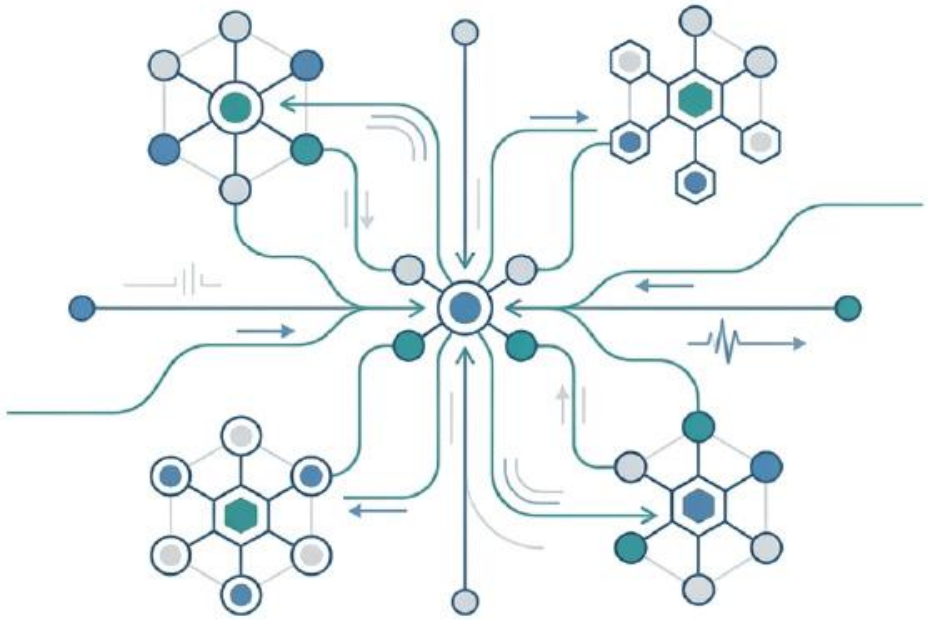
ΑΡΧΗ ΠΡΟΣΤΑΣΙΑΣ ΔΕΔΟΜΕΝΩΝ



Επιστημονική Ημερίδα της Αρχής:
2η Ημέρα Διαλόγου με την Ερευνητική Κοινότητα

22 Οκτωβρίου 2026

Διαχείριση περιεχομένου στα μέσα κοινωνικής δικτύωσης

Τεχνητή Νοημοσύνη (ΤΝ) και προστασία της έμφυλης ισότητας

Μαρία Αγγελίδου, Καθ. Στέφανος Γκρίτζαλης / Οκτώβριος 2026

Θεματικές Ενότητες

- **Έμφυλη βία στον κυβερνοχώρο:** Μορφές και επιπτώσεις
- **Αυτοματοποιημένη διαχείριση περιεχομένου:** Λειτουργία και προκλήσεις
- **Αλγοριθμική και έμφυλη μεροληψία:** Κίνδυνοι και επιπτώσεις στα θεμελιώδη δικαιώματα
- **Ευρωπαϊκό κανονιστικό πλαίσιο:** DSA, AI Act, GDPR και Οδηγία (ΕΕ) 2024/1385
- **Πρακτική συμμόρφωση:** Διαφάνεια, λογοδοσία και ανθρώπινη εποπτεία
- **Συμπεράσματα:** Τεχνολογική καινοτομία και προστασία της έμφυλης ισότητας

Έμφυλη Βία στον Κυβερνοχώρο

Η αυτοματοποιημένη διαχείριση περιεχομένου στις πλατφόρμες κοινωνικής δικτύωσης όταν επιφέρει ως αποτέλεσμα τον περιορισμό νόμιμου περιεχομένου εγείρει ζήτημα έμμεσης έμφυλης διάκρισης

Ενδεικτικές μορφές:

- Παρενόχληση στο διαδίκτυο (online harassment).
- Συντονισμένες επιθέσεις, μισογυνικός λόγος και έμφυλη παραπληροφόρηση (online misogynistic speech, manosphere, gender disinformation).
- Μη συναινετική κοινοποίηση προσωπικών ή σεξουαλικών απεικονίσεων σε υλικό παραγόμενο από ΤΝ (NCIDs).
- Παρενοχλητική κυβερνοπαρακολούθηση(cyberstalking).

→ Η βλάβη δεν περιορίζεται στην προσβολή της προσωπικότητας αλλά δύναται να επηρεάσει την ιδιωτική ζωή, την ψυχική ακεραιότητα, τη φήμη, την επαγγελματική δραστηριότητα και τη συμμετοχή στον δημόσιο βίο.

Ασφάλεια και ελευθερία της έκφρασης

Μπορεί ένα αυτοματοποιημένο σύστημα να παρέχει αποτελεσματική προστασία χωρίς να συμμετέχει στην αναπαραγωγή των έμφυλων διακρίσεων;

- Η εκθετική αύξηση τού περιεχομένου στα μέσα κοινωνικής δικτύωσης καθιστά αναγκαία τη χρήση αυτοματοποιημένων εργαλείων διαχείρισης περιεχομένου με στόχο την αποτελεσματικότερη προστασία των χρηστών και των χρηστριών.
- Ο μισογυνιστικός λόγος στο διαδίκτυο, οι απειλές, η διαδικτυακή παρενόχληση και η μη συναινετική διάδοση υλικού σεξουαλικών ή οικείων απεικονίσεων μπορούν να αποθαρρύνουν τη συμμετοχή των γυναικών στον δημόσιο διάλογο.
- Ωστόσο, ο εσφαλμένος περιορισμός περιεχομένου δύναται να περιθωριοποιήσει ακόμη περισσότερο τις ευάλωτες κοινωνικές ομάδες, εμποδίζοντας την ελευθερίας της έκφρασης, ιδίως γυναικών δημοσιογράφων, πολιτικών και ακτιβιστριών.

Πώς λειτουργεί η διαχείριση περιεχομένου;

Η ανίχνευση μιας λέξης δεν ισοδυναμεί με νομική αξιολόγηση. Ο ειρωνικός σχολιασμός, η καταγγελία ή/και η αναπαραγωγή συμβάντων κακοποίησης και η δημοσιογραφική τεκμηρίωση μπορεί να αλλάζουν πλήρως το νόημα μιας δημοσίευσης.

- **Ανίχνευση:** Ταξινόμηση κειμένου, εικόνων, βίντεο και ήχου.
- **Αξιολόγηση κινδύνου:** Εκτίμηση πιθανότητας παραβίασης κανόνων ή νομοθεσίας.
- **Αυτοματοποιημένη ενέργεια:** Επισήμανση, υποβάθμιση προβολής, περιορισμός ή αφαίρεση περιεχομένου.
- **Ανθρώπινη επανεξέταση:** Αξιολόγηση σύνθετων περιπτώσεων, προσφυγών και αμφισβητούμενων αποφάσεων.

→ Η ΤΝ μπορεί να αυξάνει την ταχύτητα και την κλίμακα της εποπτείας, αλλά δεν εγγυάται πάντα την ουδετερότητα ή την ορθή ερμηνευτική αντίληψη του συμφραζόμενου. Η εμφάνιση της αλγοριθμικής μεροληψίας απαιτεί να εξετάζεται κατά πόσον οι αποφάσεις διαχείρισης περιεχομένου επηρεάζουν την ισότιμη απόλαυση των θεμελιωδών δικαιωμάτων.

Τί είναι η έμφυλη μεροληψία των αλγορίθμων;

Τα αποτελέσματα της ΤΝ μπορεί να αναπαράγουν τα έμφυλα στερεότυπα ή να προκαλούν έμφυλες διακρίσεις. Η διαπίστωση νομικά απαγορευμένης διάκρισης απαιτεί ειδική αξιολόγηση.

Πιθανές αιτίες μεροληψίας της ΤΝ:

- Ανεπαρκής ή μη αντιπροσωπευτική εκπαίδευση των μοντέλων – μεροληπτικά δεδομένα εκπαίδευσης.
- Διαφορετικά ποσοστά σφαλμάτων ανά γλώσσα, διάλεκτο ή κοινωνικό πλαίσιο.
- Εσφαλμένη επισήμανση καταγγελιών, μαρτυριών ή εκπαιδευτικού περιεχομένου ως επιβλαβών.
- Ανεπαρκής αναγνώριση νέων μορφών συνθετικού περιεχομένου.

→ *Η αλγοριθμική μεροληψία μπορεί να οδηγήσει είτε σε ελλιπή αντιμετώπιση της έμφυλης βίας είτε σε εσφαλμένο περιορισμό νόμιμου περιεχομένου, με συνέπειες για την ισότητα και την ελευθερία της έκφρασης.*

Κανονισμός (ΕΕ) 2022/2065 Digital Services Act

Η Πράξη DSA θεσπίζει ενιαίο πλαίσιο υποχρεώσεων για τους παρόχους ψηφιακών υπηρεσιών και ενισχυμένες υποχρεώσεις για τις πολύ μεγάλες επιγραμμικές πλατφόρμες και μηχανές αναζήτησης.

Βασικές διατάξεις:

- **Άρθρο 14:** Κανόνες διαχείρισης περιεχομένου – διαφάνεια όρων και προϋποθέσεων χρήσης.
- **Άρθρα 16–17:** Μηχανισμός ειδοποίησης παράνομου περιεχομένου και αιτιολόγηση ορισμένων αποφάσεων περιορισμού.
- **Άρθρα 20-21:** Υποβολή καταγγελιών και Εξωδικαστική Επίλυση.
- **Άρθρα 34–35:** Εκτίμηση κινδύνων για VLOPs/VLOSEs – ένταξη της έμφυλης βίας στους συστημικούς κινδύνους.
- **Άρθρο 27:** Διαφάνεια των βασικών παραμέτρων των συστημάτων συστάσεων.

→ *Η Πράξη DSA δεν καθιστά κάθε επιβλαβές περιεχόμενο παράνομο ούτε επιβάλλει γενική παρακολούθηση. Η παρανομία απορρέει από το εφαρμοστέο δίκαιο (άρθρο 3 περ. η και 8).*

Κανονισμός (ΕΕ) 2024/1689 AI Act

Η Πράξη AIA θεσπίζει εναρμονισμένους κανόνες για τη ρύθμιση της ΤΝ, με μια κινδυνοκεντρική προσέγγιση και με σκοπό την προστασία των θεμελιωδών ανθρωπίνων δικαιωμάτων.

Βασικές διατάξεις:

- **Άρθρο 5:** Απαγορευμένες πρακτικές ΤΝ – Απαγόρευση nudifiers και NCIDs, όπως τροποποιήθηκε από τον Κανονισμό (ΕΕ) 2026/1744 (από 02.12.2026).
- **Άρθρα 6, 9–15 και παραρτήματα I/III:** Χαρακτηρισμός συστημάτων ΤΝ ως υψηλού κινδύνου και πλαίσιο υποχρεώσεων.
- **Άρθρο 50:** Υποχρεώσεις διαφάνειας για ορισμένα συστήματα ΤΝ και περιεχόμενο που παράγεται ή χειραγωγείται με ΤΝ, συμπεριλαμβανομένων των deepfakes.
- **Κανονισμός (ΕΕ) 2026/1744 Digital Omnibus για την ΤΝ νέο άρθρο 4α:** Κατ' εξαίρεση επεξεργασία ειδικών κατηγοριών δεδομένων για ανίχνευση/διόρθωση μεροληψίας, εφόσον είναι απολύτως αναγκαία και με αυστηρές εγγυήσεις. Δεν παρέχει γενική άδεια επεξεργασίας.

→ Η χρήση ΤΝ για διαχείριση περιεχομένου δεν συνεπάγεται αυτομάτως χαρακτηρισμό «υψηλού κινδύνου». Απαιτείται έλεγχος του προβλεπόμενου σκοπού και των κριτηρίων του άρθρου 6.

Το ευρύτερο νομικό πλαίσιο προστασίας των γυναικών

Οι Κανονισμοί DSA, AI Act όπως τροποποιήθηκε με το Digital Omnibus, GDPR, η Οδηγία 2024/1385 για την καταπολέμηση της έμφυλης βίας και ο ΧΘΔ λειτουργούν συμπληρωματικά, αλλά δεν έχουν ταυτόσημο πεδίο εφαρμογής.

- Η αξιολόγηση δεν περιορίζεται στους δύο Κανονισμούς.
- **Χάρτης Θεμελιωδών Δικαιωμάτων της ΕΕ:** Ανθρώπινη αξιοπρέπεια (άρθρο 1), ιδιωτική ζωή (άρθρο 7), προστασία δεδομένων (άρθρο 8), ελευθερία έκφρασης (άρθρο 11), ισότητα ενώπιον του νόμου (άρθρο 20), απαγόρευση διακρίσεων (άρθρο 21) και ισότητα γυναικών και ανδρών (άρθρο 23).
- **Κανονισμός (ΕΕ) 2016/679 (GDPR):** Όταν η διαχείριση περιεχομένου με εργαλεία περιλαμβάνει επεξεργασία δεδομένων προσωπικού χαρακτήρα. Εξετάζονται, μεταξύ άλλων, τα άρθρα 5, 6, 9 και, κατά περίπτωση, 22.
- **Οδηγία (ΕΕ) 2024/1385:** Καταπολέμηση της έμφυλης βίας με ποινικοποίηση συγκεκριμένων μορφών έμφυλης βίας υποστηριζόμενης με τεχνολογικά μέσα, όπως μη συναινετική κοινοποίηση οικείου/παραποιημένου υλικού (άρθρο 5). Προθεσμία μεταφοράς: 14.6.2027.

Τί σημαίνει πρακτικά η συμμόρφωση;

Δυσανάλογες επιπτώσεις στις γυναίκες αποτελούν ένδειξη προς διερεύνηση, όχι αυτομάτως απόδειξη παράνομης διάκρισης. Απαιτούνται στοιχεία, νομική βάση και έλεγχος αναλογικότητας.

- Χαρτογράφηση απειλών και συγκεκριμένων μορφών έμφυλης βίας.
- Έλεγχος ποιότητας δεδομένων, γλωσσικής κάλυψης και αποκλίσεων απόδοσης.
- Τεκμηρίωση των κανόνων περιορισμού και των διαδρομών ανθρώπινης επανεξέτασης και προσφυγής, όπου απαιτούνται.
- Εφαρμογή αναλογικών μέτρων και τεκμηρίωση των επιλογών.
- Παρακολούθηση αποτελεσματικότητας και επανεξέταση μετά από περιστατικά.
- Προστασία προσωπικών δεδομένων και περιορισμός πρόσβασης σε ευαίσθητο υλικό.

Συμπεράσματα

Η τεχνολογική αποτελεσματικότητα δεν αρκεί χωρίς νομική λογοδοσία και ισότιμη προστασία των θεμελιωδών δικαιωμάτων.

- Η ΤΝ μπορεί να ενισχύσει την προστασία της έμφυλης ισότητας, μόνο εφόσον αξιολογείται και ελέγχεται.
- Η προστασία των γυναικών απαιτεί αντιμετώπιση τόσο του παράνομου περιεχομένου όσο και των ευρύτερων συστημικών κινδύνων διαιώνισης και περαιτέρω καλλιέργιας διακρίσεων με ταυτόχρονη εξασφάλιση της συμμετοχής τους στον δημόσιο διάλογο.
- Η Πράξη DSA εστιάζει στις υποχρεώσεις των παρόχων ψηφιακών υπηρεσιών και στη διαχείριση κινδύνων, ενώ η Πράξη AI Act θεσπίζει κανόνες για την ΤΝ, με ειδικές απαιτήσεις διαφάνειας και συναφείς υποχρεώσεις ανάλογα με την ταξινόμηση των συστημάτων βάσει κινδύνου.
- Η διαφάνεια, η λογοδοσία, η αποτελεσματική αμφισβήτηση αποφάσεων και η ανθρώπινη εποπτεία συνιστούν κρίσιμες εγγυήσεις.

ΑΡΧΗ ΠΡΟΣΤΑΣΙΑΣ ΔΕΔΟΜΕΝΩΝ

Σας ευχαριστούμε για την προσοχή σας!

