας & Δημοκρίτειο Πανεπιστήμιο Θράκης, 2023.

Voelskow, Vanessa, Claudia Meßner, Tobias Kurth, Amelie Busam, Toivo Glatz, Natalie Ebert. “Prospective mixed-methods study evaluating the potential of a voicebot (CovBot) to relieve German health authorities during the COVID-19 infodemic.” *Digital Health* 9 (May, 2023): 1-14. <https://journals.sagepub.com/doi/full/10.1177/20552076231180677>.

Wiley. “The Digital Skills Gap Index (DSGI).” Accessed September 7, 2023. <https://dsgi.wiley.com/global-rankings/>.

Woebot Health. “Adult Mental Health.” Accessed July 9, 2023. <https://woebothealth.com/adult-mental-health/>.

Working Alliance Inventory. “WAI Home.” Accessed October 3, 2023. <https://wai.profhorvath.com/>.

Working Alliance Inventory. “Working Alliance Inventory – Short Revised (WAI-SR).” Accessed October 3, 2023. <https://wai.profhorvath.com/sites/default/files/upload/WAI-SR%20Client%20Version.pdf>.

World Health Organization. “ICD-11 International Classification of Diseases 11th Revision. The global standard for diagnostic health information.” Accessed September 7, 2023. <https://icd.who.int/en>.

World Health Organization. “Nairobi Mind Empowerment Peer Support Group - USP Kenya.” In the *Guidance on community mental health services: Promoting person-centred and rights-based approaches*. Geneva: WHO, 2021. <file:///C:/Users/estia/Downloads/9789240025707-eng.pdf>.

Wu, Philip Fei, Charlotte Summers, Arjun Panesar, Amit Kaura, Li Zhang. “AI Hesitancy and Acceptability: Perceptions of AI Chatbots for Chronic Health Management and Long-COVID Support.” *JMIR Preprints* (July, 2023). <https://s3.central-1.amazonaws.com/assets.jmir.org/assets/preprints/preprint-51086-submitted.pdf>.

Wysa. “Mental health that meets people where they are.” Accessed July 9, 2023. <https://www.wysa.io/>.

Χατζηβασιλείου, Χρυσή. «Προσωπικά δεδομένα, Τεχνητή Νοημοσύνη, υπολογιστική νέφους και Διαδίκτυο των Πραγμάτων στον τομέα της Υγείας». Διπλωματική διατριβή, Πανεπιστήμιο Μακεδονίας, Δημοκρίτειο Πανεπιστήμιο Θράκης, 2020. <https://dspace.lib.uom.gr/handle/2159/24152>.

Yeung, Karen, Andrew Howes, Ganna Pogrebna. “AI Governance by Human Rights–Centered Design, Deliberation, and Oversight. An End to Ethics Washing.” In *The Oxford Handbook of Ethics of AI*, edited by Markus D. Dubber, Frank Pasquale, Sunit Das, 77-106. USA: Oxford University Press, 2020.

Zhu, Yonghan, Rui Wang, Chengyan Pu. “I am chatbot, your virtual mental health adviser. What drives citizens’ satisfaction and continuance intention toward mental health chatbots during the COVID-19 pandemic? An empirical study in China.” *Digital Health* 8 (March, 2022): 1-15. <https://journals.sagepub.com/doi/full/10.1177/20552076221090031>.

Zioga, Ioanna, George Kosteletos, Evangelos D. Protopapadakis, Christos Papageorgiou, Konstantinos Kontoangelos, Charalabos Papageorgiou. “The Fox and the Lion: Investigating Associations between Empathy and Emotion Perspective-Taking in Aesop’s Fables.” *Psychology* 13, no. 4 (April, 2022): 482-513. <https://doi.org/10.4236/psych.2022.134033>