



SafeLine – Πιστοποιημένη Πηγή Επισήμανσης Παράνομου Περιεχομένου (DSA Trusted Flagger)

Επεξεργασία Αναφορών 2025

Σύνοψη

Η SafeLine, ως πιστοποιημένη Πηγή Επισήμανσης (Trusted Flagger) στο πλαίσιο του Κανονισμού για τις Ψηφιακές Υπηρεσίες (DSA), συνέχισε το 2025 την αποστολή της για την αναγνώριση και απομάκρυνση παράνομου διαδικτυακού περιεχομένου, συνεργαζόμενη με μεγάλες επιγραμμικές πλατφόρμες όπως Meta, TikTok, Google, Telegram, Discord, Reddit και X. Κατά το εξεταζόμενο διάστημα, η SafeLine υπέβαλε συνολικά 176 αναφορές παράνομου περιεχομένου, εκ των οποίων οι 154 (87%) οδήγησαν σε αφαίρεση περιεχομένου, ενώ 22 (13%) δεν κατέληξαν σε απομάκρυνση. Ο μέσος χρόνος ανταπόκρισης των περισσότερων παρόχων κυμάνθηκε μεταξύ 24 και 48 ωρών, γεγονός που καταδεικνύει τη λειτουργικότητα των μηχανισμών συνεργασίας.

Οι συχνότερες κατηγορίες παραβάσεων περιλάμβαναν:

- Πλαστοπροσωπία και υποκλοπή λογαριασμών (39 αναφορές).
- Ηλεκτρονικό «ψάρεμα» (19 αναφορές).
- Μη συναινετική κοινοποίηση εικόνων (24 αναφορές).
- Πορνογραφικό περιεχόμενο και σεξουαλική εκμετάλλευση.
- Υλικό σεξουαλικής κακοποίησης παιδιών (12 αναφορές).

Η έκθεση αναδεικνύει επίσης περιπτώσεις όπου η αφαίρεση δεν κατέστη δυνατή, κυρίως λόγω:

- Έλλειψης πρόσβασης στο περιεχόμενο (ιδιωτικά περιβάλλοντα).
- Ανεπαρκών στοιχείων ταυτοποίησης.
- Διαφορετικής ερμηνείας των πολιτικών των πλατφορμών.

Η SafeLine λειτουργεί με πλήρη θεσμική και οικονομική ανεξαρτησία από τις επιγραμμικές πλατφόρμες, χρηματοδοτούμενη κατά 50% από την Ευρωπαϊκή Επιτροπή και κατά 50% από ίδιους πόρους του ΙΤΕ. Ως επίσημο μέλος του INHOPE, ακολουθεί αυστηρά τον Κώδικα Ορθής Πρακτικής, διασφαλίζοντας διαφάνεια, λογοδοσία και αξιοπιστία. Η συνολική αποτελεσματικότητα της SafeLine το 2025 καταδεικνύει την κρίσιμη συμβολή των Trusted Flaggers στην ταχεία απομάκρυνση παράνομου περιεχομένου και στην ενίσχυση της ασφάλειας στο ψηφιακό περιβάλλον.



Summary

In 2025, SafeLine operated as a certified Trusted Flagger under the Digital Services Act (DSA), continuing its mission to identify and facilitate the removal of illegal online content in cooperation with major platforms including Meta, TikTok, Google, Telegram, Discord, Reddit and X. During the reporting period, SafeLine submitted 176 content reports, of which 154 (87%) resulted in content removal, while 22 (13%) were not removed. The average response time across platforms ranged between 24 and 48 hours, reflecting effective operational cooperation mechanisms.

The most frequent categories of illegal content included:

- Impersonation and account hijacking (39 reports).
- Phishing (19 reports).
- Non-consensual image sharing (24 reports).
- Pornographic and sexually exploitative content.
- Child sexual abuse material (12 reports).

Cases where content was not removed were primarily due to:

- Lack of access to private environments.
- Insufficient identification evidence.
- Platform policy interpretation differences.

SafeLine operates with full institutional and financial independence from online platforms, funded 50% by the European Commission and 50% by institutional resources of FORTH. As an official INHOPE member hotline, SafeLine adheres to the INHOPE Code of Practice, ensuring transparency, accountability and reliability. The 2025 results demonstrate the critical role of Trusted Flaggers in enabling rapid content removal and strengthening online safety ecosystems under the DSA framework.



Διαδικασίες

Παρακάτω περιγράφεται η διαδικασία ενημέρωσης επιγραμμικών πλατφορμών για τις οποίες η SafeLine έλαβε και επεξεργάστηκε αναφορές το συγκεκριμένο διάστημα:

- **Meta:**

Η αναφορά πραγματοποιείται μέσω του Καναλιού Αξιόπιστων Συνεργατών (Trusted Partner Channel), αξιοποιώντας το Σύστημα Αναφορών Περιεχομένου. Η διαδικασία περιλαμβάνει τρία στάδια: πύλη ελέγχου ταυτότητας, φόρμα αναφοράς και τελική σελίδα προορισμού. Μέσω αυτού του Συστήματος Αναφορών, υπάρχει η δυνατότητα υποβολής ένστασης σε περίπτωση απόρριψης του αιτήματος.

Μέσος χρόνο ανταπόκρισης του παρόχου: 48 ώρες

- **TikTok:**

Η αναφορά γίνεται μέσω του Καναλιού Συνεργατών Κοινότητας (Community Partner Channel) και του Εργαλείου Επιβολής Ασφάλειας, χρησιμοποιώντας το καταχωρημένο email του οργανισμού. Αυτό το Εργαλείο επιτρέπει την παροχή πρόσθετων στοιχείων, εφόσον ζητηθούν από την πλατφόρμα.

Σε περιπτώσεις πλαστοπροσωπίας, το TikTok απαιτεί την αποστολή αντιγράφου της ταυτότητας (ή άλλου επίσημου εγγράφου) του ατόμου του οποίου τα στοιχεία έχουν χρησιμοποιηθεί, ώστε να επιβεβαιώσει την ταυτότητά του και να προχωρήσει στον έλεγχο του ψεύτικου προφίλ. Κατά την επικοινωνία με τον καταγγέλλοντα, διαβεβαιώνουμε ότι το έγγραφο θα χρησιμοποιηθεί αποκλειστικά για την επικοινωνία με το TikTok και δε θα αποθηκευτεί ή χρησιμοποιηθεί για άλλο σκοπό.

Παράλληλα, στις περιπτώσεις πλαστοπροσωπίας, χρειάστηκαν περαιτέρω διευκρινήσεις ή υποβολή ένστασης στην αρχική απόφασης της επιγραμμικής πλατφόρμας, όταν το ψεύτικο προφίλ χρησιμοποιούσε τη φωτογραφία του καταγγέλλοντα, αλλά όχι το όνομά του. Αντίστοιχες διευκρινίσεις κρίθηκαν απαραίτητες και όταν το όνομα που εμφανιζόταν στο ψεύτικο προφίλ διέφερε από αυτό της ταυτότητας, καθώς επρόκειτο για υποκοριστικό του.

Μέσος χρόνο ανταπόκρισης του παρόχου: 24 ώρες

- **Google:**

Η αναφορά γίνεται μέσω του DSA POC Help Center χρησιμοποιώντας το καταχωρημένο email του οργανισμού. Αυτό το Εργαλείο επιτρέπει την παροχή πρόσθετων στοιχείων, εφόσον ζητηθούν από την πλατφόρμα. Εάν έχουμε απορίες σχετικά με το αίτημά μας, μπορούμε να απαντήσουμε στο email επιβεβαίωσης που έχουμε λάβει. Στην απάντησή συμπεριλαμβάνεται ο αριθμός υπόθεσης, ώστε να διευκολυνθεί η παρακολούθηση και η διαχείριση του αιτήματος.

Μέσος χρόνο ανταπόκρισης του παρόχου: 48 ώρες



- **Telegram:**

Η αναφορά πραγματοποιείται με την αποστολή αναλυτικού email, χρησιμοποιώντας το καταχωρημένο email του οργανισμού.

Αν και το Telegram δεν απαιτεί συγκεκριμένη μορφή για τις αναφορές περιεχομένου, κάθε πρόσθετη πληροφορία που μπορούμε να παρέχουμε συμβάλλει σημαντικά στη διευκόλυνση της διαδικασίας.

Για μεγαλύτερη αποτελεσματικότητα, στις αναφορές μας συμπεριλαμβάνουμε συγκεκριμένα στοιχεία ταυτοποίησης λογαριασμών ή περιεχομένου, όπως Telegram usernames (@username), συνδέσμους τύπου t.me/... ή Telegram IDs.

Δύο από τις αναφορές που υποβάλαμε στο Telegram αφορούσαν περιεχόμενο το οποίο βρισκόταν σε ιδιωτικό περιβάλλον, στο οποίο δεν είχαμε πρόσβαση. Ως εκ τούτου, δεν ήταν δυνατό να επιβεβαιώσουμε αν επρόκειτο πράγματι για παράνομο περιεχόμενο, γεγονός που πιθανόν εξηγεί και τη μη αφαίρεσή του.

Μέσος χρόνο ανταπόκρισης του παρόχου: 48 ώρες

- **Discord:**

Ακολουθούμε τον σχετικό σύνδεσμο και εισάγουμε το email που είναι καταχωρημένο στη λίστα DSA Trusted Flagger. Ένας σύνδεσμος θα σταλεί στο inbox μας για επιβεβαίωση της ταυτότητας.

Εντοπίζουμε το email και πατάμε το κουμπί "sign in" (ο σύνδεσμος είναι ενεργός για 10 λεπτά).

Ο σύνδεσμος εισόδου μας μεταφέρει στη σελίδα αιτημάτων του Discord, όπου εμφανίζονται οι οδηγίες και οι λεπτομέρειες της διαδικασίας.

Μέσος χρόνο ανταπόκρισης του παρόχου: 48 ώρες

- **Reddit:**

Κάνουμε χρήση της ειδικής φόρμας για την αναφορά περιεχομένου που θεωρείται παράνομο στην Ευρωπαϊκή Ένωση. Η φόρμα αυτή απευθύνεται τόσο σε ιδιώτες, όσο και σε φυσικά ή νομικά πρόσωπα που έχουν οριστεί ως Trusted Flaggers, σύμφωνα με τον Κανονισμό για τις Ψηφιακές Υπηρεσίες (DSA).

Μέσος χρόνο ανταπόκρισης του παρόχου: 48 ώρες

- **X:**

Κάνουμε login στην πλατφόρμα χρησιμοποιώντας τον λογαριασμό της SafeLine και στη συνέχεια ακολουθούμε τον ειδικό σύνδεσμο για τους Trusted Flaggers στο πλαίσιο του DSA. Αυτό μας μεταφέρει στη σελίδα αιτημάτων, όπου εμφανίζονται οι σχετικές οδηγίες και οι λεπτομέρειες της διαδικασίας.

Μέσος χρόνο ανταπόκρισης του παρόχου: 48 ώρες



Ανεξαρτησία της SafeLine

Η SafeLine είναι απολύτως ανεξάρτητη από τους διάφορους παρόχους επιγραμμικών πλατφορμών. Χρηματοδοτείται κατά το ήμισυ από πόρους της Ευρωπαϊκής Επιτροπής και κατά το υπόλοιπο ήμισυ από ίδιους πόρους του οργανισμού του ΙΤΕ. Ως χρηματοδοτούμενο ερευνητικό έργο εντός του οργανισμού του ΙΤΕ, διαθέτει δικό του κωδικό χρηματοδότησης και λογαριασμό χρέωσης. Η οικονομική της ανεξαρτησία, της διασφαλίζει και την ανεξαρτησία της από τους παρόχους επιγραμμικών πλατφορμών, από τους οποίους δε δέχεται καθόλου χορηγίες. Τέλος, η Γραμμή για την ανεξαρτησία κατά της λειτουργίας της και την επιτέλεση του σκοπού της για την απομάκρυνση του παράνομου περιεχομένου στο διαδίκτυο, ακολουθεί και τον Κώδικα Πρακτικής ([Code of Practice](#)) του ίδιου του INHOPE του οποίου είναι επίσημο μέλος.

Βάσει του Κώδικα Ορθής Πρακτικής του INHOPE (https://inhope.org/media/site/1ffcc1905-1614610382/inhope_codeofpractice.pdf) τον οποίον ακολουθεί η SafeLine ως επίσημο μέλος του, τα hotlines-μέλη του INHOPE απαιτείται να διαβουλεύονται τακτικά με τους σημαντικότερους φορείς της χώρας, όπως οι αρχές επιβολής του Νόμου, η βιομηχανία του διαδικτύου, φορείς παιδικής πρόνοιας κ.α. Παράλληλα, ο Κώδικας επιτάσσει την εφαρμογή από μέρους των hotlines των αρχών της διαφάνειας, της λογοδοσίας και της αξιοπιστίας, ενώ παράλληλα κάθε hotline απαιτείται να δημοσιεύει στοιχεία που αφορούν στην επίσημη δομή της γραμμής, στην διοίκηση και τον συντονισμό, καθώς επίσης και στην χρηματοδότησή της.

Τέλος, αξίζει να σημειωθεί ότι για το έργο του Ελληνικού Κέντρου Ασφαλούς Διαδικτύου, υπάρχει δημόσια προσβάσιμος ο εγκεκριμένος προϋπολογισμός εξόδων του έργου, μέσω της ιστοσελίδας «Διαύγεια» (<https://diavgeia.gov.gr/decision/view/67%CE%9B%CE%9E469%CE%97%CE%9A%CE%A5-%CE%A473>). Αναφορικά με την χρηματοδότηση του έργου από ευρωπαϊκούς πόρους, σχετικές πληροφορίες αναρτώνται στην ιστοσελίδα της Ευρωπαϊκής Επιτροπής (<https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/projects-details/43152860/101081266/DIGITAL?isExactMatch=true&programmePeriod=2021%20-%202027&frameworkProgramme=43152860&order=DESC&pageNumber=NaN&sortBy=title&tenders=false&openForSubmission=false>).

Αναφορές όπου το περιεχόμενο αφαιρέθηκε

[illegible]

or similar technology using a third party's features)																
K30: Δημόσια παροχή πληροφοριών προσωπικής ταυτοποίησης για ένα άτομο (Doxing: publicly providing personally identifiable information about an individual)															0	0
K32: Παρακολούθηση (Stalking)															0	0
K33: Σεξουαλική παρενόχληση (sexual harassment)															0	0
K37: Μη συμπεριληφθέν σε καμία άλλη υποκατηγορία (Πορνογραφία ή σεξουαλικό περιεχόμενο)	2	2					6	5	8	8					16	15
K39: Παιδική πορνογραφία/Υλικό σεξουαλικής κακοποίησης παιδιών (Child pornography/ Child sexual abuse material)							12	12							12	12
K40: Αποπλάνηση/σεξουαλική παρέκλυση ανηλίκων (Grooming/sexual enticement of minors)	1	1													1	1
K52: Πλαστοπροσωπία ή υποκλοπή λογαριασμού (Impersonation or account hijacking)	27	24	10	10			2	2							39	36

Κ53: Ηλεκτρονικό “ψάρεμα” (Phishing)	19	19													19	19
Κ54: Επιχειρηματικό σχήμα πυραμίδας (Pyramid schemes)	1	1													1	1
Κ55: Μη συμπεριλιφθέν σε καμία άλλη υποκατηγορία (διαδικτυακές αγορές)	1	1	1	1			2	2							4	4
Κ57: Παρακίνηση σε αυτοτραυματισμό (Incitement to self-mutilation)															0	0
Κ72: Οργανωμένη βία (Coordinated harm)															0	0
Κ73: Έμφυλη βία (Gender-based violence)															0	0
Κ74: Εκμετάλλευση ανθρώπων (Human trafficking)															0	0
Τυχρά παιχνίδια (gambling)					3	3	2	2							5	5
Προϊόντα καπνού (tobacco)					1	1									1	1
Άλλο (other)									7	2					7	2
Σύνολο	62	57	28	28	8	5	30	23	15	10	2	2	9	7	154	132

Αναφορές όπου το περιεχόμενο δεν αφαιρέθηκε

Επιγραμμική πλατφόρμα	Meta	Google	Telegram	Discord	x	not removed
Κατηγορίες παράνομου περιεχομένου						
K12: Ρητορική μίσους (Hate speech)		1			2	3
K28: Μη συναινετική κοινή χρήση εικόνων (Non-consensual image sharing)	2	2	6			10
K37: Μη συμπεριληφθέν σε καμία άλλη υποκατηγορία (Πορνογραφία ή σεξουαλικό περιεχόμενο)			1			1
K52: Πλαστοπροσωπία ή υποκλοπή λογαριασμού (Impersonation or account hijacking)	3					3
Άλλο (other)				5		5
Σύνολο	5	3	7	5	2	22

Λόγοι μη αφαίρεσης του αναφερθέντος περιεχομένου

- **Meta (Πλαστοπροσωπία)**

Η Meta δε μπόρεσε να επιβεβαιώσει παραβίαση των πολιτικών αυθεντικότητας και ζήτησε είτε το URL του αυθεντικού προφίλ είτε αντίγραφο ταυτότητας, αλλά ο καταγγέλλων δεν μας παρείχε αυτά τα στοιχεία.

- **Meta (Μη συναινετική κοινή χρήση εικόνων)**

Η Meta απάντησε ότι δεν μπορεί να υποστηρίξει αυτό το αίτημα μέσω της συγκεκριμένης φόρμας και ότι τέτοια ζητήματα πρέπει να αναφέρονται μέσω των εργαλείων αναφοράς για παραβίαση ιδιωτικότητας.

- **Google**

Η ειδοποίηση πρέπει να γίνει απευθείας από το ίδιο το άτομο που αφορά το ζήτημα ιδιωτικότητας, τον νόμιμο εκπρόσωπο ή τον κηδεμόνα του.

- **Telegram**

Χωρίς πρόσβαση στο αναφερόμενο περιεχόμενο, επομένως δεν γνωρίζουμε αν πρόκειται πράγματι για παράνομο υλικό. Η πλατφόρμα εξέτασε το περιεχόμενο και έκρινε πως δεν εμπίπτει σε λόγους αφαίρεσης



- **Discord**

Χωρίς πρόσβαση στο αναφερόμενο περιεχόμενο, επομένως δεν γνωρίζουμε αν πρόκειται πράγματι για παράνομο υλικό. Η πλατφόρμα εξέτασε το περιεχόμενο και έκρινε πως δεν εμπίπτει σε λόγους αφαίρεσης

- **X**

Εξέτασε το περιεχόμενο και έκρινε πως δεν εμπίπτει σε λόγους αφαίρεσης βάσει της κατηγορίας «Illegal or harmful speech» στην Ελλάδα.