ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ

ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΔΗΜΟΣΙΑΣ ΔΙΟΙΚΗΣΗΣ & ΑΥΤΟΔΙΟΙΚΗΣΗΣ

Ινστιτούτο **Τ**εκμηρίωσης **Έ**ρευνας & **Κ**αινοτομιών

Ι.Τ.Ε.Κ

AI or Real? Η νέα πρόκληση της ψηφιακής αλήθειας στη δημόσια διοίκηση

Αθήνα Ιούνιος 2026

Παραδοτέο μελέτης

Θέμα: Generative AI, ψηφιακή αλήθεια και θεσμική αξιοπιστία

Έκδοση εργασίας προς κατάθεση

Το παρόν παραδοτέο εξετάζει τη μετάβαση από την παραπληροφόρηση στη δομική αβεβαιότητα της ψηφιακής αλήθειας και τις συνέπειες για τη δημόσια διοίκηση.

ΕΡΕΥΝΗΤΙΚΗ ΟΜΑΔΑ – ΣΥΝΤΑΚΤΕΣ ΕΚΘΕΣΗΣ:

Παπαδόπουλος Τίμος, Χριστούλιας Δημήτριος

ΑΞΙΟΛΟΓΗΤΡΙΑ: Αδαμίδου Μαρία

ΣΥΝΤΟΝΙΣΤΡΙΑ ΕΡΓΟΥ: Αγγελοπούλου Δήμητρα

ΠΕΡΙΕΧΟΜΕΝΑ

Εκτελεστική σύνοψη… σελ. 7

Μεθοδολογική προσέγγιση… σελ. 8

Περιορισμοί της μελέτης… σελ. 9

Κεφάλαιο 1

Η ψηφιακή αλήθεια ως θεσμικό διακύβευμα… σελ. 10

Η μετάβαση από την ψηφιακή εποχή στην εποχή της συνθετικής πληροφορίας… σελ. 10

Κεφάλαιο 2

2.1 Από την ανάλυση στην παραγωγή περιεχομένου… σελ. 13

2.2 Generative AI και η νέα ψηφιακή συνθήκη πληροφορίας… σελ. 14

2.3 Η μετατόπιση της ενημέρωσης από τους θεσμούς στις πλατφόρμες… σελ. 15

2.4 Η Generative AI ως τεχνολογία γενικού σκοπού… σελ. 17

Κεφάλαιο 3

Η πραγματικότητα υπό αμφισβήτηση: τι δείχνουν οι διεθνείς μελέτες και οι θεσμικές πρακτικές

3.1 Από την παραπληροφόρηση στη δομική αβεβαιότητα… σελ. 19

3.1Α Από τα Fake News στα Synthetic Media: η εξέλιξη της ψηφιακής παραπληροφόρησης… σελ. 19

3.2 Reuters Institute Digital News Report 2024: πλατφόρμες, νέα οικοσυστήματα και Gen AI… σελ. 21

3.3 OECD Survey on Drivers of Trust in Public Institutions: η εμπιστοσύνη ως κρίσιμος θεσμικός πόρος… σελ. 22

3.4 Pennycook et al. και Lewandowsky et al.: γιατί η ακρίβεια δεν αρκεί… σελ. 24

3.5 Deepfakes και συνθετικές φωνές: τι δείχνουν οι έρευνες ανίχνευσης… σελ. 25

3.5 Α Τεκμηρίωση για τα deepflakes σελ. 26

3.5Β Το φαινόμενο του «Liar's Dividend»… σελ. 26

3.6 Παραδείγματα που δυσκόλεψαν τη δημόσια αναγνώριση… σελ. 27

3.7 OECD Public Communication Scan of France: δημόσια επικοινωνία με αποδείξεις και δείκτες… σελ. 35

3.8 Σχήμα 1: Ενδεικτική σύγκριση αυθεντικού και AI-generated περιεχομένου… σελ.36

3.9 Σχήμα 2: Πώς ένα παραπλανητικό caption αλλοιώνει αυθεντικό περιεχόμενο… σελ. 37

Κεφάλαιο 4

Η ψυχολογία της εμπιστοσύνης στην ψηφιακή εποχή

4.1 Η εμπιστοσύνη ως γνωστική και συναισθηματική διαδικασία… σελ. 39

4.2 Ευεργετικές και γνωστικές συντομεύσεις στην αξιολόγηση της πληροφορίας… σελ. 39

4.3 Εμπιστοσύνη, αυθεντία και πηγή πληροφορίας… σελ. 40

4.4 Authority Bias: γιατί εξακολουθούμε να πιστεύουμε τους «ειδικούς»… σελ. 40

4.5 Information Overload και Trust Fatigue… σελ. 41

4.6 Γνωστικά φίλτρα, συναισθηματικά triggers και affective truth… σελ. 41

4.7 Εμπιστοσύνη και θεσμική κόπωση… σελ. 43

4.8 Εμπιστοσύνη, Gen AI και αντίληψη ελέγχου… σελ. 44

Κεφάλαιο 5

Δημόσια διοίκηση, θεσμικός λόγος και μη θεσμικοί φορείς επιρροής

5.1 Εισαγωγή: Influencers, parasocial trust και δημόσιος λόγος στην εποχή της Gen AI… σελ. 45

5.1.1 Από τη θεσμική αυθεντία στην ψηφιακή επιρροή… σελ. 45

5.1.2 Η οικονομία της προσοχής και οι μηχανισμοί επιρροής… σελ. 46

5.1.3 Parasocial relationships και εμπιστοσύνη… σελ. 46

5.1.4 Οι influencers ως διαμεσολαβητές δημόσιας πληροφορίας… σελ. 47

5.1.5 Influencers και Generative AI… σελ. 48

5.1.6 Τι σημαίνει αυτό για τη δημόσια διοίκηση… σελ. 48

5.2 Συμπέρασμα: θεσμικός λόγος σε περιβάλλον κατανεμημένης επιρροής… σελ. 49

Κεφάλαιο 6

Πρόταση Διαχείρισης της Ψηφιακής Αλήθειας στη Δημόσια Διοίκηση

6.1 Γιατί απαιτείται θεσμικό πλαίσιο και όχι αποσπασματικές λύσεις… σελ. 50

6.1Α Η ελληνική δημόσια διοίκηση απέναντι στη Generative AI… σελ. 50

6.1Β Διεθνείς πρακτικές δημόσιας διοίκησης απέναντι στην Gen AI… σελ. 52

6.2 Πρώτος πυλώνας: ψηφιακή παιδεία ως θεσμική πολιτική… σελ. 59

6.3 Δεύτερος πυλώνας: διαφάνεια στη χρήση της Gen AI… σελ. 60

6.4 Τρίτος πυλώνας: ανθρώπινη κρίση και θεσμική λογοδοσία… σελ. 60

6.5 Ο ρόλος του AI Act… σελ. 60

6.5Α Τι αλλάζει πρακτικά με τον AI Act για τους δημόσιους φορείς… σελ. 61

6.6 Ελληνικά παραδείγματα θεσμικής χρήσης AI… σελ. 62

6.7 Επιπτώσεις για τα στελέχη της δημόσιας διοίκησης… σελ. 63

Κεφάλαιο 7

Συμπεράσματα και θεσμικές επιπτώσεις

7.1 Η ψηφιακή αλήθεια ως μόνιμη συνθήκη… σελ. 65

7.2 Ο μετασχηματισμός του ρόλου της δημόσιας διοίκησης… σελ. 65

7.3 Εμπιστοσύνη: από δεδομένο σε δυναμική διαδικασία… σελ. 65

7.4 Θεσμικές επιλογές αντί τεχνικών αντανακλαστικών… σελ. 66

7.5 Η δημόσια διοίκηση σε οικοσύστημα επιρροής… σελ. 66

7.6 Η ψηφιακή αλήθεια ως δημόσιο αγαθό… σελ. 66

7.7 Τελική αποτίμηση… σελ. 66

Παράρτημα Α

Α.1 Deepfakes, synthetic media και ψηφιακή αλήθεια… σελ. 70

Α.2 Παραπληροφόρηση, εμπιστοσύνη και θεσμική αξιοπιστία… σελ. 70

Α.3 Πλατφόρμες, influencers και δημόσιος λόγος… σελ. 71

Α.4 Κανονιστικό πλαίσιο, δεοντολογία και πολιτικές για Gen AI… σελ. 71

Α.5 Παραδείγματα εφαρμογής AI και Gen AI στον δημόσιο τομέα – Ελλάδα… σελ. 72

Α.6 Δημόσια διοίκηση και Gen AI – διεθνείς πρακτικές… σελ. 73

AI or Real?

Η νέα πρόκληση της ψηφιακής αλήθειας στη δημόσια διοίκηση

Εκτελεστική Σύνοψη

Η παρούσα μελέτη εξετάζει τις επιπτώσεις της Generative AI (Gen AI) στη δημόσια επικοινωνία και, ευρύτερα, στη θεσμική αξιοπιστία της δημόσιας διοίκησης. Η ραγδαία ανάπτυξη τεχνολογιών παραγωγής συνθετικού περιεχομένου, όπως εικόνες, βίντεο, φωνή και κείμενο, έχει μεταβάλει ουσιωδώς το περιβάλλον μέσα στο οποίο οι πολίτες προσλαμβάνουν, ερμηνεύουν και αξιολογούν τη δημόσια πληροφορία.

Η μελέτη υποστηρίζει ότι η πρόκληση της ψηφιακής αλήθειας δεν περιορίζεται πλέον στη σκόπιμη παραπληροφόρηση. Συνδέεται με μια ευρύτερη δομική αβεβαιότητα ως προς την αυθεντικότητα του περιεχομένου: ποιος το παρήγαγε, με ποια μέσα, με ποια πρόθεση και με ποιο βαθμό ανθρώπινης ευθύνης. Η Generative AI δεν δημιουργεί απλώς περισσότερο περιεχόμενο. Δημιουργεί περιεχόμενο που μπορεί να είναι γλωσσικά, οπτικά και αφηγηματικά πειστικό, καθιστώντας δυσκολότερη τη διάκριση ανάμεσα στο αυθεντικό και το συνθετικό.

Η ανάλυση βασίζεται σε διεθνείς μελέτες και θεσμικά κείμενα, όπως το OECD Public Communication Scan of France, το OECD Survey on Drivers of Trust in Public Institutions, το Reuters Institute Digital News Report 2024, ο AI Act της Ευρωπαϊκής Ένωσης, έρευνα ερευνητών του MIT Media Lab, το Stanford AI Index Report, το Deeptrace και το European Commission AI Watch. Το OECD Public Communication Scan of France είναι ιδιαίτερα χρήσιμο για τη μελέτη, καθώς αντιμετωπίζει τη δημόσια επικοινωνία ως λειτουργία που πρέπει να αξιολογείται με βάση την απόδοση και τον αντίκτυπό της, και όχι ως απλό μηχανισμό διάχυσης μηνυμάτων.

Ιδιαίτερη έμφαση δίνεται στον ρόλο των μη θεσμικών φορέων επιρροής, δηλαδή των influencers, ως παραγόντων που διαμεσολαβούν τον δημόσιο λόγο χωρίς να διαθέτουν θεσμική λογοδοσία. Η μελέτη δεν προτείνει να λειτουργήσει η δημόσια διοίκηση ως influencer. Αντίθετα, υποστηρίζει

ότι η διοίκηση οφείλει να κατανοήσει το οικοσύστημα επιρροής μέσα στο οποίο πλέον λειτουργεί και να διατηρήσει τον διακριτό της ρόλο: σαφής, αξιόπιστη, αναγνωρίσιμη και υπεύθυνη θεσμική φωνή.

Η μελέτη καταλήγει σε ένα πλαίσιο διαχείρισης της ψηφιακής αλήθειας στη δημόσια διοίκηση, βασισμένο σε τρεις πυλώνες: ψηφιακή παιδεία, διαφάνεια στη χρήση της Gen AI και ανθρώπινη κρίση. Οι τρεις αυτοί πυλώνες δεν συνιστούν τεχνικό εγχειρίδιο, αλλά θεσμικό πλαίσιο αρχών που μπορεί να αξιοποιηθεί στην επιμόρφωση στελεχών, στον σχεδιασμό δημόσιας επικοινωνίας και στη διαμόρφωση εσωτερικών πρωτοκόλλων.

ΜΕΘΟΔΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ

Η παρούσα μελέτη βασίζεται σε ποιοτική βιβλιογραφική και θεσμική επισκόπηση σχετικά με τη Generative AI (Gen AI), τη δημόσια επικοινωνία, την ψηφιακή εμπιστοσύνη και τη διαχείριση της πληροφορίας στη δημόσια διοίκηση.

Για τις ανάγκες της ανάλυσης αξιοποιήθηκαν:

- διεθνείς θεσμικές εκθέσεις και κανονιστικά κείμενα,
- ακαδημαϊκή βιβλιογραφία,
- μελέτες οργανισμών διεθνούς κύρους,
- επιλεγμένα case studies,
- καθώς και παραδείγματα εφαρμογής Gen AI στον δημόσιο τομέα.

Ιδιαίτερη έμφαση δόθηκε:

- στις επιπτώσεις της Gen AI στη θεσμική αξιοπιστία,
- στη μεταβολή του τρόπου πρόσληψης της δημόσιας πληροφορίας,
- στον ρόλο των μη θεσμικών φορέων επιρροής,
- και στις προκλήσεις που δημιουργούνται για τη δημόσια διοίκηση ως φορέα έγκυρης και αξιόπιστης πληροφόρησης.

Η μελέτη δεν λειτουργεί ως τεχνικός οδηγός χρήσης εργαλείων τεχνητής νοημοσύνης ούτε ως εγχειρίδιο επικοινωνιακής στρατηγικής. Αντίθετα, επιδιώκει να συμβάλει στη θεσμική κατανόηση της νέας ψηφιακής συνθήκης και στις επιπτώσεις της για τη δημόσια διοίκηση και τη δημόσια επικοινωνία.

Περιορισμοί της μελέτης

Η παρούσα μελέτη βασίζεται σε ποιοτική βιβλιογραφική και θεσμική επισκόπηση και δεν περιλαμβάνει πρωτογενή εμπειρική έρευνα, όπως συνεντεύξεις, ερωτηματολόγια ή συστηματική αξιολόγηση συγκεκριμένων ελληνικών δημόσιων φορέων. Τα παραδείγματα και τα case studies που αξιοποιούνται έχουν επιλεγεί ενδεικτικά, με στόχο να αναδείξουν τάσεις, κινδύνους και θεσμικά ζητήματα που συνδέονται με τη Generative AI και την ψηφιακή αλήθεια.

Η ανάλυση δεν αποσκοπεί στην τεχνική αξιολόγηση εργαλείων AI ούτε στην αποτίμηση της αποτελεσματικότητας συγκεκριμένων συστημάτων ανίχνευσης deepfakes ή synthetic media. Εστιάζει κυρίως στις θεσμικές, επικοινωνιακές και οργανωσιακές συνέπειες της νέας ψηφιακής συνθήκης για τη δημόσια διοίκηση.

Επιπλέον, λόγω της ταχύτητας με την οποία εξελίσσονται οι τεχνολογίες Generative AI, τα κανονιστικά πλαίσια και οι πρακτικές των δημόσιων διοικήσεων, ορισμένα στοιχεία ενδέχεται να χρειάζονται επικαιροποίηση σε μεταγενέστερο χρόνο. Συνεπώς, τα συμπεράσματα της μελέτης πρέπει να διαβάζονται ως θεσμικό και αναλυτικό πλαίσιο κατανόησης, όχι ως οριστική ή εξαντλητική καταγραφή ενός ταχέως μεταβαλλόμενου πεδίου.

Κεφάλαιο 1

Η ψηφιακή αλήθεια ως θεσμικό διακύβευμα

Η έννοια της αλήθειας στη δημόσια σφαίρα δεν είναι στατική. Διαχρονικά, η δημόσια διοίκηση λειτουργεί ως βασικός φορέας παραγωγής, πιστοποίησης και διάχυσης έγκυρης πληροφορίας. Μέσα από αυτή τη λειτουργία στηρίζεται η νομιμοποίηση των δημόσιων πολιτικών, η προβλεψιμότητα της διοικητικής δράσης και η σχέση εμπιστοσύνης κράτους–πολίτη.

Στην ψηφιακή εποχή, ο ρόλος αυτός επαναπροσδιορίζεται. Η δημόσια πληροφορία δεν διακινείται πλέον μόνο μέσω επίσημων ανακοινώσεων, θεσμικών ιστοσελίδων ή παραδοσιακών μέσων ενημέρωσης. Συνυπάρχει με αναρτήσεις κοινωνικών δικτύων, σχολιασμούς χρηστών, περιεχόμενο δημιουργών, αυτοματοποιημένες απαντήσεις και, πλέον, με περιεχόμενο που παράγεται ή υποστηρίζεται από Generative AI.

Η μετάβαση από την ψηφιακή εποχή στην εποχή της συνθετικής πληροφορίας

Η δημόσια διοίκηση έχει ήδη βιώσει δύο μεγάλες μεταβολές στον τρόπο παραγωγής και διακίνησης της πληροφορίας. Η πρώτη συνδέθηκε με τη διάδοση του διαδικτύου, που κατέστησε την πληροφορία άμεσα προσβάσιμη σε μεγάλο μέρος του πληθυσμού. Η δεύτερη συνδέθηκε με την ανάπτυξη των κοινωνικών δικτύων, τα οποία επέτρεψαν σε κάθε πολίτη να λειτουργεί ταυτόχρονα ως καταναλωτής και παραγωγός περιεχομένου.

Η ανάπτυξη της Generative AI σηματοδοτεί μια τρίτη και ποιοτικά διαφορετική μετάβαση. Για πρώτη φορά, η παραγωγή περιεχομένου δεν αποτελεί αποκλειστικό χαρακτηριστικό της ανθρώπινης δραστηριότητας. Κείμενα, εικόνες, βίντεο και ηχητικά αρχεία μπορούν πλέον να δημιουργούνται αυτόματα, σε μεγάλη κλίμακα και με ελάχιστο κόστος. Η συνθήκη αυτή δημιουργεί ένα νέο περιβάλλον πρόσληψης: ο πολίτης

δεν αναζητά πάντα ενεργά την πληροφορία, συχνά την συναντά μέσα σε ροές περιεχομένου, όπου θεσμικές πηγές και μη θεσμικές αφηγήσεις εμφανίζονται στο ίδιο επίπεδο ορατότητας. Η διάκριση μεταξύ επίσημου, ανεπίσημου, έγκυρου, σχολιαστικού και παραπλανητικού περιεχομένου δεν είναι πάντοτε αυτονόητη.

Η Generative AI εντείνει αυτή τη μεταβολή, διότι καθιστά δυνατή την παραγωγή περιεχομένου που δεν αναπαριστά απλώς την πραγματικότητα, αλλά τη συνθέτει με τρόπο λειτουργικά πειστικό. Μια εικόνα μπορεί να φαίνεται φωτογραφική χωρίς να είναι. Μια φωνή μπορεί να ακούγεται πραγματική χωρίς να ανήκει στο πρόσωπο που υποτίθεται ότι μιλά. Ένα κείμενο μπορεί να διαθέτει λογική ροή, θεσμικό ύφος και φαινομενική τεκμηρίωση χωρίς να έχει παραχθεί από αρμόδιο φορέα ή υπεύθυνο συντάκτη.

Η εξέλιξη αυτή μεταβάλλει θεμελιωδώς το πληροφοριακό περιβάλλον μέσα στο οποίο λειτουργούν οι δημόσιοι θεσμοί. Αν στο παρελθόν η πρόκληση αφορούσε την πρόσβαση στην πληροφορία ή την ταχύτητα διάδοσής της, σήμερα αφορά ολοένα περισσότερο την αναγνώριση της προέλευσης και της αυθεντικότητάς της.

Η νέα αυτή συνθήκη έχει περιγραφεί από αρκετούς ερευνητές ως μετάβαση προς ένα περιβάλλον «συνθετικής πληροφορίας» (synthetic information environment), στο οποίο σημαντικό μέρος του περιεχομένου που κυκλοφορεί ψηφιακά δεν παράγεται άμεσα από ανθρώπους αλλά από αλγοριθμικά συστήματα.

Για τη δημόσια διοίκηση, η μεταβολή αυτή δημιουργεί νέες απαιτήσεις. Η αξιοπιστία των θεσμών δεν εξαρτάται πλέον μόνο από την ακρίβεια της πληροφορίας που παρέχουν αλλά και από την ικανότητά τους να λειτουργούν ως αναγνωρίσιμες και αξιόπιστες πηγές μέσα σε ένα περιβάλλον αυξανόμενης αβεβαιότητας.

Υπό αυτή την έννοια, η πρόκληση της ψηφιακής αλήθειας δεν αποτελεί απλώς τεχνολογικό ζήτημα. Συνδέεται με τη δυνατότητα των πολιτών να αναγνωρίζουν αξιόπιστες πληροφορίες, να εμπιστεύονται τους δημόσιους

θεσμούς και να λαμβάνουν τεκμηριωμένες αποφάσεις σε ένα ολοένα πιο σύνθετο πληροφοριακό περιβάλλον.

Για τη δημόσια διοίκηση, η εξέλιξη αυτή έχει ιδιαίτερη σημασία. Η αξιοπιστία της δεν εξαρτάται μόνο από την ακρίβεια των δεδομένων που παρέχει, αλλά και από την ικανότητά της να παραμένει αναγνωρίσιμη, σαφής και θεσμικά συνεπής σε ένα περιβάλλον αυξημένης αβεβαιότητας. Η ψηφιακή αλήθεια αναδεικνύεται έτσι σε θεσμικό ζήτημα και όχι απλώς σε τεχνολογικό πρόβλημα.

Στο πλαίσιο της παρούσας μελέτης, η ψηφιακή αλήθεια δεν αναφέρεται αποκλειστικά στην ακρίβεια μιας πληροφορίας. Περιλαμβάνει τη δυνατότητα επαλήθευσης της προέλευσής της, την αναγνώριση της αυθεντικότητάς της και την ύπαρξη θεσμών που μπορούν να λειτουργούν ως αξιόπιστα σημεία αναφοράς σε ένα σύνθετο ψηφιακό περιβάλλον.

Κεφάλαιο 2

2.1 Από την ανάλυση στην παραγωγή περιεχομένου

Η τεχνητή νοημοσύνη δεν αποτελεί νέο φαινόμενο για τη δημόσια διοίκηση. Εδώ και αρκετά χρόνια αξιοποιείται σε πεδία όπως η ανάλυση δεδομένων, η ταξινόμηση πληροφοριών, οι προβλέψεις και η αυτοματοποίηση διαδικασιών. Η Generative AI διαφοροποιείται ουσιωδώς από τις προηγούμενες γενιές συστημάτων, καθώς δεν περιορίζεται στην επεξεργασία υφιστάμενης πληροφορίας αλλά παράγει νέο περιεχόμενο.

Η μετάβαση αυτή μεταβάλλει τη σχέση ανθρώπου και τεχνολογίας. Για πρώτη φορά, η τεχνητή νοημοσύνη παρεμβαίνει άμεσα στον χώρο της έκφρασης, της αφήγησης και της επικοινωνίας. Παράγει κείμενα, εικόνες, βίντεο και ηχητικό υλικό που προσομοιάζουν ολοένα και περισσότερο την ανθρώπινη δημιουργία.

Για τη δημόσια διοίκηση, η εξέλιξη αυτή έχει ιδιαίτερη σημασία, καθώς επηρεάζει όχι μόνο τις εσωτερικές λειτουργίες των οργανισμών αλλά και το ίδιο το περιβάλλον μέσα στο οποίο διαμορφώνεται η δημόσια πληροφορία. Η παρούσα μελέτη δεν αντιμετωπίζει την τεχνητή νοημοσύνη ως απειλή για τη δημόσια διοίκηση. Αντίθετα, αναγνωρίζει τις σημαντικές δυνατότητες που προσφέρει για τη βελτίωση της εξυπηρέτησης πολιτών, την αποτελεσματικότητα των υπηρεσιών και τη διεύρυνση της πρόσβασης στην πληροφορία. Η πρόκληση δεν αφορά την υιοθέτηση ή μη της τεχνολογίας, αλλά τον τρόπο με τον οποίο αυτή ενσωματώνεται σε ένα πλαίσιο διαφάνειας, λογοδοσίας και εμπιστοσύνης.

2.2 Generative AI και η νέα ψηφιακή συνθήκη πληροφορίας

Η Generative AI αποτελεί κατηγορία συστημάτων τεχνητής νοημοσύνης που είναι σε θέση να παράγουν νέο περιεχόμενο: κείμενο, εικόνες, ήχο, βίντεο ή συνδυασμούς αυτών. Σε αντίθεση με προηγούμενες εφαρμογές AI, οι οποίες επικεντρώνονταν κυρίως στην ανάλυση, την ταξινόμηση ή την πρόβλεψη, η Gen AI παρεμβαίνει άμεσα στο πεδίο της παραγωγής λόγου, εικόνας και αφήγησης. Η ανάπτυξη της Generative AI έχει οδηγήσει στη δημιουργία μιας νέας κατηγορίας περιεχομένου που συχνά περιγράφεται ως synthetic media.

Ο όρος περιλαμβάνει:

- AI-generated εικόνες
- συνθετικές φωνές
- deepfake βίντεο
- avatars
- κείμενα παραγόμενα από μεγάλα γλωσσικά μοντέλα

Η σημασία των synthetic media δεν περιορίζεται στην τεχνολογία παραγωγής τους. Το κρίσιμο στοιχείο είναι ότι μπορούν να αναπαριστούν πρόσωπα, γεγονότα ή καταστάσεις που δεν υπήρξαν ποτέ, διατηρώντας ταυτόχρονα υψηλό βαθμό πειστικότητας.

Η διάδοση των synthetic media αποτελεί μία από τις σημαντικότερες μεταβολές στο οικοσύστημα της πληροφορίας από την εμφάνιση των κοινωνικών δικτύων.

Η διάδοση εργαλείων Gen AI έχει τρεις βασικές συνέπειες για το οικοσύστημα πληροφορίας.

Πρώτον, μειώνει δραστικά το κόστος και τον χρόνο παραγωγής περιεχομένου. Υλικό που παλαιότερα απαιτούσε εξειδικευμένες δεξιότητες, ανθρώπινους πόρους και οικονομικό κόστος μπορεί πλέον να παραχθεί ταχύτατα και σε μεγάλη κλίμακα. Η εξέλιξη αυτή μεταβάλλει τις ισορροπίες του ψηφιακού οικοσυστήματος: το πρόβλημα δεν είναι

πλέον η έλλειψη πληροφορίας αλλά η υπερπροσφορά της. Σε ένα περιβάλλον όπου το περιεχόμενο πολλαπλασιάζεται διαρκώς, η προσοχή και η αξιοπιστία καθίστανται ολοένα πιο πολύτιμοι πόροι.

Δεύτερον, παράγει περιεχόμενο υψηλής αισθητικής και γλωσσικής ποιότητας. Η ποιότητα αυτή ενισχύει την πειστικότητα, διότι το κοινό συχνά συνδέει την καθαρότητα, τη συνοχή και την αισθητική αρτιότητα με αξιοπιστία.

Τρίτον, δυσχεραίνει τη διάκριση μεταξύ ανθρώπινης και μηχανικής παραγωγής, ιδίως σε περιβάλλοντα ταχείας κατανάλωσης πληροφορίας. Το πρόβλημα δεν είναι μόνο ότι το ψευδές μπορεί να μοιάζει αληθινό. Είναι ότι το πραγματικό μπορεί να αρχίσει να αντιμετωπίζεται με καχυποψία, επειδή το συνθετικό έχει γίνει εξίσου πειστικό.

Η νέα αυτή συνθήκη δεν συνεπάγεται αυτομάτως παραπληροφόρηση. Η Gen AI μπορεί να υποστηρίξει θετικές χρήσεις: βελτίωση πρόσβασης σε πληροφορία, γλωσσική απλοποίηση, υποστήριξη δημοσίων υπηρεσιών, αυτοματοποίηση επαναλαμβανόμενων διαδικασιών και καλύτερη εξυπηρέτηση πολιτών. Ωστόσο, δημιουργεί συστημική αβεβαιότητα, εφόσον ακόμη και ακριβείς πληροφορίες μπορεί να τίθενται υπό αμφισβήτηση λόγω της αδυναμίας επιβεβαίωσης της προέλευσής τους.

Για τον πολίτη, το ερώτημα μετατοπίζεται από το «είναι σωστό;» στο «είναι πραγματικό;». Για τη δημόσια διοίκηση, η πρόκληση είναι ακόμη μεγαλύτερη: πώς διατηρείται η θεσμική αξιοπιστία σε ένα περιβάλλον όπου η μορφή της πληροφορίας δεν επαρκεί πλέον ως τεκμήριο αυθεντικότητας.

2.3 Η μετατόπιση της ενημέρωσης από τους θεσμούς στις πλατφόρμες

Οι εξελίξεις στην τεχνητή νοημοσύνη δεν επηρεάζουν μόνο τον τρόπο παραγωγής της πληροφορίας αλλά και το περιβάλλον μέσα στο οποίο οι πολίτες ενημερώνονται. Το Reuters Institute Digital News Report

καταγράφει τα τελευταία χρόνια μια σταδιακή αλλά σταθερή μετατόπιση της κατανάλωσης ενημέρωσης από τα παραδοσιακά μέσα ενημέρωσης προς τις ψηφιακές πλατφόρμες και τα κοινωνικά δίκτυα.

Σύμφωνα με την έκδοση του 2024, ολοένα και περισσότεροι πολίτες, ιδιαίτερα νεότερων ηλικιών, ενημερώνονται μέσω πλατφορμών όπως το Instagram, το TikTok και το YouTube, ενώ η άμεση επαφή με παραδοσιακούς δημοσιογραφικούς οργανισμούς εμφανίζει πτωτική τάση. Η εξέλιξη αυτή δεν αφορά μόνο τη μεταβολή των καναλιών ενημέρωσης αλλά και τη μεταβολή των ίδιων των κριτηρίων αξιολόγησης της πληροφορίας.

Σε αντίθεση με το παραδοσιακό μοντέλο ενημέρωσης, όπου η αξιοπιστία συνδεόταν σε μεγάλο βαθμό με τον θεσμό που παρήγαγε την πληροφορία, στο ψηφιακό περιβάλλον η πληροφορία κυκλοφορεί αποσπασμένη από την αρχική της πηγή. Οι πολίτες έρχονται συχνά σε επαφή με αποσπάσματα περιεχομένου, σύντομα βίντεο, αναρτήσεις ή αναπαραγωγές πληροφοριών χωρίς να είναι πάντοτε σαφές ποιος είναι ο αρχικός δημιουργός ή ποια είναι η θεσμική προέλευσή τους.

Παράλληλα, το Reuters Institute επισημαίνει την αυξανόμενη επιρροή των λεγόμενων news creators και news influencers. Πρόκειται για πρόσωπα που λειτουργούν ως ενδιάμεσοι μεταξύ γεγονότων και κοινού, ερμηνεύοντας ή αναπαράγοντας ειδήσεις μέσα από προσωπικά κανάλια επικοινωνίας. Η επιρροή τους ενισχύεται ιδιαίτερα στις νεότερες ηλικιακές ομάδες, όπου η προσωπική σχέση εμπιστοσύνης με τον δημιουργό περιεχομένου συχνά υπερτερεί της αναγνωρισιμότητας ενός θεσμικού φορέα ή ενός παραδοσιακού μέσου ενημέρωσης.

Η μεταβολή αυτή δημιουργεί μια νέα πρόκληση για τη δημόσια διοίκηση. Οι δημόσιοι οργανισμοί δεν ανταγωνίζονται πλέον μόνο την παραπληροφόρηση αλλά και ένα ευρύτερο οικοσύστημα πληροφοριών στο οποίο η αξιοπιστία κρίνεται με διαφορετικά κριτήρια από εκείνα που ίσχυαν στο παρελθόν. Η ταχύτητα, η απλότητα, η προσωποποίηση και η αμεσότητα αποκτούν ολοένα μεγαλύτερη σημασία στη διαμόρφωση της δημόσιας αντίληψης.

Στο περιβάλλον αυτό, η έννοια της ψηφιακής αλήθειας αποκτά πρόσθετη πολυπλοκότητα. Η πρόκληση δεν περιορίζεται στην αναγνώριση ψευδών πληροφοριών αλλά επεκτείνεται στη δυνατότητα των πολιτών να αναγνωρίζουν αξιόπιστες πηγές μέσα σε ένα συνεχώς μεταβαλλόμενο ψηφιακό περιβάλλον. Η ενίσχυση της θεσμικής επικοινωνίας και της ψηφιακής παιδείας αναδεικνύεται έτσι ως κρίσιμος παράγοντας διατήρησης της εμπιστοσύνης στους δημόσιους θεσμούς.

2.4 Η Generative AI ως τεχνολογία γενικού σκοπού

Πολλοί διεθνείς οργανισμοί και ερευνητικά κέντρα αντιμετωπίζουν πλέον τη Generative AI ως τεχνολογία γενικού σκοπού (General Purpose Technology), αντίστοιχη σε σημασία με το διαδίκτυο ή τον ηλεκτρισμό.

Η προσέγγιση αυτή βασίζεται στο γεγονός ότι η τεχνολογία δεν επηρεάζει έναν μόνο τομέα δραστηριότητας αλλά διαχέεται οριζόντια σε ολόκληρη την οικονομία και τη δημόσια διοίκηση.

Οι επιπτώσεις της αφορούν:

- τη δημόσια επικοινωνία,
- την εκπαίδευση,
- την υγεία,
- τη δικαιοσύνη,
- την παραγωγικότητα,
- την εξυπηρέτηση πολιτών.

Για τον λόγο αυτό, η συζήτηση για τη Generative AI δεν περιορίζεται στην τεχνολογία αλλά επεκτείνεται σε ζητήματα θεσμών, διακυβέρνησης και εμπιστοσύνης.

Με αυτόν τον τρόπο, το Κεφάλαιο 2 λειτουργεί ως εννοιολογική γέφυρα ανάμεσα στη θεσμική πρόκληση της ψηφιακής αλήθειας και στην εμπειρική τεκμηρίωση που ακολουθεί. Η νέα ψηφιακή συνθήκη εξηγεί

γιατί στο επόμενο κεφάλαιο εξετάζονται διεθνείς μελέτες, παραδείγματα και θεσμικές πρακτικές που δείχνουν πώς η πραγματικότητα τίθεται πλέον υπό αμφισβήτηση.

Κεφάλαιο 3

Η πραγματικότητα υπό αμφισβήτηση: τι δείχνουν οι διεθνείς μελέτες και οι θεσμικές πρακτικές

3.1 Από την παραπληροφόρηση στη δομική αβεβαιότητα

Η διεθνής συζήτηση γύρω από την παραπληροφόρηση έχει μετατοπιστεί σημαντικά τα τελευταία χρόνια. Ενώ αρχικά το επίκεντρο βρισκόταν στη σκόπιμη διάδοση ψευδών ή παραπλανητικών πληροφοριών, η ταχεία ανάπτυξη της Generative AI έχει εισαγάγει μια νέα συνθήκη: τη δομική αβεβαιότητα ως προς την αυθεντικότητα του περιεχομένου.

Στη νέα αυτή συνθήκη, το πρόβλημα δεν περιορίζεται στο αν μια πληροφορία είναι ακριβής ή ανακριβής. Αφορά το αν είναι ανθρώπινης ή μηχανικής προέλευσης, αν έχει παραχθεί, τροποποιηθεί ή ενισχυθεί από συστήματα Gen AI, και κατά πόσο ο πολίτης μπορεί να το γνωρίζει αυτό με βεβαιότητα.

Για τη δημόσια διοίκηση, η εξέλιξη αυτή μετατρέπει την ψηφιακή αλήθεια σε θεσμικό ζήτημα πρώτης γραμμής, καθώς επηρεάζει άμεσα την αξιοπιστία της επίσημης πληροφόρησης.

3.1Α Από τα Fake News στα Synthetic Media: η εξέλιξη της ψηφιακής παραπληροφόρησης

Η συζήτηση γύρω από την παραπληροφόρηση δεν ξεκίνησε με τη Generative AI. Ήδη από τη δεκαετία του 2010, η διάδοση ψευδών ειδήσεων, παραπλανητικού περιεχομένου και οργανωμένων εκστρατειών επιρροής είχε αναδειχθεί σε σημαντική πρόκληση για τη δημοκρατία, τη δημοσιογραφία και τη δημόσια διοίκηση.

Η έννοια των fake news κυριάρχησε στη δημόσια συζήτηση ιδίως μετά το δημοψήφισμα για το Brexit και τις αμερικανικές εκλογές του 2016. Ωστόσο, η διεθνής βιβλιογραφία γρήγορα επισήμανε ότι ο όρος είναι υπερβολικά γενικός και συχνά ανεπαρκής για να περιγράψει τις διαφορετικές μορφές παραπληροφόρησης.

Η UNESCO (Wardle & Derakhshan, 2017) πρότεινε μια περισσότερο αναλυτική τυπολογία:

- misinformation: ανακριβές περιεχόμενο χωρίς πρόθεση εξαπάτησης,
- disinformation: ανακριβές περιεχόμενο με πρόθεση παραπλάνησης,
- malinformation: αληθινό περιεχόμενο που χρησιμοποιείται εκτός πλαισίου ή με στόχο την πρόκληση βλάβης.

Η διάκριση αυτή είναι ιδιαίτερα χρήσιμη για τη δημόσια διοίκηση, διότι αναδεικνύει ότι το πρόβλημα δεν περιορίζεται στο ψευδές περιεχόμενο. Αφορά και τον τρόπο με τον οποίο αληθινές πληροφορίες μπορούν να αποκοπούν από το πλαίσιο τους και να χρησιμοποιηθούν για τη δημιουργία λανθασμένων εντυπώσεων.

Η εμφάνιση της Generative AI μεταβάλλει ποιοτικά το τοπίο αυτό. Ενώ η παραδοσιακή παραπληροφόρηση βασιζόταν κυρίως στη διαστρέβλωση υπαρκτών πληροφοριών, η Gen AI καθιστά δυνατή την παραγωγή εξ ολοκλήρου νέου περιεχομένου. Εικόνες, βίντεο, φωνές και κείμενα μπορούν πλέον να δημιουργούνται χωρίς να αντιστοιχούν σε πραγματικά γεγονότα.

Το κρίσιμο στοιχείο είναι ότι η Gen AI μειώνει δραστικά το κόστος παραγωγής πειστικού περιεχομένου. Περιεχόμενο που παλαιότερα απαιτούσε εξειδικευμένο εξοπλισμό και σημαντικούς πόρους μπορεί πλέον να παραχθεί από έναν μόνο χρήστη μέσα σε λίγα λεπτά.

Η εξέλιξη αυτή οδηγεί σε μετάβαση από την εποχή των fake news στην εποχή των synthetic media. Ο όρος synthetic media χρησιμοποιείται για να περιγράψει περιεχόμενο που παράγεται ή τροποποιείται μέσω συστημάτων τεχνητής νοημοσύνης και περιλαμβάνει εικόνες, βίντεο, φωνές, avatars και κείμενα.

Σύμφωνα με το World Economic Forum (2024), η παραπληροφόρηση και η παραπληροφόρηση μέσω AI συγκαταλέγονται στους

σημαντικότερους παγκόσμιους κινδύνους για τα επόμενα χρόνια, ακριβώς επειδή υπονομεύουν την κοινή βάση πραγματικότητας πάνω στην οποία στηρίζονται οι δημοκρατικοί θεσμοί.

Για τη δημόσια διοίκηση, η μετάβαση αυτή σημαίνει ότι η αξιοπιστία δεν μπορεί πλέον να θεωρείται αυτονόητη συνέπεια της ακρίβειας. Απαιτείται ενεργή διαχείριση της αυθεντικότητας, της προέλευσης και της επαληθευσιμότητας της πληροφορίας.

3.2 Reuters Institute Digital News Report 2024: πλατφόρμες, νέα οικοσυστήματα και Gen AI

Το Reuters Institute Digital News Report 2024 αποτελεί μία από τις πιο χρήσιμες πηγές για την κατανόηση του σύγχρονου οικοσυστήματος ειδήσεων. Η έκθεση βασίζεται σε έρευνα της YouGov σε περισσότερα από 95.000 άτομα σε 47 χώρες, καλύπτοντας πληθυσμούς που αντιστοιχούν περίπου στο μισό του παγκόσμιου πληθυσμού.

Τα ευρήματα της έκθεσης έχουν άμεση σημασία για τη δημόσια διοίκηση. Το Reuters Institute καταγράφει την αυξανόμενη σημασία των πλατφορμών, ιδίως των πιο οπτικών και video-led περιβαλλόντων όπως TikTok, Instagram και YouTube, στην κατανάλωση και παραγωγή ειδήσεων. Αυτό σημαίνει ότι η δημόσια πληροφορία δεν προσλαμβάνεται πλέον απαραίτητα μέσα από οργανωμένα ενημερωτικά περιβάλλοντα, αλλά μέσα από πλατφόρμες όπου το θεσμικό περιεχόμενο συνυπάρχει με ψυχαγωγία, προσωπικές αφηγήσεις, influencers και περιεχόμενο αμφίβολης προέλευσης.

Η σημασία της έκθεσης δεν βρίσκεται μόνο στο ότι καταγράφει αλλαγές στη δημοσιογραφία. Βρίσκεται στο ότι περιγράφει ένα νέο περιβάλλον πρόσληψης της δημόσιας πληροφορίας: περισσότερο οπτικό, περισσότερο προσωποποιημένο, περισσότερο αποσπασματικό και λιγότερο συνδεδεμένο με παραδοσιακές πηγές αυθεντίας. Σε αυτό το πλαίσιο, η Gen AI λειτουργεί ως επιταχυντής, καθώς αυξάνει την

ποσότητα και την πειστικότητα του περιεχομένου που κυκλοφορεί σε αυτές τις πλατφόρμες.

3.3 OECD Survey on Drivers of Trust in Public Institutions: η εμπιστοσύνη ως κρίσιμος θεσμικός πόρος

Το OECD Survey on Drivers of Trust in Public Institutions – 2024 Results βασίζεται στη δεύτερη διακρατική έρευνα του OECD για τους παράγοντες που επηρεάζουν την εμπιστοσύνη στους δημόσιους θεσμούς. Η έρευνα πραγματοποιήθηκε στο τέλος του 2023, με σχεδόν 60.000 απαντήσεις, αντιπροσωπευτικές του ενήλικου πληθυσμού σε 30 χώρες του OECD.

Το κρίσιμο εύρημα για την παρούσα μελέτη είναι ότι η εμπιστοσύνη δεν είναι αφηρημένη στάση. Συνδέεται με τις εμπειρίες και τις προσδοκίες των πολιτών από τους δημόσιους θεσμούς: από την καθημερινή επαφή με δημόσιες υπηρεσίες έως την αντίληψη για τη λήψη αποφάσεων σε σύνθετα ζητήματα δημόσιας πολιτικής.

Η σχέση αυτή είναι σημαντική στην εποχή της Gen AI. Όταν ο πολίτης δυσκολεύεται να αξιολογήσει την προέλευση και την αυθεντικότητα της πληροφορίας, η εμπιστοσύνη στον θεσμό λειτουργεί ως ερμηνευτικό πλαίσιο. Δεν εξαλείφει την αβεβαιότητα, αλλά επηρεάζει τον τρόπο με τον οποίο ο πολίτης την επεξεργάζεται. Η εμπιστοσύνη, επομένως, δεν μπορεί να θεωρείται δεδομένο υπόβαθρο της δημόσιας επικοινωνίας. Αποτελεί δυναμικό θεσμικό κεφάλαιο που πρέπει να συντηρείται μέσω συνέπειας, διαφάνειας και λογοδοσίας.

Τι δημιουργεί εμπιστοσύνη στους δημόσιους θεσμούς: τα ευρήματα του OECD

Η προσέγγιση του OECD βασίζεται σε πέντε βασικές διαστάσεις: αξιοπιστία (reliability), ανταπόκριση (responsiveness), ανοιχτότητα (openness), ακεραιότητα (integrity) και δικαιοσύνη (fairness). Οι

διαστάσεις αυτές δεν αφορούν μόνο την ποιότητα των δημόσιων υπηρεσιών αλλά και τον τρόπο με τον οποίο οι πολίτες αντιλαμβάνονται τη λειτουργία των θεσμών.

Η αξιοπιστία συνδέεται με την ικανότητα του κράτους να παρέχει σταθερές και συνεπείς υπηρεσίες και πληροφορίες. Στο περιβάλλον της Generative AI, η διάσταση αυτή αποκτά νέα σημασία. Όταν οι πολίτες δυσκολεύονται να διακρίνουν την προέλευση μιας πληροφορίας, η εμπιστοσύνη μετατοπίζεται από το ίδιο το περιεχόμενο προς τον θεσμό που το παράγει ή το επικυρώνει. Η αξιοπιστία λειτουργεί έτσι ως μηχανισμός μείωσης της αβεβαιότητας.

Η ανταπόκριση αφορά την ικανότητα των δημόσιων οργανισμών να κατανοούν και να απαντούν στις ανάγκες των πολιτών. Η τεχνητή νοημοσύνη μπορεί να βελτιώσει σημαντικά την ταχύτητα και την προσβασιμότητα των δημόσιων υπηρεσιών. Ωστόσο, η χρήση της δεν ενισχύει αυτομάτως την εμπιστοσύνη. Αντίθετα, όταν οι πολίτες δεν κατανοούν πώς λαμβάνονται οι αποφάσεις ή πώς παράγονται οι απαντήσεις, η αίσθηση ανταπόκρισης μπορεί να υποχωρήσει παρά την τεχνολογική πρόοδο.

Η ανοιχτότητα συνδέεται με τη διαφάνεια και τη δυνατότητα των πολιτών να κατανοούν τις διαδικασίες λήψης αποφάσεων. Η διάσταση αυτή παρουσιάζει ιδιαίτερο ενδιαφέρον στην εποχή της Gen AI. Όσο περισσότερο οι δημόσιοι οργανισμοί αξιοποιούν συστήματα τεχνητής νοημοσύνης, τόσο μεγαλύτερη γίνεται η ανάγκη να εξηγούν πότε χρησιμοποιούνται τα εργαλεία αυτά, ποιος τα εποπτεύει και ποια είναι τα όριά τους.

Η ακεραιότητα αφορά την πεποίθηση ότι οι θεσμοί ενεργούν με συνέπεια και σύμφωνα με το δημόσιο συμφέρον. Σε περιβάλλον όπου το συνθετικό περιεχόμενο πολλαπλασιάζεται, η ακεραιότητα αποκτά πρόσθετη διάσταση: οι πολίτες χρειάζεται να γνωρίζουν ότι πίσω από κάθε δημόσια πληροφορία εξακολουθεί να υπάρχει σαφής ανθρώπινη ευθύνη και θεσμική λογοδοσία.

Τέλος, η δικαιοσύνη συνδέεται με την αντίληψη ότι οι δημόσιες πολιτικές και υπηρεσίες εφαρμόζονται ισότιμα. Η διεθνής συζήτηση για την τεχνητή νοημοσύνη έχει αναδείξει τον κίνδυνο αλγοριθμικών προκαταλήψεων και άνισης μεταχείρισης. Για τον λόγο αυτό, η διατήρηση της ανθρώπινης εποπτείας και η δυνατότητα ελέγχου των συστημάτων AI αποτελούν κρίσιμες προϋποθέσεις για τη διατήρηση της θεσμικής εμπιστοσύνης.

Υπό το πρίσμα αυτό, η εμπιστοσύνη δεν μπορεί να θεωρείται δεδομένο αποτέλεσμα της τεχνολογικής καινοτομίας. Αντίθετα, προκύπτει από τη συνδυασμένη λειτουργία αξιοπιστίας, διαφάνειας, λογοδοσίας και δίκαιης διακυβέρνησης. Η ψηφιακή αλήθεια αναδεικνύεται έτσι όχι μόνο ως επικοινωνιακό ή τεχνολογικό ζήτημα, αλλά ως κρίσιμη διάσταση της σχέσης μεταξύ κράτους και πολίτη.

3.4 Pennycook et al. και Lewandowsky et al.: γιατί η ακρίβεια δεν αρκεί

Η έρευνα των Pennycook et al. (2021) στο Nature δείχνει ότι οι άνθρωποι συχνά μοιράζονται παραπλανητικό περιεχόμενο όχι επειδή αδιαφορούν για την ακρίβεια, αλλά επειδή η προσοχή τους εστιάζεται σε άλλους παράγοντες. Στα πειράματα και στη δοκιμή πεδίου που περιγράφει η μελέτη, η απλή μετατόπιση της προσοχής προς την ακρίβεια βελτίωσε την ποιότητα του περιεχομένου που οι χρήστες ήταν διατεθειμένοι να μοιραστούν.

Το εύρημα αυτό είναι θεμελιώδες για τη δημόσια επικοινωνία. Δείχνει ότι η παραπληροφόρηση δεν οφείλεται πάντοτε σε πλήρη απόρριψη της αλήθειας. Συχνά οφείλεται σε περιβάλλοντα όπου η προσοχή αποσπάται από την ακρίβεια και στρέφεται σε άλλα στοιχεία: συναίσθημα, ταχύτητα, κοινωνική επιβεβαίωση, οικειότητα ή οπτική πειστικότητα.

Η μελέτη των Lewandowsky et al. (2012) εξετάζει το φαινόμενο της παραμονής της παραπληροφόρησης ακόμη και μετά τη διόρθωσή της. Η εργασία δείχνει ότι οι διορθώσεις δεν ακυρώνουν πάντοτε την επίδραση

της αρχικής λανθασμένης πληροφορίας, καθώς οι άνθρωποι διατηρούν συχνά ένα νοητικό μοντέλο του γεγονότος ακόμη και όταν μέρος του έχει διαψευστεί.

Για τη δημόσια διοίκηση, τα δύο αυτά ευρήματα είναι κρίσιμα. Η απλή διάθεση ορθής πληροφορίας δεν αρκεί. Η δημόσια επικοινωνία πρέπει να σχεδιάζεται έτσι ώστε να ενισχύει την προσοχή στην ακρίβεια, να προλαμβάνει παρερμηνείες και να διαμορφώνει σαφές πλαίσιο πριν εδραιωθούν λανθασμένες αφηγήσεις.

3.5 Deepfakes και συνθετικές φωνές: τι δείχνουν οι έρευνες ανίχνευσης

Η δυσκολία αναγνώρισης deepfakes δεν είναι θεωρητική. Μελέτες για την ανθρώπινη ικανότητα ανίχνευσης deepfake περιεχομένου δείχνουν ότι οι χρήστες δεν μπορούν πάντοτε να βασίζονται σε άμεση αισθητηριακή κρίση.

Η μελέτη “Warning: Humans cannot reliably detect speech deepfakes”, δημοσιευμένη στο PLOS ONE, εξέτασε την ικανότητα 529 συμμετεχόντων να διακρίνουν αυθεντική από συνθετική φωνή σε αγγλικά και μανδαρινικά. Τα αποτελέσματα έδειξαν ότι οι συμμετέχοντες εντόπισαν deepfake ομιλία με ακρίβεια περίπου 73%, γεγονός που σημαίνει ότι πάνω από το ένα τέταρτο των δειγμάτων δεν ανιχνεύθηκε σωστά.

Αντίστοιχα, η εργασία “Deepfake detection with and without content warnings” της Royal Society Open Science εξετάζει την ικανότητα ατόμων να εντοπίζουν deepfakes ακόμη και όταν υπάρχουν προειδοποιήσεις. Η μελέτη είναι σημαντική διότι θέτει το ερώτημα κατά πόσο η σήμανση ή η προειδοποίηση επαρκεί από μόνη της, αν οι χρήστες δεν διαθέτουν τα κατάλληλα εργαλεία, χρόνο ή δεξιότητες για αξιολόγηση.

Το συμπέρασμα για τη δημόσια διοίκηση είναι σαφές: η διάκριση αυθεντικού και συνθετικού περιεχομένου δεν μπορεί να αφεθεί μόνο στην ατομική κρίση του πολίτη. Απαιτεί συνδυασμό σήμανσης, θεσμικής επαλήθευσης, τεχνικών εργαλείων και ανθρώπινης κρίσης.

3.5Α Τεκμηρίωση για τα deepfakes

Η διεθνής τεκμηρίωση για τα deepfakes δείχνει ότι η πρόκληση δεν αφορά μόνο την τεχνική δυνατότητα δημιουργίας ψευδούς περιεχομένου, αλλά και την ταχύτητα με την οποία αυτό μπορεί να παραχθεί, να διακινηθεί και να γίνει πιστευτό. Εκθέσεις όπως η Deeptrace, το Stanford AI Index Report, το European Commission AI Watch και έρευνα ερευνητών του MIT Media Lab συγκλίνουν στο ότι τα εργαλεία ανίχνευσης είναι χρήσιμα, αλλά δεν επαρκούν από μόνα τους.

Η συμπληρωματική αυτή τεκμηρίωση ενισχύει το επιχείρημα ότι η αντιμετώπιση των deepfakes, απαιτεί συνδυασμό τεχνολογικής ανίχνευσης, ψηφιακής παιδείας, θεσμικής επαλήθευσης και σαφούς δημόσιας επικοινωνίας, Για τη δημόσια διοίκηση, το κρίσιμο ζήτημα είναι η δημιουργία έγκαιρης επιβεβαίωσης, επίσημης διάψευσης και αξιόπιστης ενημέρωσης, όταν κυκλοφορεί περιεχόμενο αμφίβολης αυθεντικότητας.

3.5Β Το φαινόμενο του «Liar's Dividend»

Οι Chesney και Citron (2019) εισήγαγαν τον όρο "Liar's Dividend" για να περιγράψουν ένα παράδοξο αποτέλεσμα της διάδοσης των deepfakes.

Καθώς αυξάνεται η δημόσια επίγνωση ότι εικόνες, βίντεο και ηχητικά αρχεία μπορούν να παραποιηθούν, αυξάνεται παράλληλα και η δυνατότητα αμφισβήτησης αυθεντικού περιεχομένου.

Με άλλα λόγια, η ύπαρξη των deepfakes δεν απειλεί μόνο επειδή επιτρέπει τη δημιουργία ψευδών αποδεικτικών στοιχείων. Απειλεί και επειδή επιτρέπει την απόρριψη πραγματικών αποδεικτικών στοιχείων ως δήθεν πλαστών.

Το φαινόμενο αυτό έχει ιδιαίτερη σημασία για τη δημόσια διοίκηση. Σε συνθήκες κρίσης, πολιτικής αντιπαράθεσης ή κοινωνικής έντασης, η αμφισβήτηση της αυθεντικότητας μπορεί να μετατραπεί σε εργαλείο υπονόμευσης της θεσμικής επικοινωνίας.

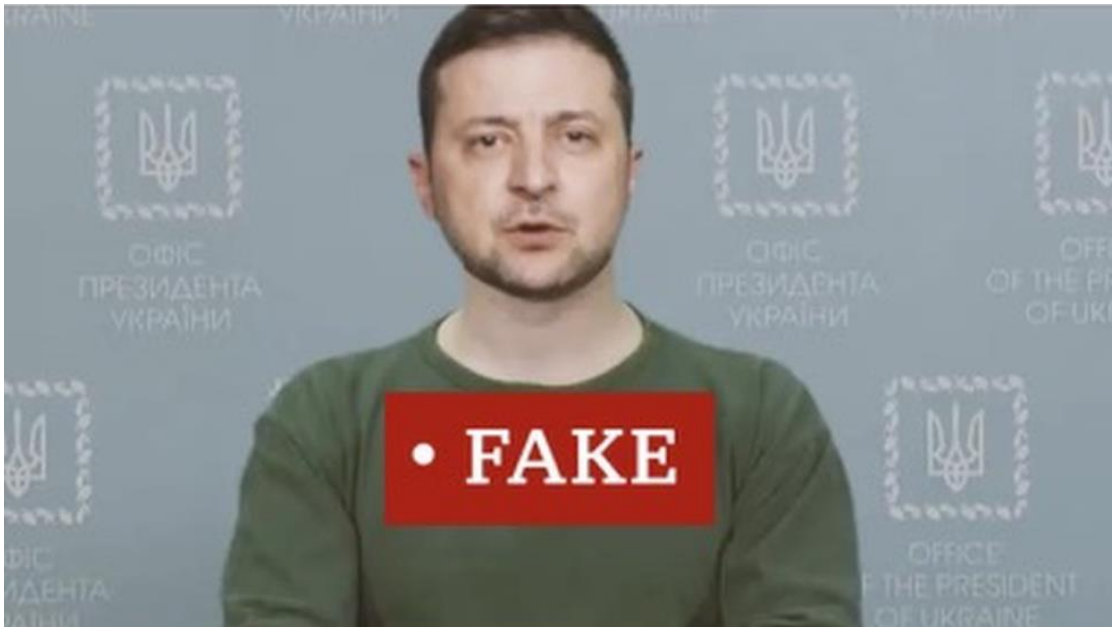
Η έννοια του Liar's Dividend αναδεικνύει ότι η πρόκληση της ψηφιακής αλήθειας δεν αφορά μόνο την ανίχνευση του ψευδούς, αλλά και τη διατήρηση της αξιοπιστίας του αληθινού.

3.6 Παραδείγματα που δυσκόλεψαν τη δημόσια αναγνώριση

Η διάδοση της Generative AI έχει ήδη οδηγήσει σε περιστατικά που ανέδειξαν τις επιπτώσεις των synthetic media στη δημόσια ζωή, την πολιτική και τη θεσμική επικοινωνία.

α) Το deepfake Zelenskyy

Το 2022 κυκλοφόρησε deepfake βίντεο που εμφάνιζε τον Πρόεδρο της Ουκρανίας Volodymyr Zelenskyy να καλεί τους Ουκρανούς να παραδοθούν. Το περιεχόμενο διακινήθηκε online και συνδέθηκε με κυβερνοπαραβίαση ουκρανικού ειδησεογραφικού φορέα, πριν διαψευστεί και αφαιρεθεί. Η University of Virginia παρουσίασε το περιστατικό ως ενδεικτικό της απειλής που δημιουργούν τα deepfakes για τη δημοκρατία και την ασφάλεια σε περιβάλλοντα κρίσης.



Η σημασία του παραδείγματος δεν βρίσκεται μόνο στο ότι το περιεχόμενο ήταν ψευδές. Βρίσκεται στο ότι εμφανίστηκε σε συνθήκες πολέμου, όπου η ταχύτητα, η σύγχυση και η συναισθηματική ένταση ενισχύουν τον κίνδυνο παραπλάνησης.

β) Το AI-generated image του Πάπα με puffer jacket

Το 2023 κυκλοφόρησε ευρέως εικόνα που εμφάνιζε τον Πάπα Φραγκίσκο να φορά λευκό puffer jacket. Η εικόνα έγινε viral και πολλοί χρήστες την εξέλαβαν ως πραγματική. Το AFP Fact Check επιβεβαίωσε ότι η εικόνα δεν ήταν αυθεντική και εντόπισε ενδείξεις AI manipulation, ενώ ειδικοί του Βατικανού ανέφεραν ότι το συγκεκριμένο ένδυμα δεν αντιστοιχούσε σε παπική ενδυμασία.



Το παράδειγμα είναι χρήσιμο επειδή δεν αφορά πολιτική κρίση ή κακόβουλη κρατική επιχείρηση. Δείχνει ότι ακόμη και "ελαφρύ" ή πολιτισμικό περιεχόμενο μπορεί να λειτουργήσει ως μαζικό τεστ πειστικότητας της Gen AI και να εξοικειώσει το κοινό με συνθετικές εικόνες που μοιάζουν αυθεντικές.

γ) Το AI-generated robocall με τη φωνή Biden

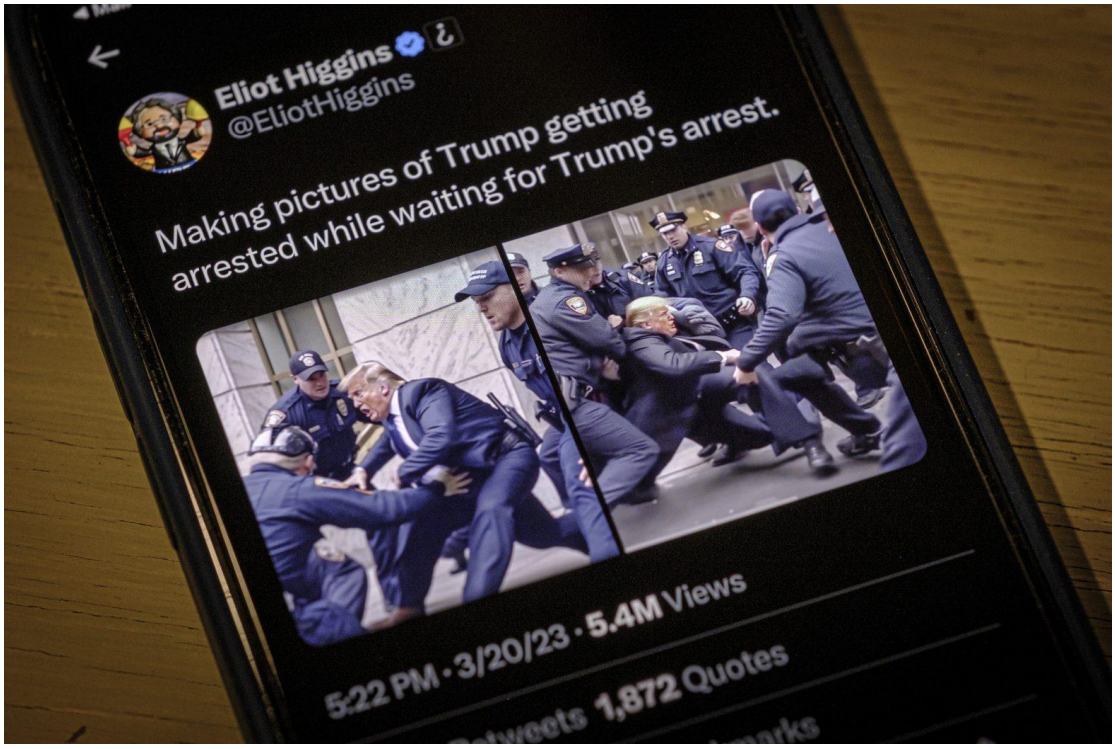
Το 2024, στις ΗΠΑ, εντοπίστηκε robocall με AI-generated voice clone του Προέδρου Joe Biden, το οποίο καλούσε πολίτες στο New Hampshire να μην ψηφίσουν στις προκριματικές εκλογές. Το New Hampshire Department of Justice ανακοίνωσε την έρευνα για το περιστατικό, ενώ η FCC πρότεινε πρόστιμο 6 εκατομμυρίων δολαρίων για τις παράνομες κλήσεις.



Το περιστατικό δείχνει ότι η Gen AI δεν αφορά μόνο εικόνες και κείμενα. Η συνθετική φωνή μπορεί να επηρεάσει άμεσα πολιτικές διαδικασίες και να υπονομεύσει την εμπιστοσύνη στην αυθεντικότητα της επικοινωνίας.

δ) Οι εικόνες σύλληψης του Donald Trump

Το 2023 κυκλοφόρησαν ευρέως εικόνες που παρουσίαζαν τον Donald Trump να συλλαμβάνεται από αστυνομικούς. Οι εικόνες είχαν δημιουργηθεί με Midjourney και δεν απεικόνιζαν πραγματικά γεγονότα.



Ωστόσο, διαδόθηκαν εκτενώς στα κοινωνικά δίκτυα και χρησιμοποιήθηκαν ως παράδειγμα της δυσκολίας διάκρισης μεταξύ πραγματικού και συνθετικού περιεχομένου.

ε) Οι εκλογές στη Σλοβακία

Λίγο πριν τις βουλευτικές εκλογές του 2023, κυκλοφόρησε ηχητικό deepfake που εμφάνιζε πολιτικό πρόσωπο να συζητά για νοθεία της εκλογικής διαδικασίας. Παρότι το περιεχόμενο διαψεύστηκε, το περιστατικό ανέδειξε τη δυσκολία άμεσης αντίδρασης όταν συνθετικό περιεχόμενο εμφανίζεται σε κρίσιμες χρονικές στιγμές.

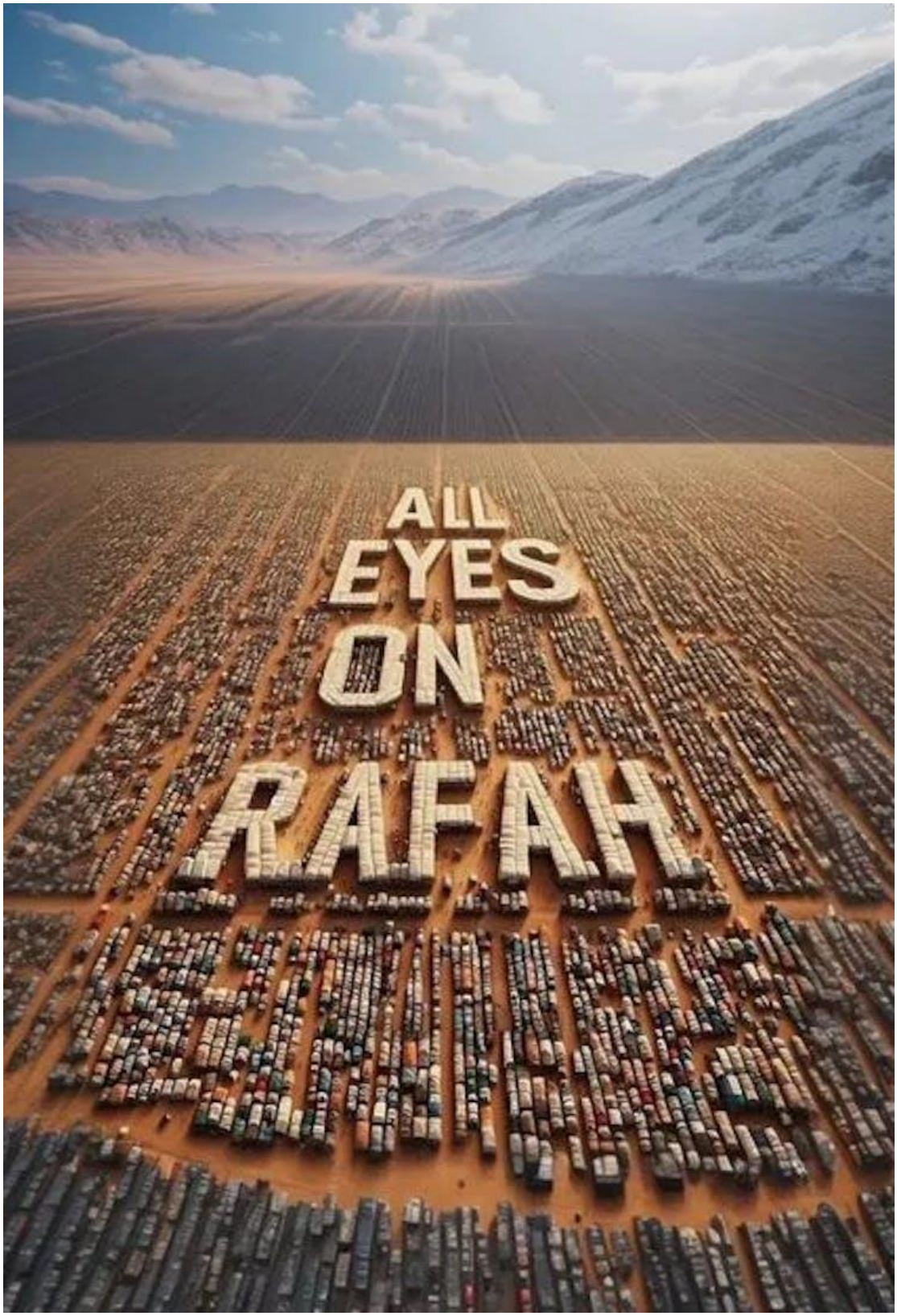
στ) AI-generated εικόνες από τη σύγκρουση στη Γάζα

Το 2023 και το 2024 κυκλοφόρησαν χιλιάδες AI-generated εικόνες που υποτίθεται ότι απεικόνιζαν γεγονότα από τη σύγκρουση στη Μέση

Ανατολή.



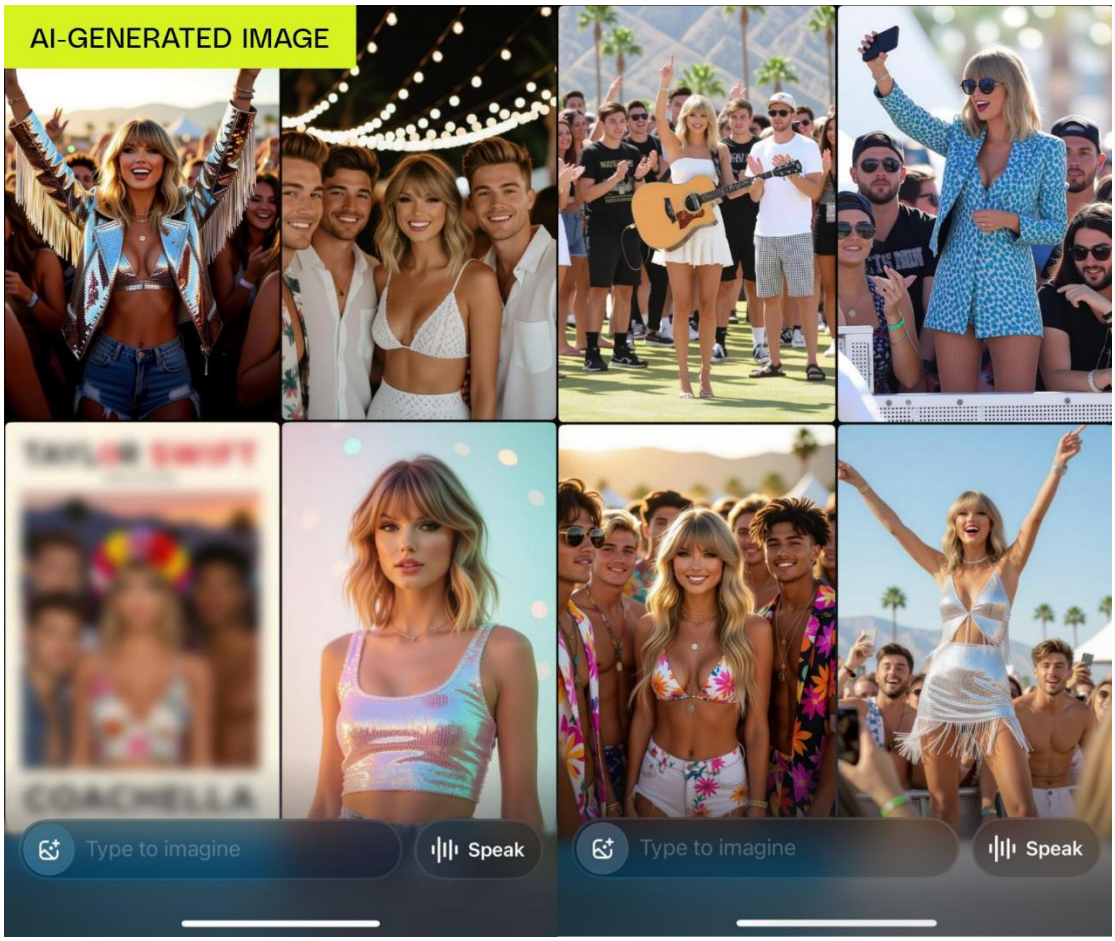




Δημοσιογραφικοί οργανισμοί και fact-checkers αναγκάστηκαν να αναπτύξουν νέες διαδικασίες επαλήθευσης, καθώς πολλές από τις εικόνες παρουσίαζαν υψηλό βαθμό οπτικής πειστικότητας.

ζ) Taylor Swift AI images

Το 2024 η κυκλοφορία ψεύτικων εικόνων της Taylor Swift προκάλεσε διεθνή συζήτηση για τα όρια των generative systems και την ανάγκη ισχυρότερων μηχανισμών προστασίας από συνθετικό περιεχόμενο.



Τα περιστατικά αυτά δείχνουν ότι η Gen AI δεν αποτελεί μελλοντικό σενάριο. Αποτελεί ήδη παράγοντα που επηρεάζει τη δημόσια σφαίρα, τη διαμόρφωση στάσεων και την εμπιστοσύνη στην πληροφορία.

3.7 OECD Public Communication Scan of France: δημόσια επικοινωνία με αποδείξεις και δείκτες

Το OECD Public Communication Scan of France προσφέρει ιδιαίτερα χρήσιμη θεσμική οπτική. Η έκθεση αναλύει πώς η δημόσια επικοινωνία μπορεί να συμβάλει καλύτερα στους κυβερνητικούς στόχους, με έμφαση στη μέτρηση της απόδοσης και του αντίκτυπου. Εξετάζει πρακτικές των γαλλικών διευθύνσεων κυβερνητικής επικοινωνίας και προτείνει την υιοθέτηση πιο προηγμένων μεθόδων evidence-based communication και συστηματικής αξιολόγησης.

Το ενδιαφέρον για την παρούσα μελέτη είναι διπλό. Πρώτον, η OECD δεν αντιμετωπίζει τη δημόσια επικοινωνία ως απλή παραγωγή μηνυμάτων, αλλά ως λειτουργία που πρέπει να συνδέεται με στόχους δημόσιας πολιτικής. Δεύτερον, η έμφαση στη μέτρηση του αντίκτυπου δείχνει ότι η αξιοπιστία δεν μπορεί να θεωρείται αφηρημένη έννοια. Πρέπει να συνδέεται με δείκτες κατανόησης, εμπιστοσύνης, σαφήνειας και αποτελεσματικότητας.

Σε περιβάλλον Gen AI, η λογική αυτή καθίσταται ακόμη πιο σημαντική. Η δημόσια διοίκηση δεν αρκεί να “ενημερώνει”. Πρέπει να γνωρίζει αν η ενημέρωση γίνεται κατανοητή, αν αναγνωρίζεται ως θεσμική, αν μειώνει την αβεβαιότητα και αν ενισχύει την εμπιστοσύνη.

3.8 Σχήμα 1: Ενδεικτική σύγκριση αυθεντικού και AI-generated περιεχομένου

Κριτήριο	Αυθεντικό περιεχόμενο	Συνθετικό / AI-generated περιεχόμενο
Πηγή	Σαφώς προσδιορισμένη, π.χ. θεσμικός φορέας ή επίσημο μέσο	Συχνά ασαφής ή έμμεση
Ιστορικό δημοσίευσης	Διαθέσιμο και επαληθεύσιμο	Περιορισμένο ή ανύπαρκτο
Μεταδεδομένα	Συνεπή με χρόνο, τόπο και πλαίσιο	Ελλιπή, αλλοιωμένα ή απουσιάζουν
Οπτική / γλωσσική ποιότητα	Ρεαλιστική, ενίοτε ατελής	Υψηλή πειστικότητα, συχνά υπερβολικά ομοιογενής
Σχέση με πραγματικό γεγονός	Συνδέεται με επιβεβαιώσιμη πραγματικότητα	Μπορεί να μιμείται γεγονός χωρίς να έχει υπάρξει
Δυνατότητα άμεσης διάκρισης	Υψηλότερη όταν η πηγή είναι γνωστή	Χαμηλή χωρίς τεχνικά εργαλεία
Κίνδυνος παραπλάνησης	Περιορισμένος	Υψηλός, ακόμη και χωρίς πρόθεση εξαπάτησης
Ανάγκη θεσμικής επαλήθευσης	Συνήθως όχι κρίσιμη	Συχνά απαραίτητη

Το σχήμα αποτυπώνει ότι η διάκριση μεταξύ αυθεντικού και AI-generated περιεχομένου δεν βασίζεται πλέον σε προφανή μορφολογικά χαρακτηριστικά. Αντίθετα, απαιτεί συνδυασμό πληροφοριών, όπως πηγή,

μεταδεδομένα και πλαίσιο, οι οποίες σπάνια είναι άμεσα διαθέσιμες στον μέσο χρήστη.

3.9 Σχήμα 2: Πώς ένα παραπλανητικό caption αλλοιώνει αυθεντικό περιεχόμενο

Στοιχείο	Αυθεντικό περιεχόμενο χωρίς παραπλανητικό caption	Το ίδιο περιεχόμενο με παραπλανητικό caption
Οπτικό ή κειμενικό υλικό	Πραγματική εικόνα ή απόσπασμα δήλωσης	Το ίδιο ακριβώς υλικό
Πηγή περιεχομένου	Γνωστή και επαληθεύσιμη	Συχνά αποσιωπάται ή υποβαθμίζεται
Caption / συνοδευτικό κείμενο	Περιγραφικό, ουδέτερο ή επεξηγηματικό	Συναισθηματικά φορτισμένο ή αποσπασματικό
Πλαίσιο ερμηνείας	Συνδέεται με πραγματικό γεγονός και χρονική στιγμή	Αποσυνδέεται από το αρχικό πλαίσιο
Αντίληψη χρήστη	Κατανόηση του γεγονότος ως έχει	Δημιουργία εσφαλμένων συμπερασμάτων
Ρόλος Gen AI	Δεν υπάρχει ή είναι περιορισμένος	Το caption μπορεί να είναι AI-generated
Κίνδυνος παραπληροφόρησης	Περιορισμένος	Υψηλός, χωρίς αλλοίωση του ίδιου υλικού

Ανάγκη θεσμικής αποσαφήνισης	Συνήθως όχι κρίσιμη	Συχνά απαραίτητη

Το σχήμα αναδεικνύει ότι η παραπληροφόρηση δεν προκύπτει μόνο από συνθετικό ή ψευδές περιεχόμενο. Συχνά προκύπτει από την αλλοίωση του ερμηνευτικού πλαισίου μέσω συνοδευτικού κειμένου. Στις περιπτώσεις αυτές, το υλικό μπορεί να είναι αυθεντικό, αλλά η πρόσληψή του να είναι παραπλανητική.

Κεφάλαιο 4

Η ψυχολογία της εμπιστοσύνης στην ψηφιακή εποχή

4.1 Η εμπιστοσύνη ως γνωστική και συναισθηματική διαδικασία

Η εμπιστοσύνη δεν αποτελεί αποκλειστικά ορθολογική κρίση βασισμένη σε δεδομένα και αποδείξεις. Συγκροτείται ως σύνθετη διαδικασία, όπου συνυπάρχουν γνωστικά, συναισθηματικά και κοινωνικά στοιχεία.

Ο Kahneman (2011) διακρίνει δύο τρόπους σκέψης: τον ταχύ, διαισθητικό και συναισθηματικά φορτισμένο τρόπο, και τον αργό, αναλυτικό και απαιτητικό σε γνωστικούς πόρους. Στο ψηφιακό περιβάλλον, ιδίως σε συνθήκες συνεχούς ροής πληροφορίας, η αξιολόγηση της αξιοπιστίας πραγματοποιείται συχνά μέσω ταχύτερων και λιγότερο αναλυτικών διαδικασιών.

Για τη δημόσια διοίκηση, αυτό σημαίνει ότι η θεσμική αξιοπιστία δεν κρίνεται μόνο από την ακρίβεια των δεδομένων, αλλά από τον τρόπο με τον οποίο αυτά προσλαμβάνονται από τον πολίτη.

4.2 Ευεργετικές και γνωστικές συντομεύσεις στην αξιολόγηση της πληροφορίας

Η ανθρώπινη λήψη αποφάσεων δεν βασίζεται πάντοτε σε εκτενή επεξεργασία πληροφοριών. Αντιθέτως, οι άνθρωποι χρησιμοποιούν συχνά ευρετικές (heuristics), δηλαδή νοητικές συντομεύσεις που επιτρέπουν γρήγορες κρίσεις υπό συνθήκες αβεβαιότητας. Οι Tversky και Kahneman (1974) έδειξαν ότι οι ευρετικές αυτές είναι απαραίτητες για την καθημερινή λειτουργία του ατόμου, αλλά ταυτόχρονα μπορούν να οδηγήσουν σε συστηματικά σφάλματα κρίσης.

Στο ψηφιακό περιβάλλον, όπου οι πολίτες εκτίθενται καθημερινά σε χιλιάδες πληροφοριακά ερεθίσματα, η χρήση τέτοιων γνωστικών μηχανισμών γίνεται ακόμη πιο έντονη. Οι χρήστες δεν αξιολογούν κάθε πληροφορία με τον ίδιο βαθμό προσοχής. Αντίθετα, συχνά βασίζονται σε

ενδείξεις όπως η εμφάνιση ενός περιεχομένου, ο αριθμός κοινοποιήσεων ή η αναγνωρισιμότητα του προσώπου που το δημοσιεύει.

Η διαπίστωση αυτή αποκτά ιδιαίτερη σημασία στην εποχή της Generative AI. Όσο πιο ρεαλιστικό γίνεται το συνθετικό περιεχόμενο, τόσο περισσότερο ενεργοποιούνται οι ίδιες γνωστικές συντομεύσεις που χρησιμοποιούνται και για την αξιολόγηση αυθεντικού περιεχομένου. Ως αποτέλεσμα, η διάκριση μεταξύ αληθινού και συνθετικού περιεχομένου γίνεται δυσκολότερη όχι μόνο λόγω της τεχνολογίας αλλά και λόγω του τρόπου με τον οποίο λειτουργεί η ανθρώπινη αντίληψη.

4.3 Εμπιστοσύνη, αυθεντία και πηγή πληροφορίας

Η θεσμική αυθεντία λειτουργεί παραδοσιακά ως μηχανισμός μείωσης αβεβαιότητας. Ο πολίτης δεν επαληθεύει κάθε πληροφορία από την αρχή, αλλά βασίζεται στην αξιοπιστία της πηγής.

Ωστόσο, το ψηφιακό περιβάλλον μεταβάλλει αυτή τη δυναμική. Η θεσμική πηγή ανταγωνίζεται πλέον μη θεσμικές πηγές μέσα στο ίδιο περιβάλλον. Η αναγνωρισιμότητα ενός θεσμού δεν εγγυάται αυτόματα αυξημένη προσοχή. Η εμπιστοσύνη λειτουργεί περισσότερο ως πλαίσιο ερμηνείας παρά ως φίλτρο που αποκλείει την αμφιβολία.

Το OECD Trust Survey δείχνει ότι η εμπιστοσύνη συνδέεται με εμπειρίες και προσδοκίες από τους θεσμούς. Αυτό σημαίνει ότι η δημόσια επικοινωνία δεν μπορεί να αποσπαστεί από τη συνολική εμπειρία του πολίτη από τη διοίκηση.

4.4 Authority Bias: γιατί εξακολουθούμε να πιστεύουμε τους «ειδικούς»

Η αξιολόγηση της πληροφορίας επηρεάζεται σημαντικά από την αντίληψη αυθεντίας. Το φαινόμενο Authority Bias περιγράφει την τάση των ατόμων να αποδίδουν αυξημένη αξιοπιστία σε πληροφορίες που

προέρχονται από πρόσωπα ή φορείς που θεωρούνται ειδικοί, ακόμη και όταν δεν έχουν επαληθεύσει ανεξάρτητα το περιεχόμενό τους.

Στο περιβάλλον των social media, η έννοια της αυθεντίας έχει μετασχηματιστεί. Η θεσμική ιδιότητα δεν αποτελεί πλέον τη μοναδική πηγή κύρους. Influencers, creators και δημοφιλείς σχολιαστές συχνά αποκτούν υψηλά επίπεδα εμπιστοσύνης από το κοινό, χωρίς να διαθέτουν επίσημο θεσμικό ρόλο.

Η εξέλιξη αυτή δημιουργεί νέες προκλήσεις για τη δημόσια διοίκηση. Οι πολίτες ενδέχεται να αξιολογούν μια πληροφορία όχι με βάση την πηγή της αλλά με βάση το πρόσωπο που την αναπαράγει ή την ερμηνεύει.

4.5 Information Overload και Trust Fatigue

Η υπερπληθώρα πληροφοριών αποτελεί ένα από τα χαρακτηριστικά της ψηφιακής εποχής. Σύμφωνα με τη σχετική βιβλιογραφία, η συνεχής έκθεση σε μεγάλους όγκους πληροφοριών αυξάνει το γνωστικό φορτίο των χρηστών και περιορίζει τη δυνατότητα προσεκτικής αξιολόγησης κάθε πληροφορίας ξεχωριστά.

Το φαινόμενο αυτό συνδέεται με την έννοια της trust fatigue. Όταν οι πολίτες εκτίθενται διαρκώς σε αντικρουόμενες πληροφορίες, επαληθεύσεις, διαψεύσεις και αντικρουόμενες αφηγήσεις, ενδέχεται να οδηγηθούν όχι σε μεγαλύτερη προσοχή αλλά σε γενικευμένη δυσπιστία.

Στην περίπτωση της δημόσιας διοίκησης, η trust fatigue μπορεί να οδηγήσει σε αποδυνάμωση της αποτελεσματικότητας ακόμη και έγκυρων ενημερωτικών παρεμβάσεων.

4.6 Γνωστικά φίλτρα, συναισθηματικά triggers και affective truth

Η σύγχρονη έρευνα για την παραπληροφόρηση δείχνει ότι οι πολίτες δεν εξαπατώνται κυρίως λόγω έλλειψης γνώσης, αλλά λόγω της λειτουργίας

γνωστικών φίλτρων, περιορισμένης προσοχής και συναισθηματικών triggers.

Η εργασία των Pennycook et al. (2021) δείχνει ότι η μετατόπιση της προσοχής προς την ακρίβεια μπορεί να μειώσει τη διάδοση παραπληροφόρησης. Το εύρημα αυτό είναι σημαντικό διότι υποδηλώνει ότι οι πολίτες συχνά διαθέτουν προτίμηση για ακριβή πληροφόρηση, αλλά δεν την ενεργοποιούν πάντοτε μέσα σε περιβάλλοντα ταχύτητας και κοινωνικής πίεσης.

Στο πλαίσιο της Gen AI, τα συναισθηματικά triggers ενισχύονται. Το περιεχόμενο μπορεί να είναι πιο καλογραμμένο, πιο οπτικά πειστικό και πιο προσαρμοσμένο στις προσδοκίες του κοινού. Αυτό σημαίνει ότι η εμπιστοσύνη ή η αποδοχή μπορεί να προηγείται της κριτικής επεξεργασίας.

Η έννοια της affective truth, ή «συναισθηματικής αλήθειας», περιγράφει την τάση των ανθρώπων να αποδέχονται ως αληθινό περιεχόμενο που ανταποκρίνεται στα συναισθήματα, στις προσδοκίες ή στις προϋπάρχουσες αντιλήψεις τους, ακόμη και όταν η πραγματική ακρίβεια του περιεχομένου δεν έχει επαληθευθεί.

Η έννοια της affective truth περιγράφει την τάση των ανθρώπων να αποδέχονται ως αληθινό περιεχόμενο που ανταποκρίνεται στα συναισθήματα, στις προσδοκίες ή στις προϋπάρχουσες αντιλήψεις τους, ακόμη και όταν η πραγματική ακρίβεια του περιεχομένου δεν έχει επαληθευθεί. Στο ψηφιακό περιβάλλον, η αξιολόγηση μιας πληροφορίας συχνά προηγείται της επαλήθευσής της. Το ερώτημα δεν είναι πάντοτε «είναι ακριβές;», αλλά «μου φαίνεται εύλογο;», «ταιριάζει με όσα ήδη πιστεύω;» ή «προκαλεί συναισθηματική ανταπόκριση;».

Η λογική αυτή συνδέεται άμεσα με τα φαινόμενα confirmation bias και cognitive ease που περιγράφονται στη σχετική βιβλιογραφία. Όσο πιο οικείο, κατανοητό ή συναισθηματικά φορτισμένο είναι ένα μήνυμα, τόσο αυξάνεται η πιθανότητα να γίνει αποδεκτό χωρίς εκτεταμένη κριτική επεξεργασία. Στην εποχή της Generative AI, όπου η παραγωγή ιδιαίτερα

πειστικών αφηγήσεων, εικόνων και βίντεο είναι τεχνικά εύκολη, η affective truth αποκτά ιδιαίτερη σημασία για τη δημόσια διοίκηση, καθώς η εμπιστοσύνη των πολιτών μπορεί να επηρεάζεται περισσότερο από την πειστικότητα μιας αφήγησης παρά από την πραγματική της ακρίβεια.

Για τον λόγο αυτό, η θεσμική επικοινωνία χρειάζεται να συνδυάζει ακρίβεια με ερμηνευτική σαφήνεια. Δεν αρκεί να παρέχει ορθά στοιχεία: πρέπει να εξηγεί το πλαίσιο, να προλαμβάνει παρανοήσεις και να μειώνει το περιθώριο συναισθηματικής ή αποσπασματικής ανάγνωσης της πληροφορίας.

4.7 Εμπιστοσύνη και θεσμική κόπωση

Σε περιβάλλοντα υπερπληροφόρησης, οι πολίτες δεν χάνουν απαραίτητα την εμπιστοσύνη στους θεσμούς. Μπορεί όμως να εμφανίζουν μειωμένη διάθεση εμπλοκής, περιορισμένη προσοχή στον επίσημο λόγο και αυξημένη εξάρτηση από απλουστευτικές αφηγήσεις.

Η Gen AI επιτείνει το φαινόμενο, διότι αυξάνει τον όγκο και την πολυπλοκότητα της πληροφορίας χωρίς αντίστοιχη αύξηση της γνωστικής αντοχής των πολιτών. Η δημόσια διοίκηση οφείλει, επομένως, να σχεδιάζει τον λόγο της όχι μόνο με βάση την πληρότητα, αλλά και με βάση την κατανόηση.

Η θεσμική κόπωση συνδέεται επίσης με τη συνεχή ανάγκη επαλήθευσης. Όταν ο πολίτης αισθάνεται ότι κάθε εικόνα, βίντεο, δήλωση ή επίσημη πληροφορία πρέπει να ελέγχεται εκ νέου, η εμπιστοσύνη δεν αντικαθίσταται πάντα από κριτική σκέψη· συχνά αντικαθίσταται από αποστασιοποίηση, αδιαφορία ή γενικευμένη καχυποψία. Σε αυτό το περιβάλλον, η δημόσια διοίκηση χρειάζεται να μειώνει το γνωστικό βάρος του πολίτη, παρέχοντας σαφή, συνεπή και εύκολα επαληθεύσιμη πληροφόρηση.

Σε πρακτικό επίπεδο, αυτό σημαίνει ότι ένα μήνυμα μπορεί να γίνει αποδεκτό ως «αληθοφανές» ακόμη και πριν ελεγχθεί θεσμικά. Η δημόσια διοίκηση οφείλει επομένως να δίνει ιδιαίτερη προσοχή όχι μόνο στο περιεχόμενο της πληροφορίας, αλλά και στον τρόπο με τον οποίο αυτή

πλαισιώνεται, συνοδεύεται και επανακυκλοφορεί σε ψηφιακά περιβάλλοντα.

4.8 Εμπιστοσύνη, Gen AI και αντίληψη ελέγχου

Ένα επαναλαμβανόμενο εύρημα στη βιβλιογραφία για την τεχνολογία και την εμπιστοσύνη είναι ότι οι πολίτες δεν ανησυχούν μόνο για την τεχνολογία καθαυτή, αλλά για την αίσθηση απώλειας ελέγχου.

Η UNESCO, στην καθοδήγησή της για τη Generative AI στην εκπαίδευση και την έρευνα, επισημαίνει την ανάγκη ανθρώπινης-κεντρικής προσέγγισης, ανάπτυξης πολιτικών, προστασίας δεδομένων και ενίσχυσης ανθρώπινων ικανοτήτων. Παρότι το κείμενο εστιάζει στην εκπαίδευση και την έρευνα, τα συμπεράσματά του έχουν ευρύτερη θεσμική σημασία: η Gen AI χρειάζεται πλαίσιο, διαφάνεια και ανθρώπινη ευθύνη.

Τα φαινόμενα αυτά δεν επηρεάζουν μόνο τους δημόσιους θεσμούς αλλά και οργανισμούς, επιχειρήσεις και brands, των οποίων η αξιοπιστία συνδέεται πλέον με την ικανότητα επαλήθευσης της προέλευσης του ψηφιακού περιεχομένου.

Για τη δημόσια διοίκηση, η εμπιστοσύνη συνδέεται άρρηκτα με τη δυνατότητα του πολίτη να γνωρίζει πότε χρησιμοποιείται Gen AI, για ποιο σκοπό και ποιος φέρει την τελική ευθύνη.

Η ανάλυση της εμπιστοσύνης οδηγεί έτσι στο επόμενο ερώτημα: ποιοι επηρεάζουν σήμερα την πρόσληψη της δημόσιας πληροφορίας και με ποιους μηχανισμούς διαμορφώνεται η αξιοπιστία στο ψηφιακό περιβάλλον. Το ζήτημα αυτό εξετάζεται στο επόμενο κεφάλαιο, μέσα από τον ρόλο των μη θεσμικών φορέων επιρροής.

Κεφάλαιο 5

Δημόσια διοίκηση, θεσμικός λόγος και μη θεσμικοί φορείς επιρροής

5.1 Εισαγωγή: Influencers, parasocial trust και δημόσιος λόγος στην εποχή της Gen AI

5.1.1 Από τη θεσμική αυθεντία στην ψηφιακή επιρροή

Για μεγάλο μέρος του 20ού αιώνα, η επιρροή στη δημόσια σφαίρα συνδεόταν κυρίως με θεσμούς. Κυβερνήσεις, μέσα ενημέρωσης, πανεπιστήμια και δημόσιοι οργανισμοί λειτουργούσαν ως βασικοί διαμεσολαβητές της πληροφορίας. Η αξιοπιστία ενός μηνύματος αντλούσε σημαντικό μέρος της ισχύος της από την ιδιότητα και το κύρος του φορέα που το εξέπεμπε.

Η εξάπλωση των κοινωνικών δικτύων τροποποίησε σταδιακά αυτή τη σχέση. Η δυνατότητα άμεσης δημοσίευσης περιεχομένου επέτρεψε την εμφάνιση νέων μορφών επιρροής που δεν βασίζονται απαραίτητα σε θεσμική νομιμοποίηση αλλά σε προσωπική προβολή, ψηφιακή εγγύτητα και συνεχή αλληλεπίδραση με το κοινό.

Στο νέο αυτό περιβάλλον, η επιρροή δεν αποτελεί αποκλειστικό προνόμιο των θεσμών. Δημιουργοί περιεχομένου, σχολιαστές, influencers και ψηφιακές κοινότητες μπορούν να επηρεάζουν τη δημόσια συζήτηση, τη διαμόρφωση στάσεων και την αντίληψη για δημόσιες πολιτικές με τρόπο αντίστοιχο ή ακόμη και ισχυρότερο από παραδοσιακούς θεσμικούς φορείς.

Η εξέλιξη αυτή δεν συνεπάγεται απαραίτητα υποχώρηση της σημασίας των θεσμών. Υποδηλώνει όμως ότι η θεσμική αξιοπιστία λειτουργεί πλέον σε ένα περισσότερο ανταγωνιστικό περιβάλλον προσοχής και επιρροής.

5.1.2 Η οικονομία της προσοχής και οι μηχανισμοί επιρροής

Οι πλατφόρμες κοινωνικής δικτύωσης λειτουργούν με βάση την οικονομία της προσοχής (attention economy). Σε ένα περιβάλλον όπου η πληροφορία είναι άφθονη αλλά η ανθρώπινη προσοχή περιορισμένη, η επιτυχία ενός μηνύματος εξαρτάται από την ικανότητά του να προσελκύει και να διατηρεί το ενδιαφέρον του χρήστη.

Οι αλγόριθμοι των πλατφορμών ευνοούν συχνά περιεχόμενο που:

- προκαλεί συναισθηματική αντίδραση,
- δημιουργεί αίσθηση ταύτισης,
- ενθαρρύνει αλληλεπίδραση,
- και οδηγεί σε αυξημένο χρόνο παραμονής.

Αυτό σημαίνει ότι η ορατότητα μιας πληροφορίας δεν καθορίζεται αποκλειστικά από την ακρίβεια ή τη σημασία της αλλά και από την ικανότητά της να ανταγωνίζεται χιλιάδες άλλα ερεθίσματα.

Για τη δημόσια διοίκηση, η πραγματικότητα αυτή δημιουργεί μια σημαντική πρόκληση: η εγκυρότητα μιας πληροφορίας δεν αρκεί πλέον από μόνη της για να εξασφαλίσει ότι η πληροφορία θα φτάσει στους πολίτες.

5.1.3 Parasocial relationships και εμπιστοσύνη

Η βιβλιογραφία για τα μέσα επικοινωνίας χρησιμοποιεί τον όρο parasocial relationships για να περιγράψει τις μονομερείς σχέσεις που αναπτύσσουν οι πολίτες με δημόσια πρόσωπα, παρουσιαστές, δημιουργούς περιεχομένου ή influencers.

Παρότι οι σχέσεις αυτές δεν είναι πραγματικά αμοιβαίες, δημιουργούν συχνά αίσθηση οικειότητας, προσωπικής σύνδεσης και εμπιστοσύνης.

Οι χρήστες μπορεί να αισθάνονται ότι:

- γνωρίζουν τον δημιουργό περιεχομένου,

- κατανοούν τον τρόπο σκέψης του,
- μοιράζονται κοινές εμπειρίες,
- ή συμμετέχουν σε μια κοινότητα κοινών αξιών.

Η αίσθηση αυτή συχνά λειτουργεί ως μηχανισμός ενίσχυσης της αξιοπιστίας. Οι πληροφορίες που μεταδίδονται μέσω προσώπων με τα οποία έχει αναπτυχθεί parasocial σχέση γίνονται ευκολότερα αποδεκτές σε σύγκριση με πληροφορίες που προέρχονται από απρόσωπους οργανισμούς.

Η διαπίστωση αυτή δεν σημαίνει ότι οι πολίτες απορρίπτουν τους θεσμούς. Δείχνει όμως ότι η εμπιστοσύνη συγκροτείται πλέον μέσα από διαφορετικούς μηχανισμούς από εκείνους που κυριαρχούσαν στις προηγούμενες δεκαετίες**.**

5.1.4 Οι influencers ως διαμεσολαβητές δημόσιας πληροφορίας

Τα τελευταία χρόνια παρατηρείται διεθνώς η εμφάνιση των λεγόμενων news influencers, δηλαδή δημιουργών περιεχομένου που σχολιάζουν πολιτικές εξελίξεις, κοινωνικά ζητήματα και δημόσιες πολιτικές.

Σύμφωνα με το Reuters Digital News Report 2024, οι νεότερες ηλικίες αντλούν ολοένα και μεγαλύτερο μέρος της ενημέρωσής τους από πλατφόρμες όπως το TikTok, το Instagram και το YouTube. Η μεταβολή αυτή σημαίνει ότι η πρόσβαση στην πληροφορία πραγματοποιείται συχνά μέσω προσώπων και όχι απευθείας μέσω δημοσιογραφικών ή θεσμικών οργανισμών.

Οι influencers δεν υποκαθιστούν απαραίτητα τη δημοσιογραφία ή τους θεσμούς. Λειτουργούν όμως ως ενδιάμεσοι ερμηνευτές της πραγματικότητας. Επιλέγουν ποια ζητήματα θα αναδείξουν, πώς θα τα παρουσιάσουν και ποια στοιχεία θα τονίσουν.

Από τη σκοπιά της δημόσιας διοίκησης, η εξέλιξη αυτή σημαίνει ότι η δημόσια πληροφορία δεν διακινείται πλέον μόνο μέσω θεσμικών καναλιών αλλά και μέσω ενός σύνθετου οικοσυστήματος ψηφιακής επιρροής**.**

5.1.5 Influencers και Generative AI

Η διάδοση της Generative AI προσθέτει μια νέα διάσταση στο οικοσύστημα επιρροής.

Η δυνατότητα παραγωγής εικόνων, βίντεο, ηχητικών αποσπασμάτων και κειμένων υψηλής ποιότητας επιτρέπει τη δημιουργία περιεχομένου σε πρωτοφανή κλίμακα. Παράλληλα, δημιουργεί νέους κινδύνους για τη διασπορά παραπλανητικού περιεχομένου.

Οι influencers και οι creators μπορούν να λειτουργήσουν είτε ως πολλαπλασιαστές παραπληροφόρησης είτε ως φορείς αποσαφήνισης και επαλήθευσης. Η διττή αυτή δυνατότητα αναδεικνύει τη σημασία της ψηφιακής παιδείας όχι μόνο για τους πολίτες αλλά και για όσους συμμετέχουν ενεργά στη δημόσια επικοινωνία.

Η δημόσια διοίκηση καλείται συνεπώς να κατανοήσει ότι η αντιμετώπιση των κινδύνων της Gen AI δεν αφορά μόνο την τεχνολογία αλλά και το ευρύτερο οικοσύστημα επιρροής μέσα στο οποίο κυκλοφορεί η πληροφορία.

5.1.6 Τι σημαίνει αυτό για τη δημόσια διοίκηση

Η ύπαρξη ενός περιβάλλοντος όπου η επιρροή κατανέμεται μεταξύ θεσμών, μέσων ενημέρωσης, influencers και ψηφιακών κοινοτήτων δεν συνεπάγεται ότι η δημόσια διοίκηση πρέπει να υιοθετήσει πρακτικές influencing.

Αντίθετα, η πρόκληση είναι διαφορετική.

Η δημόσια διοίκηση οφείλει να παραμείνει θεσμικός φορέας αξιοπιστίας, διατηρώντας ταυτόχρονα την ικανότητα να επικοινωνεί αποτελεσματικά σε ένα περιβάλλον όπου η προσοχή είναι περιορισμένη και ο ανταγωνισμός για ορατότητα ιδιαίτερα έντονος.

Η στρατηγική απάντηση δεν είναι η μίμηση των influencers αλλά η ενίσχυση:

- της σαφήνειας,
- της διαφάνειας,
- της συνέπειας,
- και της επαληθευσιμότητας της δημόσιας πληροφορίας.

Με τον τρόπο αυτό, η δημόσια διοίκηση μπορεί να διατηρήσει τον ρόλο της ως σημείο αναφοράς αξιοπιστίας ακόμη και σε ένα οικοσύστημα όπου η επιρροή δεν ασκείται πλέον αποκλειστικά από θεσμούς.

5.2 Συμπέρασμα: θεσμικός λόγος σε περιβάλλον κατανεμημένης επιρροής

Η ανάλυση του κεφαλαίου δείχνει ότι η δημόσια διοίκηση λειτουργεί πλέον σε ένα περιβάλλον όπου η επιρροή δεν ασκείται αποκλειστικά από θεσμούς, μέσα ενημέρωσης ή επίσημες πηγές. Influencers, creators και ψηφιακές κοινότητες μπορούν να λειτουργούν ως δευτερογενείς ερμηνευτές της δημόσιας πληροφορίας, επηρεάζοντας τον τρόπο με τον οποίο οι πολίτες κατανοούν δημόσιες πολιτικές, κρίσεις και θεσμικά μηνύματα.

Το κρίσιμο συμπέρασμα δεν είναι ότι η δημόσια διοίκηση πρέπει να μιμηθεί πρακτικές influencing. Αντίθετα, οφείλει να διατηρήσει τον διακριτό θεσμικό της ρόλο, ενισχύοντας τη σαφήνεια, τη συνέπεια, τη διαφάνεια και την επαληθευσιμότητα του δημόσιου λόγου της. Η θεσμική αξιοπιστία δεν εξαρτάται από την προσπάθεια ανταγωνισμού με μη θεσμικούς φορείς στο πεδίο της ορατότητας, αλλά από την ικανότητα της διοίκησης να παραμένει σταθερό σημείο αναφοράς μέσα σε ένα ρευστό πληροφοριακό περιβάλλον.

Η διαπίστωση αυτή οδηγεί στο επόμενο ερώτημα της μελέτης: ποιο θεσμικό πλαίσιο χρειάζεται ώστε η δημόσια διοίκηση να μπορεί να διαχειρίζεται την ψηφιακή αλήθεια, τη χρήση της Gen AI και την εμπιστοσύνη των πολιτών με όρους λογοδοσίας και δημόσιου συμφέροντος.

Κεφάλαιο 6

Πρόταση Διαχείρισης της Ψηφιακής Αλήθειας στη Δημόσια Διοίκηση

6.1 Γιατί απαιτείται θεσμικό πλαίσιο και όχι αποσπασματικές λύσεις

Η πρόκληση της ψηφιακής αλήθειας στην εποχή της Gen AI δεν μπορεί να αντιμετωπιστεί με μεμονωμένες παρεμβάσεις, τεχνικές λύσεις ή αποσπασματικές επικοινωνιακές πρακτικές. Η αβεβαιότητα ως προς την αυθεντικότητα του περιεχομένου, η κόπωση από την ανάγκη διαρκούς επαλήθευσης και η συνύπαρξη θεσμικών και μη θεσμικών φορέων επιρροής συγκροτούν μια δομική συνθήκη.

Σε αυτό το πλαίσιο, η δημόσια διοίκηση δεν καλείται να ελέγξει το σύνολο της πληροφορίας, αλλά να διαχειριστεί θεσμικά την αλήθεια ως δημόσιο αγαθό.

Πριν διατυπωθούν οι επιμέρους πυλώνες μιας τέτοιας προσέγγισης, είναι χρήσιμο να αποτυπωθεί πρώτα το ελληνικό πλαίσιο μέσα στο οποίο καλείται να εφαρμοστεί η διαχείριση της ψηφιακής αλήθειας.

6.1Α Η ελληνική δημόσια διοίκηση απέναντι στη Generative AI

Η Ελλάδα βρίσκεται σε φάση σταδιακής ενσωμάτωσης της τεχνητής νοημοσύνης στον δημόσιο τομέα, με έμφαση τόσο στον ψηφιακό μετασχηματισμό υπηρεσιών όσο και στη διαμόρφωση στρατηγικού και κανονιστικού πλαισίου.

Η Βίβλος Ψηφιακού Μετασχηματισμού 2020–2025 εντάσσει την τεχνητή νοημοσύνη στους βασικούς άξονες του ψηφιακού μετασχηματισμού της χώρας, ενώ το Υπουργείο Ψηφιακής Διακυβέρνησης έχει αναλάβει κεντρικό ρόλο στον σχεδιασμό σχετικών πολιτικών και πρωτοβουλιών.

Η πρωτοβουλία European Commission AI Watch επισημαίνει ότι η επιτυχής ενσωμάτωση της τεχνητής νοημοσύνης στον δημόσιο τομέα δεν εξαρτάται αποκλειστικά από την τεχνολογική υποδομή, αλλά και από την

ανάπτυξη δεξιοτήτων, την ενίσχυση της εμπιστοσύνης των πολιτών και τη διαμόρφωση μηχανισμών διακυβέρνησης. Στην περίπτωση της Ελλάδας, η σχετική ανάλυση συνδέει την ανάπτυξη της AI με τον ευρύτερο ψηφιακό μετασχηματισμό της δημόσιας διοίκησης και την ανάγκη θεσμικής προσαρμογής σε ζητήματα διαφάνειας και λογοδοσίας.

Η έκθεση European Commission AI Watch για την Ελλάδα αναφέρει ότι η χώρα έχει προχωρήσει στη διαμόρφωση εθνικής στρατηγικής για την AI, με έμφαση:

- στη δημόσια διοίκηση,
- στην καινοτομία,
- στις ψηφιακές υπηρεσίες,
- και στη δημιουργία πλαισίου εμπιστοσύνης για τη χρήση συστημάτων AI.

Παράλληλα, η μελέτη "Generative AI Greece 2030", που εκπονήθηκε από την Ειδική Γραμματεία Μακροπρόθεσμου Σχεδιασμού, το ΕΚΚΕ και το ΕΚΕΦΕ "Δημόκριτος", αποτυπώνει πιθανά σενάρια εξέλιξης της Gen AI στην Ελλάδα και αναδεικνύει κρίσιμα ζητήματα:

- θεσμικής προσαρμογής,
- ψηφιακής παιδείας,
- ανθρώπινου δυναμικού,
- και δημοκρατικής διακυβέρνησης.

Οι εξελίξεις αυτές δείχνουν ότι η ελληνική δημόσια διοίκηση δεν βρίσκεται εκτός της παγκόσμιας μετάβασης προς την Gen AI. Αντίθετα, εισέρχεται σταδιακά σε φάση πρακτικής αξιοποίησης εργαλείων AI, γεγονός που καθιστά ακόμη πιο σημαντική την ανάγκη για:

- διαφάνεια,
- ανθρώπινη εποπτεία,
- θεσμική λογοδοσία,

- και σαφές πλαίσιο δημόσιας επικοινωνίας.

Η ελληνική περίπτωση αναδεικνύει την ανάγκη προσαρμογής του θεσμικού πλαισίου στις νέες τεχνολογικές συνθήκες. Για να αποτιμηθεί καλύτερα αυτή η ανάγκη, είναι χρήσιμη η σύγκριση με διεθνείς διοικήσεις που έχουν ήδη αναπτύξει πιο ώριμες πρακτικές για τη χρήση της Gen AI.

6.1Β Διεθνείς πρακτικές δημόσιας διοίκησης απέναντι στην Gen AI

Η διεθνής εμπειρία δείχνει ότι η αντιμετώπιση των προκλήσεων που δημιουργεί η Generative AI δεν περιορίζεται στην τεχνολογική ανάπτυξη ή στη ρυθμιστική παρέμβαση. Πολλές δημόσιες διοικήσεις έχουν ήδη αναπτύξει στρατηγικές που συνδυάζουν εκπαίδευση στελεχών, θεσμική επικοινωνία, διακυβέρνηση δεδομένων και μηχανισμούς λογοδοσίας.

Η προσέγγιση αυτή αναγνωρίζει ότι η εμπιστοσύνη προς τις δημόσιες υπηρεσίες εξαρτάται όχι μόνο από την ποιότητα των τεχνολογικών εργαλείων αλλά και από τον τρόπο με τον οποίο οι οργανισμοί επικοινωνούν τη χρήση τους.

Γαλλία: Η μέτρηση της δημόσιας επικοινωνίας ως εργαλείο ενίσχυσης της εμπιστοσύνης

Η περίπτωση της Γαλλίας παρουσιάζει ιδιαίτερο ενδιαφέρον, καθώς δεν εστιάζει πρωτίστως στην τεχνολογία αλλά στη θεσμική αξιοπιστία της δημόσιας επικοινωνίας. Το OECD Public Communication Scan of France αντιμετωπίζει τη δημόσια επικοινωνία ως στρατηγική λειτουργία του κράτους και όχι ως απλό μηχανισμό διάχυσης πληροφοριών. Βασική διαπίστωση της μελέτης είναι ότι οι δημόσιοι οργανισμοί δεν αρκεί να επικοινωνούν· οφείλουν να γνωρίζουν αν η επικοινωνία τους παράγει κατανόηση, εμπιστοσύνη και αλλαγή συμπεριφοράς.

Ο ΟΟΣΑ προτείνει την υιοθέτηση συστηματικών μεθόδων αξιολόγησης της δημόσιας επικοινωνίας, τη χρήση δεδομένων για τον σχεδιασμό παρεμβάσεων και τη στενότερη σύνδεση μεταξύ υπευθύνων επικοινωνίας και υπευθύνων δημόσιας πολιτικής. Η επικοινωνία δεν αντιμετωπίζεται

ως παράλληλη δραστηριότητα αλλά ως στοιχείο της ίδιας της διακυβέρνησης.

Ιδιαίτερη σημασία για την παρούσα μελέτη έχει το γεγονός ότι η γαλλική προσέγγιση αναγνωρίζει το σύνθετο πληροφοριακό περιβάλλον στο οποίο λειτουργούν σήμερα οι δημόσιοι θεσμοί. Σε συνθήκες υπερπληροφόρησης, παραπληροφόρησης και αυξανόμενης χρήσης εργαλείων τεχνητής νοημοσύνης, η αποτελεσματικότητα της δημόσιας επικοινωνίας δεν μπορεί να μετράται αποκλειστικά με δείκτες απήχησης ή δημοσιότητας. Χρειάζεται να συνδέεται με δείκτες κατανόησης, αξιοπιστίας και εμπιστοσύνης των πολιτών.

Η γαλλική περίπτωση δείχνει ότι η ενίσχυση της ψηφιακής αλήθειας δεν εξαρτάται μόνο από τεχνολογικά εργαλεία επαλήθευσης, αλλά και από την ικανότητα των δημόσιων φορέων να αξιολογούν συστηματικά αν η θεσμική πληροφορία γίνεται αντιληπτή ως αξιόπιστη και αναγνωρίσιμη.

Σιγκαπούρη: Η τεχνητή νοημοσύνη ως δημόσιο αγαθό

Η Σιγκαπούρη αποτελεί ένα από τα πλέον χαρακτηριστικά παραδείγματα σύνδεσης της τεχνητής νοημοσύνης με τη δημόσια αξία. Η National AI Strategy 2.0, που παρουσιάστηκε το 2023, υιοθετεί ως κεντρικό όραμα το «AI for the Public Good, for Singapore and the World». Η προσέγγιση αυτή αντιμετωπίζει την AI όχι μόνο ως εργαλείο οικονομικής ανάπτυξης αλλά και ως μηχανισμό βελτίωσης της δημόσιας διοίκησης και της ποιότητας ζωής των πολιτών.

Η στρατηγική δίνει ιδιαίτερη έμφαση σε δύο έννοιες: την αριστεία (Excellence) και την ενδυνάμωση (Empowerment). Παράλληλα, αναγνωρίζει ότι η επιτυχία της AI εξαρτάται από την ύπαρξη εμπιστοσύνης, διαφάνειας και υπεύθυνης χρήσης των συστημάτων. Για τον λόγο αυτό, η ανάπτυξη τεχνολογικών δυνατοτήτων συνοδεύεται από επενδύσεις σε δεξιότητες, θεσμική διακυβέρνηση και μηχανισμούς λογοδοσίας.

Ιδιαίτερο ενδιαφέρον παρουσιάζει η έμφαση που δίνεται στην έννοια του «δημόσιου οφέλους». Η AI δεν παρουσιάζεται ως ουδέτερη τεχνολογία

αλλά ως μέσο που πρέπει να υπηρετεί σαφείς κοινωνικούς και διοικητικούς στόχους. Η λογική αυτή συνδέεται άμεσα με το ζήτημα της ψηφιακής αλήθειας, καθώς μεταφέρει τη συζήτηση από το ερώτημα «τι μπορεί να κάνει η τεχνολογία» στο ερώτημα «πώς η τεχνολογία μπορεί να χρησιμοποιηθεί με τρόπο που να διατηρεί την εμπιστοσύνη των πολιτών».

Η περίπτωση της Σιγκαπούρης αναδεικνύει ότι η διαφάνεια, η κατανόηση του τρόπου λειτουργίας των συστημάτων και η σαφής ανάληψη ευθύνης από τους δημόσιους φορείς αποτελούν προϋποθέσεις για τη βιώσιμη ενσωμάτωση της τεχνητής νοημοσύνης στη δημόσια διοίκηση.

Εσθονία: Η εμπιστοσύνη ως αποτέλεσμα ψηφιακής διακυβέρνησης

Η Εσθονία θεωρείται διεθνώς ένα από τα πλέον ώριμα ψηφιακά κράτη. Η προσέγγισή της στην τεχνητή νοημοσύνη δεν αναπτύσσεται αποσπασματικά αλλά εντάσσεται σε ένα ευρύτερο μοντέλο ψηφιακής διακυβέρνησης που βασίζεται στη διαφάνεια, στη διαχείριση δεδομένων και στη λογοδοσία.

To White Paper for Data and Artificial Intelligence 2024–2030 θέτει ως βασικούς στόχους τη διακυβέρνηση βάσει δεδομένων, την αξιοποίηση της τεχνητής νοημοσύνης για τη βελτίωση των δημόσιων υπηρεσιών και την ανάπτυξη αξιόπιστης και ανθρωποκεντρικής AI. Η αξιοπιστία και η ανθρώπινη εποπτεία αποτελούν κεντρικές αρχές της στρατηγικής και όχι δευτερεύουσες παραμέτρους.

Παράλληλα, η Εσθονία έχει επενδύσει συστηματικά στην ανάπτυξη ψηφιακών υπηρεσιών που βασίζονται στην εμπιστοσύνη του πολίτη προς το κράτος. Η χρήση εργαλείων AI συνδέεται με σαφείς κανόνες διαχείρισης δεδομένων και με τη διατήρηση ανθρώπινης ευθύνης για τις διοικητικές αποφάσεις. Η χώρα συνεχίζει να αναπτύσσει το πρόγραμμα Bürokratt, ένα οικοσύστημα ψηφιακών βοηθών για την εξυπηρέτηση πολιτών, εντάσσοντας την τεχνητή νοημοσύνη στο πλαίσιο του ψηφιακού κράτους και όχι ως αυτόνομη τεχνολογική πρωτοβουλία.

Η εσθονική εμπειρία παρουσιάζει ιδιαίτερη αξία για τη συζήτηση περί ψηφιακής αλήθειας. Δείχνει ότι η εμπιστοσύνη δεν οικοδομείται εκ των υστέρων μέσω επικοινωνιακών παρεμβάσεων, αλλά ενσωματώνεται εξαρχής στον σχεδιασμό των ψηφιακών υπηρεσιών, των δεδομένων και των συστημάτων τεχνητής νοημοσύνης. Η αξιοπιστία προκύπτει ως αποτέλεσμα θεσμικής συνέπειας, διαφάνειας και σαφούς λογοδοσίας.

Ηνωμένο Βασίλειο: Η θεσμική χρήση της Generative AI στη δημόσια επικοινωνία

Το Ηνωμένο Βασίλειο αποτελεί μία από τις πρώτες χώρες που προχώρησαν στη διαμόρφωση εξειδικευμένων κατευθύνσεων για τη χρήση της Generative AI στη δημόσια διοίκηση και ειδικότερα στη δημόσια επικοινωνία. Η προσέγγισή του παρουσιάζει ιδιαίτερο ενδιαφέρον για την παρούσα μελέτη, καθώς συνδέει άμεσα την αξιοποίηση της τεχνητής νοημοσύνης με ζητήματα αξιοπιστίας, διαφάνειας και λογοδοσίας.

Το 2024 το Government Communication Service (GCS), ο κεντρικός επαγγελματικός φορέας επικοινωνίας της βρετανικής κυβέρνησης, δημοσίευσε επίσημη πολιτική για τη χρήση Generative AI στις κυβερνητικές επικοινωνίες. Η πολιτική αυτή αναγνωρίζει ότι τα εργαλεία τεχνητής νοημοσύνης μπορούν να ενισχύσουν την αποτελεσματικότητα της δημόσιας επικοινωνίας, αλλά ταυτόχρονα θέτει σαφή όρια στη χρήση τους. Βασική αρχή είναι ότι καμία επικοινωνία προς το κοινό δεν πρέπει να παράγεται ή να διανέμεται χωρίς ουσιαστική ανθρώπινη εποπτεία. Η ανθρώπινη ευθύνη παραμένει το τελικό σημείο ελέγχου τόσο για στατικό περιεχόμενο όσο και για διαδραστικές υπηρεσίες, όπως chatbots και ψηφιακοί βοηθοί.

Η πολιτική του GCS προβλέπει επίσης ότι οι πολίτες πρέπει να γνωρίζουν πότε αλληλοεπιδρούν με σύστημα τεχνητής νοημοσύνης αντί με άνθρωπο. Η απαίτηση αυτή αντιμετωπίζει τη διαφάνεια όχι ως τεχνική λεπτομέρεια αλλά ως προϋπόθεση διατήρησης της εμπιστοσύνης. Παράλληλα, απαγορεύεται η χρήση Generative AI με τρόπο που θα μπορούσε να παραβιάσει τις αρχές του Civil Service Code ή να

υπονομεύσει την ακρίβεια και την αμεροληψία της δημόσιας επικοινωνίας.

Σε ευρύτερο επίπεδο, το Cabinet Office και το Central Digital and Data Office δημοσίευσαν το Generative AI Framework for HM Government, ένα ολοκληρωμένο πλαίσιο που καθορίζει δέκα βασικές αρχές για την ασφαλή και υπεύθυνη χρήση της Generative AI στον δημόσιο τομέα. Μεταξύ αυτών περιλαμβάνονται η ανθρώπινη εποπτεία, η διαχείριση κινδύνων, η ασφάλεια δεδομένων, η διαφάνεια, η ανάπτυξη δεξιοτήτων και η ύπαρξη κατάλληλων μηχανισμών διακυβέρνησης. Το πλαίσιο επισημαίνει ότι τα συστήματα Generative AI δεν πρέπει να υποκαθιστούν τη στρατηγική κρίση ή τη λήψη σημαντικών διοικητικών αποφάσεων και ότι η τεχνολογία οφείλει να λειτουργεί υποστηρικτικά προς τους δημόσιους λειτουργούς.

Ιδιαίτερη σημασία έχει το γεγονός ότι το Ηνωμένο Βασίλειο δεν αντιμετωπίζει την τεχνητή νοημοσύνη αποκλειστικά ως ζήτημα παραγωγικότητας. Το πλαίσιο δίνει έμφαση στη νομιμότητα, στην ηθική χρήση, στην προστασία προσωπικών δεδομένων και στην αποφυγή παραπληροφόρησης. Η αξιοπιστία του δημόσιου λόγου θεωρείται αναγκαία προϋπόθεση για την επιτυχή ενσωμάτωση της AI στις δημόσιες υπηρεσίες.

Η περίπτωση του Ηνωμένου Βασιλείου αναδεικνύει ότι η αποτελεσματική αξιοποίηση της Generative AI στη δημόσια διοίκηση δεν βασίζεται μόνο στην τεχνολογική καινοτομία. Προϋποθέτει σαφείς κανόνες χρήσης, ανθρώπινη λογοδοσία και διαρκή ενημέρωση των πολιτών για τον τρόπο με τον οποίο χρησιμοποιούνται τα συστήματα τεχνητής νοημοσύνης. Υπό αυτή την έννοια, η βρετανική εμπειρία συνδέεται άμεσα με το κεντρικό ερώτημα της παρούσας μελέτης: πώς μπορεί να διατηρηθεί η εμπιστοσύνη και η θεσμική αξιοπιστία σε ένα περιβάλλον όπου η παραγωγή περιεχομένου δεν αποτελεί πλέον αποκλειστικά ανθρώπινη δραστηριότητα.

Συγκριτικά συμπεράσματα από τη διεθνή εμπειρία

Η εξέταση των περιπτώσεων της Γαλλίας, του Ηνωμένου Βασιλείου, της Σιγκαπούρης και της Εσθονίας αναδεικνύει ότι, παρά τις διαφορετικές διοικητικές παραδόσεις και τα διαφορετικά επίπεδα ψηφιακής ωριμότητας, οι δημόσιες διοικήσεις συγκλίνουν σε ορισμένες κοινές αρχές ως προς τη διαχείριση της τεχνητής νοημοσύνης και της ψηφιακής αξιοπιστίας.

Πρώτον, καμία από τις χώρες δεν αντιμετωπίζει την τεχνητή νοημοσύνη αποκλειστικά ως τεχνολογικό ζήτημα. Αντίθετα, η χρήση της εντάσσεται σε ευρύτερα πλαίσια διακυβέρνησης, δημόσιας επικοινωνίας και θεσμικής λογοδοσίας. Η τεχνολογία αντιμετωπίζεται ως εργαλείο υποστήριξης της δημόσιας αποστολής και όχι ως αυτοσκοπός.

Δεύτερον, η ανθρώπινη εποπτεία παραμένει κεντρική αρχή σε όλες τις προσεγγίσεις. Είτε πρόκειται για τις κατευθύνσεις του Government Communication Service στο Ηνωμένο Βασίλειο, είτε για τη στρατηγική της Σιγκαπούρης, είτε για το ψηφιακό μοντέλο της Εσθονίας, η τελική ευθύνη για το περιεχόμενο, τις αποφάσεις και την επικοινωνία εξακολουθεί να ανήκει στους δημόσιους λειτουργούς και όχι στα συστήματα τεχνητής νοημοσύνης.

Τρίτον, η διαφάνεια αναδεικνύεται σε βασική προϋπόθεση διατήρησης της εμπιστοσύνης. Οι διεθνείς πρακτικές συγκλίνουν στην ανάγκη οι πολίτες να γνωρίζουν πότε χρησιμοποιείται τεχνητή νοημοσύνη, με ποιον τρόπο και για ποιον σκοπό. Η διαφάνεια δεν παρουσιάζεται ως ζήτημα συμμόρφωσης αλλά ως όρος θεσμικής αξιοπιστίας.

Τέταρτον, η εμπιστοσύνη δεν θεωρείται δεδομένη. Αντίθετα, αντιμετωπίζεται ως αποτέλεσμα συνεχούς προσπάθειας που συνδέεται με τη σαφήνεια της δημόσιας επικοινωνίας, τη συνέπεια των θεσμών και την ικανότητα των διοικήσεων να εξηγούν τις επιλογές τους σε ένα περιβάλλον αυξημένης ψηφιακής πολυπλοκότητας.

Με διαφορετικούς τρόπους, όλες οι χώρες που εξετάστηκαν καταλήγουν στο ίδιο βασικό συμπέρασμα: η αξιοπιστία του δημόσιου τομέα στην

εποχή της Generative AI θα εξαρτηθεί λιγότερο από τις ίδιες τις τεχνολογίες και περισσότερο από την ικανότητα των θεσμών να διατηρήσουν την εμπιστοσύνη των πολιτών.

Από τη διεθνή εμπειρία στους τρεις πυλώνες διαχείρισης της ψηφιακής αλήθειας

Η ανάλυση των διεθνών παραδειγμάτων αναδεικνύει ότι, παρά τις διαφορετικές διοικητικές και θεσμικές προσεγγίσεις, οι δημόσιες διοικήσεις που ενσωματώνουν την τεχνητή νοημοσύνη συγκλίνουν σε ορισμένες κοινές αρχές. Η έμφαση στην ανάπτυξη δεξιοτήτων και στην κατανόηση των νέων τεχνολογιών, η ανάγκη διαφάνειας ως προς τον τρόπο χρήσης των συστημάτων AI και η διατήρηση της ανθρώπινης ευθύνης και λογοδοσίας εμφανίζονται επανειλημμένα στις στρατηγικές και τα κανονιστικά πλαίσια που εξετάστηκαν.

Η διαπίστωση αυτή είναι ιδιαίτερα σημαντική για τη δημόσια διοίκηση. Δείχνει ότι η διαχείριση της ψηφιακής αλήθειας δεν μπορεί να στηριχθεί σε μία μόνο παρέμβαση, ούτε να αντιμετωπιστεί αποκλειστικά ως τεχνολογικό ή επικοινωνιακό ζήτημα. Απαιτεί έναν συνδυασμό θεσμικών, οργανωσιακών και ανθρώπινων παραγόντων που λειτουργούν συμπληρωματικά.

Υπό το πρίσμα αυτό, η παρούσα μελέτη προτείνει τρεις αλληλένδετους πυλώνες για τη διαχείριση της ψηφιακής αλήθειας στη δημόσια διοίκηση. Οι πυλώνες αυτοί δεν αποτελούν τεχνικές οδηγίες ή εξαντλητικό πλαίσιο πολιτικής. Αποτυπώνουν τις βασικές αρχές που προκύπτουν από τη διεθνή εμπειρία και μπορούν να λειτουργήσουν ως σημεία αναφοράς για την προσαρμογή των δημόσιων οργανισμών στη νέα ψηφιακή συνθήκη.

Ο πρώτος πυλώνας αφορά την ψηφιακή παιδεία, δηλαδή την ανάπτυξη γνώσεων και δεξιοτήτων που επιτρέπουν την κριτική αξιολόγηση της πληροφορίας και την υπεύθυνη χρήση της τεχνητής νοημοσύνης. Ο δεύτερος αφορά τη διαφάνεια στη χρήση της Gen AI, ώστε οι πολίτες να γνωρίζουν πότε και με ποιον τρόπο χρησιμοποιούνται τέτοιες

τεχνολογίες. Ο τρίτος αφορά τη διατήρηση της ανθρώπινης κρίσης και της θεσμικής λογοδοσίας ως τελικής εγγύησης αξιοπιστίας.

Οι τρεις αυτοί πυλώνες αναλύονται στις επόμενες ενότητες.

Με βάση αυτή τη λογική, η ανάλυση που ακολουθεί δεν αντιμετωπίζει τους πυλώνες ως απομονωμένες δράσεις, αλλά ως αλληλένδετα πεδία θεσμικής ικανότητας: γνώση, διαφάνεια και υπεύθυνη ανθρώπινη κρίση.

6.2 Πρώτος πυλώνας: ψηφιακή παιδεία ως θεσμική πολιτική

Η ψηφιακή παιδεία δεν αποτελεί μόνο ατομική δεξιότητα. Αποτελεί πεδίο δημόσιας πολιτικής. Η UNESCO, στην καθοδήγησή της για τη Generative AI, τονίζει ότι οι κυβερνήσεις χρειάζεται να σχεδιάζουν ρυθμίσεις, πολιτικές και ανθρώπινες ικανότητες ώστε η Gen AI να υπηρετεί μια ανθρώπινη-κεντρική προσέγγιση.

Σύμφωνα με έρευνα ερευνητών του MIT Media Lab, αναδεικνύεται ο κίνδυνος της υπερβολικής εξάρτησης από συστήματα AI για την επαλήθευση πληροφοριών. Σύμφωνα με την έρευνα, συμμετέχοντες που βασίστηκαν συστηματικά σε AI εργαλεία για fact-checking εμφάνισαν στη συνέχεια μειωμένη ικανότητα να εντοπίζουν παραπληροφόρηση χωρίς τη βοήθειά τους. Το εύρημα αυτό ενισχύει την ανάγκη για μια νέα μορφή AI literacy: όχι μόνο εξοικείωση με τα εργαλεία, αλλά καλλιέργεια κριτικής σκέψης, ικανότητας επαλήθευσης και συνειδητής ανθρώπινης εποπτείας.

Για τη δημόσια διοίκηση, η ψηφιακή παιδεία αφορά την κατανόηση της προέλευσης της πληροφορίας, τη διάκριση μεταξύ θεσμικής πληροφόρησης και ερμηνείας, και την ανάπτυξη ρεαλιστικών προσδοκιών απέναντι στην Gen AI.

Η ψηφιακή παιδεία αποτελεί επομένως την προϋπόθεση ώστε τα στελέχη και οι πολίτες να μπορούν να αναγνωρίζουν τα όρια της τεχνολογίας. Ωστόσο, η γνώση από μόνη της δεν επαρκεί αν δεν συνοδεύεται από

σαφή ενημέρωση για το πότε και πώς χρησιμοποιείται η Gen AI από τους δημόσιους φορείς.

6.3 Δεύτερος πυλώνας: διαφάνεια στη χρήση της Gen AI

Η διαφάνεια αποτελεί κεντρικό άξονα διαχείρισης της ψηφιακής αλήθειας. Στην πράξη, δεν αφορά μόνο την τεχνική περιγραφή των συστημάτων, αλλά τη σαφή σχέση του πολίτη με την πληροφορία.

Η δημόσια διοίκηση οφείλει να καθιστά σαφές πότε χρησιμοποιείται Gen AI, για ποιο σκοπό, με ποια όρια και με ποια ανθρώπινη εποπτεία. Η διαφάνεια δεν αποδυναμώνει τον θεσμικό λόγο. Αντίθετα, λειτουργεί ως μηχανισμός αποκατάστασης ελέγχου και ενίσχυσης λογοδοσίας.

Η διαφάνεια, όμως, δεν έχει ουσιαστικό αποτέλεσμα αν δεν συνοδεύεται από σαφή ανάληψη ευθύνης. Για τον λόγο αυτό, ο τρίτος πυλώνας αφορά την ανθρώπινη κρίση ως τελικό σημείο ελέγχου και λογοδοσίας.

6.4 Τρίτος πυλώνας: ανθρώπινη κρίση και θεσμική λογοδοσία

Η διατήρηση ανθρώπινης κρίσης ως τελικού σημείου λήψης αποφάσεων αποτελεί κοινό παρονομαστή των διεθνών προσεγγίσεων για αξιόπιστη AI. Για τη δημόσια διοίκηση, η ανθρώπινη εποπτεία δεν είναι τεχνική επιλογή, αλλά θεσμική εγγύηση λογοδοσίας.

Η τεχνολογία μπορεί να υποστηρίζει, να επιταχύνει και να οργανώνει. Δεν μπορεί όμως να αντικαθιστά την ευθύνη του δημόσιου φορέα για το περιεχόμενο, τη διαδικασία και τις συνέπειες της επικοινωνίας.

Οι τρεις πυλώνες συγκροτούν ένα λειτουργικό πλαίσιο αρχών. Η κανονιστική τους σημασία ενισχύεται περαιτέρω από το ευρωπαϊκό πλαίσιο για την τεχνητή νοημοσύνη, το οποίο μετατρέπει πολλές από αυτές τις αρχές σε συγκεκριμένες υποχρεώσεις για τους δημόσιους φορείς.

6.5 Ο ρόλος του AI Act

Ο Ευρωπαϊκός Κανονισμός για την Τεχνητή Νοημοσύνη, Regulation (EU) 2024/1689, αποτελεί το πρώτο ολοκληρωμένο κανονιστικό πλαίσιο

της ΕΕ για την AI. Ο Κανονισμός θεσπίζει εναρμονισμένους κανόνες για την AI, με στόχο την προώθηση ανθρώπινης-κεντρικής και αξιόπιστης AI, την προστασία θεμελιωδών δικαιωμάτων, της δημοκρατίας και του κράτους δικαίου, καθώς και την υποστήριξη της καινοτομίας.

Για τη δημόσια διοίκηση, ο AI Act έχει ιδιαίτερη σημασία για τρεις λόγους. Πρώτον, καθιερώνει υποχρεώσεις διαφάνειας για συστήματα που αλληλεπιδρούν με φυσικά πρόσωπα ή παράγουν συνθετικό περιεχόμενο. Δεύτερον, προβλέπει ανθρώπινη εποπτεία για συστήματα υψηλού κινδύνου. Τρίτον, αναγνωρίζει ότι οι δημόσιες αρχές δεν είναι μόνο ρυθμιστές, αλλά και χρήστες συστημάτων AI.

Ο AI Act μετατρέπει την αυθεντικότητα, τη διαφάνεια και τη λογοδοσία από ηθικές αρχές σε κανονιστικές υποχρεώσεις.

6.5Α Τι αλλάζει πρακτικά με τον AI Act για τους δημόσιους φορείς

Ο AI Act δεν αποτελεί μόνο τεχνολογική ή νομική ρύθμιση. Δημιουργεί νέες υποχρεώσεις οργανωσιακής ετοιμότητας για τους δημόσιους φορείς.

Οι δημόσιες υπηρεσίες που χρησιμοποιούν ή προμηθεύονται συστήματα AI καλούνται να αναπτύξουν διαδικασίες:

- αξιολόγησης κινδύνου,
- διαφάνειας,
- τεκμηρίωσης,
- λογοδοσίας,
- και ανθρώπινης εποπτείας.

Η συμμόρφωση δεν περιορίζεται στην τεχνική λειτουργία των συστημάτων. Επεκτείνεται στον τρόπο με τον οποίο οι δημόσιοι οργανισμοί εξηγούν στους πολίτες πότε και πώς χρησιμοποιείται η AI.

Για τον λόγο αυτό, ο AI Act συνδέεται άμεσα με τη δημόσια επικοινωνία και όχι μόνο με την τεχνολογική διακυβέρνηση.

Η σημασία των υποχρεώσεων αυτών γίνεται πιο συγκεκριμένη όταν εξετάζεται πώς η AI έχει ήδη αρχίσει να ενσωματώνεται σε πραγματικές δημόσιες υπηρεσίες και διοικητικές λειτουργίες στην Ελλάδα.

6.6 Ελληνικά παραδείγματα θεσμικής χρήσης AI

Στην Ελλάδα έχουν ήδη αναπτυχθεί πρωτοβουλίες που δείχνουν την κατεύθυνση θεσμικής ενσωμάτωσης της AI στον δημόσιο τομέα.

Το mAiGreece, σύμφωνα με το gov.gr, είναι ψηφιακός βοηθός τεχνητής νοημοσύνης που δημιουργήθηκε από το Υπουργείο Τουρισμού και το Υπουργείο Ψηφιακής Διακυβέρνησης, με στόχο την παροχή πληροφοριών και άμεσης πρόσβασης σε κρίσιμη ενημέρωση για επισκέπτες της χώρας.

Η μελέτη Generative AI Greece 2030, από την Ειδική Γραμματεία Μακροπρόθεσμου Σχεδιασμού, το ΕΚΚΕ και το ΕΚΕΦΕ “Δημόκριτος”, παρουσιάζεται ως η πρώτη εμπειρική στρατηγική foresight προσέγγιση για τη χρήση της Gen AI στην Ελλάδα, με ανάλυση τάσεων, ευκαιριών, προκλήσεων και πιθανών επιλογών για το μέλλον του οικοσυστήματος Gen AI στη χώρα.

Το Hellenic Cadastre αποτελεί επίσης παράδειγμα χρήσης Azure OpenAI Service για επιτάχυνση της αξιολόγησης εγγράφων και ανάπτυξη ψηφιακού βοηθού για απαντήσεις σε ερωτήσεις πολιτών, σύμφωνα με το σχετικό Microsoft customer story.

Τέλος, η έκθεση European Commission AI Watch για την Ελλάδα αναφέρει ότι η χώρα αναπτύσσει την εθνική της στρατηγική για την AI, με συντονισμό από το Υπουργείο Ψηφιακής Διακυβέρνησης και με την AI ως έναν από τους βασικούς άξονες της Βίβλου Ψηφιακού Μετασχηματισμού.

Τα παραδείγματα αυτά δεν σημαίνουν ότι έχει λυθεί το ζήτημα της ψηφιακής αλήθειας. Δείχνουν όμως ότι η ελληνική δημόσια διοίκηση έχει ήδη αρχίσει να εισέρχεται σε πεδίο πρακτικής χρήσης AI, γεγονός

που καθιστά ακόμη πιο αναγκαίο ένα πλαίσιο διαφάνειας, ανθρώπινης εποπτείας και θεσμικής λογοδοσίας.

Από τα παραδείγματα αυτά προκύπτει ότι η διαχείριση της ψηφιακής αλήθειας δεν αφορά μόνο κεντρικές πολιτικές ή κανονιστικές επιλογές. Αφορά και τις καθημερινές δεξιότητες, κρίσεις και ευθύνες των στελεχών που σχεδιάζουν, εφαρμόζουν και επικοινωνούν δημόσιες υπηρεσίες.

6.7 Επιπτώσεις για τα στελέχη της δημόσιας διοίκησης

Η ενσωμάτωση της Generative AI στο περιβάλλον δημόσιας επικοινωνίας δεν επηρεάζει μόνο τις τεχνολογικές υποδομές ή τις διαδικασίες παραγωγής περιεχομένου. Επηρεάζει άμεσα και τον ρόλο των στελεχών της δημόσιας διοίκησης.

Τα στελέχη καλούνται πλέον:

- να αξιολογούν την αξιοπιστία ψηφιακού περιεχομένου,
- να κατανοούν βασικές αρχές λειτουργίας της Gen AI,
- να διαχειρίζονται περιπτώσεις πιθανής παραπληροφόρησης,
- να αναγνωρίζουν κινδύνους αλλοίωσης θεσμικού λόγου,
- και να λειτουργούν ως σημεία θεσμικής αποσαφήνισης.

Η ανάγκη αυτή δεν αφορά μόνο εξειδικευμένα στελέχη επικοινωνίας ή πληροφορικής. Αφορά συνολικά τη διοικητική λειτουργία, καθώς η διάκριση μεταξύ αυθεντικού και συνθετικού περιεχομένου συνδέεται πλέον με:

- την αξιοπιστία των δημόσιων φορέων,
- τη σχέση εμπιστοσύνης με τους πολίτες,
- και την αποτελεσματικότητα της δημόσιας πολιτικής.

Στο νέο αυτό περιβάλλον, η ανθρώπινη κρίση αποκτά ακόμη μεγαλύτερη σημασία. Η Gen AI μπορεί να υποστηρίζει την επεξεργασία πληροφορίας, δεν μπορεί όμως να αντικαθιστά:

- τη θεσμική ευθύνη,
- την πολιτική κρίση,
- και τη λογοδοσία της δημόσιας διοίκησης.

Η επιμόρφωση των στελεχών σε ζητήματα ψηφιακής αξιοπιστίας, επαλήθευσης περιεχομένου και θεσμικής επικοινωνίας αναδεικνύεται, επομένως, σε κρίσιμο πεδίο διοικητικής προσαρμογής για τα επόμενα χρόνια.

Κεφάλαιο 7

Συμπεράσματα και θεσμικές επιπτώσεις

7.1 Η ψηφιακή αλήθεια ως μόνιμη συνθήκη

Η ανάλυση που προηγήθηκε δείχνει ότι η αμφισβήτηση της αυθεντικότητας της πληροφορίας δεν αποτελεί παροδικό φαινόμενο. Η Gen AI επιταχύνει μια μετάβαση προς ένα ψηφιακό περιβάλλον όπου η αλήθεια δεν αμφισβητείται μόνο ως προς το περιεχόμενό της, αλλά και ως προς την προέλευσή της.

Η συνθήκη αυτή είναι μόνιμη. Δεν θα αντιστραφεί με τεχνικές λύσεις ή με αυστηρότερη επιμέλεια περιεχομένου. Για τη δημόσια διοίκηση, αυτό σημαίνει ότι η ψηφιακή αλήθεια πρέπει να αντιμετωπιστεί ως διαρκές θεσμικό πεδίο ευθύνης.

7.2 Ο μετασχηματισμός του ρόλου της δημόσιας διοίκησης

Η δημόσια διοίκηση δεν λειτουργεί πλέον σε περιβάλλον μονοσήμαντης αυθεντίας. Η θεσμική πληροφορία συνυπάρχει με δημοσιογραφικό περιεχόμενο, μη θεσμικούς φορείς επιρροής και AI-generated αφηγήσεις.

Ο ρόλος της διοίκησης μετασχηματίζεται από αποκλειστικό πομπό πληροφορίας σε θεσμικό σημείο αναφοράς. Η αξία της θεσμικής πληροφορίας δεν έγκειται μόνο στην ακρίβειά της, αλλά στη δυνατότητά της να μειώνει την αβεβαιότητα και να προσφέρει ερμηνευτικό πλαίσιο.

7.3 Εμπιστοσύνη: από δεδομένο σε δυναμική διαδικασία

Η εμπιστοσύνη στους θεσμούς δεν μπορεί πλέον να θεωρείται σταθερό κεφάλαιο. Λειτουργεί ως δυναμική διαδικασία που επηρεάζεται από τον τρόπο επικοινωνίας, τη διαφάνεια στη χρήση τεχνολογίας και τη διατήρηση ανθρώπινης κρίσης.

Η εμπιστοσύνη δεν διασφαλίζεται με περισσότερες πληροφορίες, αλλά με σαφήνεια, συνέπεια και αναγνωρίσιμη θεσμική παρουσία.

7.4 Θεσμικές επιλογές αντί τεχνικών αντανακλαστικών

Η μελέτη δείχνει ότι οι αποτελεσματικές απαντήσεις στην πρόκληση της ψηφιακής αλήθειας δεν είναι μόνο τεχνικές. Είναι θεσμικές.

Η υπερβολική έμφαση σε ανιχνευτές, φίλτρα και αυτοματισμούς κινδυνεύει να μετατοπίσει την ευθύνη από τον θεσμό στην τεχνολογία. Αντίθετα, οι κρίσιμες επιλογές είναι η επένδυση στην ψηφιακή παιδεία, η διαφάνεια στη χρήση Gen AI και η διατήρηση της ανθρώπινης κρίσης ως τελικού κριτηρίου.

7.5 Η δημόσια διοίκηση σε οικοσύστημα επιρροής

Η επιρροή στον δημόσιο διάλογο είναι πλέον κατανεμημένη. Η δημόσια διοίκηση δεν μπορεί και δεν πρέπει να ανταγωνιστεί influencers στο πεδίο της ορατότητας ή της συναισθηματικής απήχησης.

Αντίθετα, οφείλει να αναγνωρίσει το οικοσύστημα επιρροής, να διασφαλίσει αξιόπιστη θεσμική πληροφορία και να λειτουργήσει ως σημείο αναφοράς για επαλήθευση και κατανόηση.

7.6 Η ψηφιακή αλήθεια ως δημόσιο αγαθό

Η ψηφιακή αλήθεια πρέπει να αντιμετωπιστεί ως δημόσιο αγαθό. Όπως η διαφάνεια, η λογοδοσία και η πρόσβαση στην πληροφορία, έτσι και η διαχείριση της αυθεντικότητας του δημόσιου λόγου αποτελεί θεμελιώδη λειτουργία της δημοκρατικής διοίκησης.

Η δημόσια διοίκηση καλείται να προνοεί, να ενσωματώνει νέες τεχνολογίες με όρους θεσμικής ευθύνης και να αξιολογεί διαρκώς τον αντίκτυπο των επιλογών της.

7.7 Τελική αποτίμηση

Η ανάπτυξη της Generative AI σηματοδοτεί μια νέα φάση στην εξέλιξη του ψηφιακού περιβάλλοντος. Σε αντίθεση με προηγούμενες τεχνολογικές μεταβολές, η τεχνητή νοημοσύνη δεν επηρεάζει μόνο τον τρόπο διάδοσης της πληροφορίας αλλά και την ίδια την παραγωγή της. Η δυνατότητα δημιουργίας κειμένων, εικόνων, βίντεο και ηχητικού

περιεχομένου με υψηλό βαθμό πειστικότητας μεταβάλλει θεμελιωδώς τις συνθήκες μέσα στις οποίες οι πολίτες ενημερώνονται, αλληλεπιδρούν και διαμορφώνουν αντιλήψεις για την πραγματικότητα.

Η εξέλιξη αυτή δημιουργεί σημαντικές ευκαιρίες για τη δημόσια διοίκηση. Η τεχνητή νοημοσύνη μπορεί να συμβάλει στη βελτίωση της εξυπηρέτησης πολιτών, στην αποτελεσματικότερη διαχείριση πληροφοριών, στην ενίσχυση της προσβασιμότητας και στην αναβάθμιση της δημόσιας επικοινωνίας. Παράλληλα, όμως, δημιουργεί νέες προκλήσεις που συνδέονται με την αξιοπιστία της πληροφορίας, τη διαφάνεια των διαδικασιών και τη διατήρηση της εμπιστοσύνης στους δημόσιους θεσμούς.

Η διεθνής εμπειρία που εξετάστηκε στη μελέτη δείχνει ότι οι πιο ώριμες διοικητικές προσεγγίσεις δεν αντιμετωπίζουν την τεχνητή νοημοσύνη αποκλειστικά ως τεχνολογικό εργαλείο. Αντίθετα, τη συνδέουν με ζητήματα διακυβέρνησης, λογοδοσίας, δημόσιας επικοινωνίας και θεσμικής αξιοπιστίας. Η Γαλλία δίνει έμφαση στην αξιολόγηση της δημόσιας επικοινωνίας και στην οικοδόμηση εμπιστοσύνης. Το Ηνωμένο Βασίλειο αναπτύσσει σαφείς κανόνες για την υπεύθυνη χρήση της Generative AI στη δημόσια επικοινωνία. Η Σιγκαπούρη προσεγγίζει την τεχνητή νοημοσύνη ως μέσο παραγωγής δημόσιας αξίας, ενώ η Εσθονία εντάσσει την αξιοποίησή της σε ένα ευρύτερο πλαίσιο ψηφιακής διακυβέρνησης και θεσμικής διαφάνειας. Παρά τις επιμέρους διαφοροποιήσεις, όλες οι προσεγγίσεις συγκλίνουν στην ανάγκη διατήρησης της ανθρώπινης ευθύνης και της εμπιστοσύνης των πολιτών.

Η πρόκληση της ψηφιακής αλήθειας δεν περιορίζεται στην αντιμετώπιση ψευδών πληροφοριών ή deepfakes. Η ευρύτερη πρόκληση αφορά τη δυνατότητα αναγνώρισης του αυθεντικού σε ένα περιβάλλον όπου η παραγωγή συνθετικού περιεχομένου καθίσταται ολοένα ευκολότερη και μαζικότερη. Η συζήτηση μετατοπίζεται σταδιακά από την παραπληροφόρηση προς την αυθεντικότητα, από την ακρίβεια της πληροφορίας προς την αξιοπιστία της πηγής και από τον έλεγχο του περιεχομένου προς την ενίσχυση της θεσμικής εμπιστοσύνης.

Υπό το πρίσμα αυτό, η ψηφιακή αλήθεια δεν μπορεί να διασφαλιστεί αποκλειστικά μέσω τεχνικών λύσεων ή κανονιστικών παρεμβάσεων. Απαιτεί συνδυασμό ψηφιακής παιδείας, διαφάνειας, υπεύθυνης χρήσης της τεχνητής νοημοσύνης και διατήρησης της ανθρώπινης κρίσης ως τελικού σημείου λογοδοσίας. Οι τρεις πυλώνες που αναπτύχθηκαν στην παρούσα μελέτη αναδεικνύουν ακριβώς αυτή την ανάγκη μιας πολυδιάστατης προσέγγισης.

Τελικά, το κρίσιμο ερώτημα για τη δημόσια διοίκηση δεν είναι αν θα χρησιμοποιήσει την τεχνητή νοημοσύνη. Η ενσωμάτωσή της αποτελεί ήδη πραγματικότητα. Το ουσιαστικό ερώτημα είναι αν θα μπορέσει να αξιοποιήσει τις δυνατότητές της διατηρώντας παράλληλα την εμπιστοσύνη των πολιτών. Στην εποχή της Generative AI, η αξιοπιστία των θεσμών αναδεικνύεται σε στρατηγικό πόρο. Όσο δυσκολότερη γίνεται η διάκριση μεταξύ αυθεντικού και συνθετικού περιεχομένου, τόσο μεγαλύτερη σημασία αποκτά η ύπαρξη θεσμών που λειτουργούν ως σταθερά και αξιόπιστα σημεία αναφοράς για την κοινωνία.

Η επόμενη δεκαετία δεν θα κριθεί μόνο από την ικανότητα των δημόσιων οργανισμών να υιοθετήσουν νέες τεχνολογίες. Θα κριθεί κυρίως από την ικανότητά τους να διατηρήσουν την εμπιστοσύνη, τη διαφάνεια και τη λογοδοσία σε ένα περιβάλλον όπου η πληροφορία θα είναι περισσότερο άφθονη από ποτέ, αλλά η βεβαιότητα για την προέλευσή της ενδεχομένως πιο περιορισμένη από οποιαδήποτε άλλη στιγμή στο παρελθόν.

Η πρόκληση της ψηφιακής αλήθειας δεν μπορεί να αντιμετωπιστεί με όρους άμυνας ή αποτροπής. Απαιτεί θεσμική ωρίμανση, στρατηγικό σχεδιασμό και αναγνώριση ότι η Gen AI αποτελεί μέρος της νέας κανονικότητας.

Υπό το πρίσμα αυτό, η επιμόρφωση των στελεχών της δημόσιας διοίκησης σε ζητήματα ψηφιακής αξιοπιστίας, αναγνώρισης συνθετικού περιεχομένου, θεσμικής επικοινωνίας και υπεύθυνης χρήσης της Gen AI αναδεικνύεται σε κρίσιμο πεδίο δράσης για το Εθνικό Κέντρο Δημόσιας Διοίκησης και Αυτοδιοίκησης τα επόμενα χρόνια.

Η δημόσια διοίκηση διαθέτει εργαλεία, κανονιστικό πλαίσιο και παραδείγματα για να ανταποκριθεί. Το κρίσιμο ζητούμενο είναι η συνέπεια, η πρόθεση και η διατήρηση του ανθρώπινου παράγοντα στο κέντρο της ψηφιακής διακυβέρνησης.

Παράρτημα – Βιβλιογραφία και πηγές

A.1 Deepfakes, synthetic media και ψηφιακή αλήθεια

Sartor, G. et al. / Royal Society Open Science (2023). *Deepfake detection with and without content warnings*. Διαθέσιμο στο: https://royalsocietypublishing.org/doi/10.1098/rsos.231214

Mai, K.T., Bray, S.D. and Davies, T. (2023). "Warning: Humans cannot reliably detect speech deepfakes". *PLOS ONE*, 18(8), e0285333. Διαθέσιμο στο: https://doi.org/10.1371/journal.pone.0285333

A.2 Παραπληροφόρηση, εμπιστοσύνη και θεσμική αξιοπιστία

Wardle, C. and Derakhshan, H. (2017). *Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making*. Strasbourg: Council of Europe. Διαθέσιμο στο: https://rm.coe.int/information-disorder-report-version-2018/16808c9c77

World Economic Forum (2024). *Global Risks Report 2024*. Geneva: World Economic Forum. Διαθέσιμο στο: https://www.weforum.org/reports/global-risks-report-2024/

Chesney, R. and Citron, D. (2019). "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security". *California Law Review*, 107(6), pp. 1753–1819. Διαθέσιμο στο: https://www.californialawreview.org/print/deep-fakes-a-looming-challenge-for-privacy-democracy-and-national-security/

Pennycook, G. et al. (2021). "Shifting attention to accuracy can reduce misinformation online". *Nature*, 592, pp. 590–595. Διαθέσιμο στο: https://www.nature.com/articles/s41586-021-03344-2

Kahneman, D. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

Tversky, A. and Kahneman, D. (1974). "Judgment under Uncertainty: Heuristics and Biases". *Science*, 185(4157), pp. 1124–1131. Διαθέσιμο στο: https://www.science.org/doi/10.1126/science.185.4157.1124

OECD (2024). *OECD Survey on Drivers of Trust in Public Institutions – 2024 Results: Building Trust in a Complex Policy Environment*. OECD Publishing. Διαθέσιμο στο: https://www.oecd.org/en/publications/oecd-survey-on-drivers-of-trust-in-public-institutions-2024-results_9a20554b-en.html

Α.3 Πλατφόρμες, influencers και δημόσιος λόγος

Abidin, C. (2018). *Internet Celebrity: Understanding Fame Online*. Bingley: Emerald Publishing. Διαθέσιμο στο: https://www.emerald.com/books/monograph/13023/Internet-Celebrity-Understanding-Fame-Online

Reuters Institute for the Study of Journalism (2024). *Digital News Report 2024*. Oxford: University of Oxford. Διαθέσιμο στο: https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2024

MIT News / MIT Media Lab (2026). “The consequences of relying on AI for accurate news”. Διαθέσιμο στο: https://news.mit.edu/2026/consequences-of-relying-on-ai-for-accurate-news-0609

Freberg, K., Graham, K., McGaughey, K. and Freberg, L. (2011). “Who are the Social Media Influencers?”. *Public Relations Review*, 37(1), pp. 90–92. Διαθέσιμο στο: https://doi.org/10.1016/j.pubrev.2010.11.001

Α.4 Κανονιστικό πλαίσιο, δεοντολογία και πολιτικές για Gen AI

European Union (2024). *Regulation (EU) 2024/1689 – Artificial Intelligence Act*. Official Journal of the European Union. Διαθέσιμο στο: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

UNESCO (2023). *Guidance for Generative AI in Education and Research*. Paris: UNESCO. Διαθέσιμο στο: https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research

Floridi, L. (2023). *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities*. Oxford: Oxford University Press. Διαθέσιμο στο: https://academic.oup.com/book/46748

Stanford Institute for Human-Centered Artificial Intelligence (2025). *AI Index Report 2025*. Stanford University. Διαθέσιμο στο: https://hai.stanford.edu/ai-index/2025-ai-index-report

Stanford Institute for Human-Centered Artificial Intelligence (2026). *AI Index Report 2026*. Stanford University. Διαθέσιμο στο: https://hai.stanford.edu/ai-index/2026-ai-index-report

Α.5 Παραδείγματα εφαρμογής AI και Gen AI στον δημόσιο τομέα – Ελλάδα

European Commission AI Watch (2021). *Greece AI Strategy Report*. Brussels: European Commission. Διαθέσιμο στο: https://ai-watch.ec.europa.eu/countries/greece/greece-ai-strategy-report_en

Special Secretariat of Foresight (2024). *Generative AI Greece 2030: Possible Futures of Generative Artificial Intelligence in Greece*. Athens: Presidency of the Government. Διαθέσιμο στο: https://foresight.gov.gr/en/studies/Generative-AI-Greece-2030-Possible-Futures-of-Generative-Artificial-Intelligence-in-Greece/

Hellenic Republic (2024). *mAiGreece: Digital assistant for visitors*. Διαθέσιμο στο: https://www.gov.gr/en/ipiresies/politismos-athletismos-kai-tourismos/tourismos/maigreece

Microsoft (2024). *From months to days: Hellenic Cadastre speeds up land registry assessments with Azure OpenAI Service*. Διαθέσιμο στο: https://www.microsoft.com/en/customers/story/20308-hellenic-cadastre-azure-open-ai-service

Α.6 Δημόσια διοίκηση και Gen AI – διεθνείς πρακτικές

OECD (2024). *Public Communication Scan of France*. OECD Publishing. Διαθέσιμο στο: https://www.oecd.org/en/publications/public-communication-scan-of-france_7450361a-en.html

UK Government Communication Service (2025). *Generative AI policy*. Διαθέσιμο στο: https://www.communications.gov.uk/publications/gcs-generative-ai-policy/

UK Government Digital Service (2025). *Artificial Intelligence Playbook for the UK Government*. Διαθέσιμο στο: https://www.gov.uk/government/publications/ai-playbook-for-the-uk-government/artificial-intelligence-playbook-for-the-uk-government-html

Singapore Government / Smart Nation (2024). *National AI Strategy 2.0*. Διαθέσιμο στο: https://www.smartnation.gov.sg/nais/