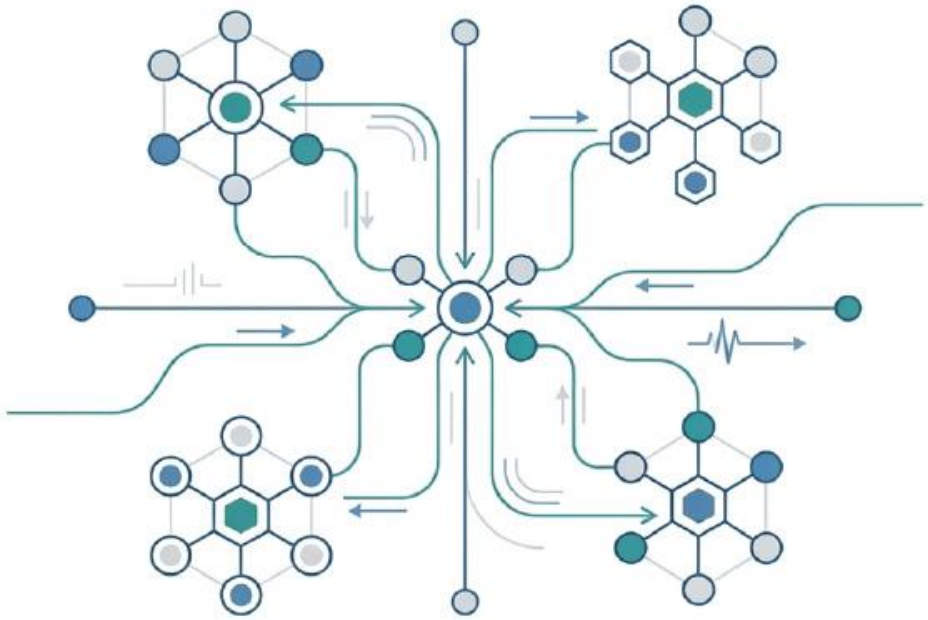
ΑΡΧΗ ΠΡΟΣΤΑΣΙΑΣ ΔΕΔΟΜΕΝΩΝ



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

22 Οκτωβρίου 2026

Τεχνητή Νοημοσύνη και Εντοπισμός Υλικού Παιδικής Σεξουαλικής Κακοποίησης στο Διαδίκτυο
Τεχνικές, κανονιστικό πλαίσιο και ζητήματα προστασίας δεδομένων

Εμμανουέλα Κοκολάκη

FORTH-ICS | Ερευνήτρια PreventCSA@EU

Τεχνολογίες Τεχνητής Νοημοσύνης για τον εντοπισμό προηγουμένως μη εντοπισμένου ΣεΚαΕΠ-π.

Ανάλυση εικόνας και βίντεο

Υπολογιστική όραση, εντοπισμός χαρακτηριστικών, ανάλυση περιεχομένου σε πραγματικό χρόνο.

Εργαλεία: Google Cloud Vision AI, Amazon Rekognition, Salus

Ταξινόμηση περιεχομένου και σημασιολογική επισήμανση

Κατηγοριοποίηση, σημασιολογικές ετικέτες και πιθανοτικές εκτιμήσεις.

Εργαλεία: Thorn Safer Predict, Google Content Safety API, HIVE AI, Sightengine

Ανάλυση συμπεριφοράς χρηστών (AI-based Behavioral Analysis)

Ανάλυση μοτίβων δραστηριότητας, αλληλεπιδράσεων και μεταδεδομένων για τον εντοπισμό ενδείξεων ΣεΚαΕΠ-π.

Τεχνικές: ML-based Anomaly Detection, Intent Classification

Εξέλιξη του ενωσιακού πλαισίου για την ανίχνευση ΣεΚαΕΠ-π.* από ιδιωτικούς παρόχους

- **Οδηγία 2000/31/ΕΚ (e-Commerce Directive) → Κανονισμός (ΕΕ) 2022/2065 (DSA)**

E-Commerce Directive: άρθρο 14 παρ. 1(β). Μεταφορά ευθύνης ενδιάμεσων παρόχων στα άρθρα 4-6 της DSA.

Άρθρο 8 DSA: απουσία γενικής υποχρέωσης παρακολούθησης και ενεργητικής αναζήτησης παράνομου περιεχομένου.

Άρθρο 7 DSA: δυνατότητα διενέργειας εθελοντικών ερευνών για τον εντοπισμό παράνομου περιεχομένου.

- **Κανονισμός (ΕΕ) 2021/1232 — Προσωρινή παρέκκλιση από την Οδηγία 2002/58/ΕΚ (έληξε 3 Απριλίου 2026)**

Δυνατότητα εθελοντικού εντοπισμού ΣεΚαΕΠ-π. από πλευράς messaging apps («υπηρεσίες διαπροσωπικών επικοινωνιών ανεξαρτήτως αριθμών»):

- Instant messaging services
- Webmail services
- VoIP - υπηρεσίες φωνητικής επικοινωνίας μέσω διαδικτύου (εξαιρείται η σάρωση ηχητικών επικοινωνιών — άρθρο 1 παρ. 2)

- **Κανονισμός (ΕΕ) 2026/1881 — Επαναφορά προσωρινής παρέκκλισης από την Οδηγία (2002/58/ΕΚ) (31 Ιουλίου 2026 – 3 Απριλίου 2028)**

*** ΣεΚαΕΠ-π.:** Υλικό Σεξουαλικής Κακοποίησης Παιδιών (συντομογραφία χρησιμοποιούμενη στο πλαίσιο του έργου PreventCSA@EU).

Εξέλιξη του ενωσιακού πλαισίου για την ανίχνευση ΣεΚαΕΠ-π. από ιδιωτικούς παρόχους

- **Προτεινόμενος Κανονισμός CSAR — COM (2022) 209 (υπό διαπραγμάτευση)**

Ευρωπαϊκή Επιτροπή (2022): Υποχρεώσεις αξιολόγησης και μετριασμού κινδύνων, εντολές ανίχνευσης και αφαίρεσης ΣεΚαΕΠ-π. προς τους παρόχους.

Δανική Προεδρία του Συμβουλίου (2025): Αφαίρεση εντολών ανίχνευσης, ενίσχυση της αξιολόγησης και μετριασμού κινδύνων, διατήρηση της εθελοντικής ανίχνευσης.

Κυπριακή Προεδρία του Συμβουλίου (2026): Πρόταση για στοχευμένες εντολές ανίχνευσης σε συγκεκριμένους χρήστες και διαφοροποιημένο καθεστώς ανίχνευσης μεταξύ δημόσια και μη δημόσια προσβάσιμου περιεχομένου.

Διοργανικοί Τρίλογοι (2026): Διαπραγματεύσεις μεταξύ Ευρωπαϊκού Κοινοβουλίου και Συμβουλίου, με συμμετοχή της Επιτροπής. Επιμέρους προσωρινές συμφωνίες, χωρίς οριστική υιοθέτηση του Κανονισμού.

Πηγές: COM(2022) 209 · ST 15318/25 (Δανική Προεδρία) · ST 9659/26 (Κυπριακή Προεδρία) · ST 6946/26, ST 8196/26 (διοργανικές διαπραγματεύσεις).

ΑΡΧΗ ΠΡΟΣΤΑΣΙΑΣ ΔΕΔΟΜΕΝΩΝ

Αρχές επιβολής του νόμου (LEAs): επεξεργασία δεδομένων και ρόλος τους στον CSAR

- **Οδηγία (ΕΕ) 2016/680 — Law Enforcement Directive (LED)**

Επεξεργασία προσωπικών δεδομένων για σκοπούς πρόληψης, διερεύνησης και δίωξης ποινικών αδικημάτων, υπό τις αρχές της νομιμότητας, αναγκαιότητας και αναλογικότητας.

Ενσωμάτωση της LED στην ελληνική έννομη τάξη (Μέρος Δ΄ του Ν. 4624/2019). Ειδικές προϋποθέσεις για την άρση του απορρήτου των επικοινωνιών βάσει του Ν. 5002/2022, όπου εφαρμόζεται.

- **Προτεινόμενος Κανονισμός CSAR — COM (2022) 209 - Ρόλος των αρχών επιβολής του νόμου (LEAs)**

- Αξιολόγηση αναφερθέντος υλικού (άρθρο 36§3): Τα κράτη μέλη διασφαλίζουν την αξιολόγηση νέου ΣεΚαΕΠ-π. και υλικού άγρας παιδιών που αναφέρεται στις LEAs. Το επιβεβαιωμένο υλικό υποβάλλεται στο Ευρωπαϊκό Κέντρο από τις εθνικές Συντονιστικές Αρχές.
- Πρόσβαση σε δείκτες (άρθρο 46§4): Πρόσβαση των LEAs και της Europol στη βάση δεικτών του Ευρωπαϊκού Κέντρου για τη διερεύνηση αδικημάτων.
- Παραλαβή αναφορών (άρθρο 48§3): Διαβίβαση αναφορών από το Ευρωπαϊκό Κέντρο στα αρμόδια LEAs και στη Europol για περαιτέρω διερεύνηση.

Πηγή: COM(2022) 209, αρχική πρόταση CSAR (2022) — διατάξεις υπό διαπραγμάτευση.

ΑΡΧΗ ΠΡΟΣΤΑΣΙΑΣ ΔΕΔΟΜΕΝΩΝ

AI Act: ορισμός συστημάτων ΤΝ ως υψηλού κινδύνου βάσει σκοπού χρήσης

Αρχές επιβολής του νόμου (AI Act)

- Η χρήση ΤΝ από Αρχές Επιβολής του Νόμου δεν χαρακτηρίζεται αυτομάτως υψηλού κινδύνου.
- Παράρτημα III§6: ορισμένες χρήσεις ΤΝ χαρακτηρίζονται υψηλού κινδύνου (π.χ. αξιολόγηση κινδύνου τέλεσης ποινικού αδικήματος, αξιολόγηση της αξιοπιστίας αποδεικτικών στοιχείων κατά τη διερεύνηση ή δίωξη ποινικών αδικημάτων κ.α.)
- Εφόσον κριθεί σύστημα υψηλού κινδύνου (άρθρο 27 AI Act) απαιτείται Εκτίμηση Αντικτύπου στα Θεμελιώδη Δικαιώματα (FRIA), πέρα από την DPIA (υπό άρθρο 27 LED).

Ιδιωτικοί πάροχοι (AI Act)

- Η χρήση ΤΝ από παρόχους για ανίχνευση ΣεΚαΕΠ-π. δεν αποτελεί κατηγορία υψηλού κινδύνου στο Παράρτημα III.
- Ως εκ τούτου, δεν απαιτείται Εκτίμηση Αντικτύπου στα Θεμελιώδη Δικαιώματα (FRIA).
- Κανονισμός (ΕΕ) 2026/1881 — Νέο προσωρινό καθεστώς: απαιτείται DPIA και προηγούμενη διαβούλευση (άρθρα 35-36 ΓΚΠΔ) πριν την έναρξη πράξεων επεξεργασίας με χρήση σχετικών τεχνολογιών.

Κίνδυνοι για τα Δεδομένα Προσωπικού Χαρακτήρα

- **CSAR: κίνδυνος γενικής και αδιάκριτης παρακολούθησης**

Κοινή Γνωμοδότηση 4/2022 ΕΣΠΔ/ΕΕΠΔ: ανίχνευση σε ολόκληρη υπηρεσία και όχι στοχευμένα σε ύποπτους λογαριασμούς ενδέχεται να οδηγήσει σε de facto γενικευμένη σάρωση.

Ευρύ πεδίο εφαρμογής μέτρων με ανώτατη διάρκεια έως και 24 μήνες (άρθρο 7 παρ. 9, για γνωστό και νέο υλικό ΣεΚαΕΠ-π.)

Απαιτείται αξιολόγηση αναγκαιότητας και αναλογικότητας ανά τεχνολογία (Άρθρα 7, 8 και 52§1 ΧΘΔΕ).

- **Ακρίβεια των δεδομένων και κίνδυνοι κατάρτισης προφίλ (άρθρα 4(4) και 5(1)(δ) ΓΚΠΔ, σκέψη 71 ΓΚΠΔ).**

Η χρήση ΤΝ για την ανίχνευση ΣεΚαΕΠ-π. μπορεί να οδηγεί στην παραγωγή αλγοριθμικά συναγόμενων δεδομένων προσωπικού χαρακτήρα, τα οποία ενδέχεται να είναι ανακριβή.

Meta: Χρήση μεταδεδομένων και συμπεριφορικών σημάτων για τον εντοπισμό ύποπτων προτύπων συμπεριφοράς, καθώς και μηχανισμών recidivism detection για την αναγνώριση χρηστών που επιχειρούν να δημιουργήσουν νέους λογαριασμούς μετά από αποκλεισμό.

Κίνδυνοι για τα Δεδομένα Προσωπικού Χαρακτήρα

- **Χρόνος Διατήρησης δεδομένων**

Κανονισμός (ΕΕ) 2026/1881: διατήρηση για συγκεκριμένους σκοπούς (αναφορά περιστατικών, περιορισμός λογαριασμών, δημιουργία hashes κ.α.) για όσο είναι απολύτως αναγκαίο - ανώτατο όριο διατήρησης οι 12 μήνες από τον εντοπισμό της υπόνοιας.

Η αναγκαιότητα της διατήρησης πρέπει να τεκμηριώνεται για όσο αυτή διαρκεί (άρθρο 3(1)(h)–(i), αιτιολογική σκέψη 23).

- **Δικαίωμα σε εξήγηση της ατομικής λήψης αποφάσεων:**

ΓΚΠΔ άρθρα 15(1)(η), 22: Ουσιαστική πληροφόρηση για τη λογική της αποκλειστικά αυτοματοποιημένης λήψης αποφάσεων, συμπεριλαμβανομένου του τρόπου χρήσης των προσωπικών δεδομένων.

Όχι γενική περιγραφή του αλγορίθμου (ΔΕΕ, C-203/22, Dun & Bradstreet Austria).

AI Act άρθρο 86: Δικαίωμα εξήγησης του ρόλου της ΤΝ και των βασικών στοιχείων της απόφασης.

Συστήματα Content Moderation: Δεν καλύπτονται κατ' ανάγκη, καθώς δεν χαρακτηρίζονται αυτομάτως ως υψηλού κινδύνου.

DSA άρθρο 17: Υποχρέωση αιτιολόγησης περιορισμών περιεχομένου ή λογαριασμών, με αναφορά στη χρήση αυτοματοποιημένων μέσων για τον εντοπισμό του περιεχομένου ή τη λήψη της απόφασης.

Διασυνδεδεμένο οικοσύστημα ανίχνευσης ΣεΚαΕΠ-π. – CSAR: Η ανταλλαγή πληροφοριών μεταξύ παρόχων, Κέντρου της ΕΕ και διωκτικών αρχών αναδεικνύει διαφορετικές απαιτήσεις διαφάνειας και αιτιολόγησης (ΓΚΠΔ, AI Act, DSA).

ΑΡΧΗ ΠΡΟΣΤΑΣΙΑΣ ΔΕΔΟΜΕΝΩΝ

Κίνδυνοι για τα Δεδομένα Προσωπικού Χαρακτήρα

- **Αξιοπιστία τεχνολογιών και αρχή της ακρίβειας των δεδομένων (άρθρο 5 παρ. 1 στοιχ. δ΄ ΓΚΠΔ)**

Χαμηλή ακρίβεια τεχνολογιών - Έκθεση Εκτίμησης Επιπτώσεων που συνοδεύει την Πρόταση Κανονισμού CSAR (2022):

Ο Thorn Classifier έχει ακρίβεια 99,9%, με δυνατότητα εντοπισμού του 80% ΣεΚαΕΠ-π. στο εξεταζόμενο σύνολο δεδομένων.

Οι τεχνολογίες ανίχνευσης άγρας παιδιών έχουν ακρίβεια 88%.

- ❑ False positives: Κίνδυνος παραγωγής ανακριβών δεδομένων και αδικαιολόγητων αναφορών στις αρμόδιες αρχές.
- ❑ False negatives: Κίνδυνος μη εντοπισμού πραγματικού υλικού ΣεΚαΕΠ-π. – συνέχιση κυκλοφορίας παράνομου υλικού.

Συνεκτίμηση των ανωτέρω κατά τον έλεγχο καταλληλότητας, αναγκαιότητας και αναλογικότητας (άρθρο 52 παρ. 1 ΧΘΔΕΕ).

ΑΡΧΗ ΠΡΟΣΤΑΣΙΑΣ ΔΕΔΟΜΕΝΩΝ

Κίνδυνοι για τα Δεδομένα Προσωπικού Χαρακτήρα από τη δράση Αρχών Επιβολής του Νόμου

- **Ακρίβεια δεδομένων και εσφαλμένες υπόνοιες**:

Κίνδυνος αδικαιολόγητης αστυνομικής διερεύνησης λόγω false positives, εσφαλμένης κατηγοριοποίησης προσώπων και αντιμετώπισης αλγοριθμικών εκτιμήσεων ως εξακριβωμένων γεγονότων (άρθρα 4(1)(δ), 6 και 7 LED).

- **Απαγορευμένη προγνωστική αξιολόγηση εγκληματικής συμπεριφοράς (άρθρο 5 παρ. 1 στοιχ. δ΄ AI Act)**

Απαγορεύεται η εκτίμηση ή πρόβλεψη του κινδύνου τέλεσης αξιόποινης πράξης από φυσικό πρόσωπο, όταν βασίζεται αποκλειστικά σε profiling ή στην αξιολόγηση γνωρισμάτων και χαρακτηριστικών προσωπικότητας.

Η απαγόρευση δεν ισχύει για συστήματα ΤΝ που υποστηρίζουν ανθρώπινη αξιολόγηση συμμετοχής σε εγκληματική δραστηριότητα, ήδη βασισμένη σε αντικειμενικά και επαληθεύσιμα γεγονότα άμεσα συνδεόμενα με αυτή.

- **Αλγοριθμικό profiling και κίνδυνος αδικαιολόγητης αστυνομικής στοχοποίησης**

Η χρήση μεροληπτικών δεδομένων σε συστήματα προγνωστικής αξιολόγησης μπορεί να αναπαράγει διακρίσεις και να ενισχύει την αδικαιολόγητη στοχοποίηση και δυσανάλογη αστυνομική επιτήρηση (FRA, 2022, σ. 9–11, ιδίως FRA Opinion 1).

ΑΡΧΗ ΠΡΟΣΤΑΣΙΑΣ ΔΕΔΟΜΕΝΩΝ

ΑΡΧΗ ΠΡΟΣΤΑΣΙΑΣ ΔΕΔΟΜΕΝΩΝ

Σας ευχαριστώ