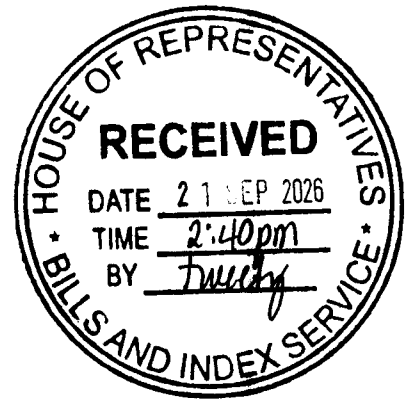




Republic of the Philippines
HOUSE OF REPRESENTATIVES
Quezon City, Metro Manila

TWENTIETH CONGRESS
Second Regular Session



HOUSE BILL NO. 11397

Introduced by Representative DE CARLO "OYO" L. UY

EXPLANATORY NOTE

The rapid growth of social media and livestreaming platforms has increased the speed and reach by which actual acts of extreme violence may be recorded, distributed, and repeatedly circulated online. When such content is algorithmically amplified or monetized, it may expose children and the public to harmful material while potentially encouraging further circulation and imitation.

Existing laws and platform policies provide mechanisms for addressing unlawful online content, but gaps remain in addressing the rapid livestreaming, repeated re-uploading, algorithmic recommendation, and monetization of actual extreme violence. These risks require preventive measures that are responsive to the speed and scale of digital platforms.

The proposed measure establishes responsibilities for covered online platforms to detect and interrupt livestreams depicting actual extreme violence, limit repeated circulation and algorithmic amplification, prevent the recommendation of such content to children, and prohibit its monetization. It likewise provides for emergency coordination with Philippine authorities and the preservation of relevant digital evidence for lawful investigations.

The measure balances these obligations with constitutional rights and legitimate public-interest uses of violent-content documentation, including journalism, education, research, human rights documentation, and law enforcement. It adopts a risk-based approach to platform compliance and provides proportionate administrative penalties for serious or repeated violations.

The immediate approval of this urgent bill is earnestly sought.



DE CARLO "OYO" L. UY
Representative
First District, Davao del Norte



Republic of the Philippines
HOUSE OF REPRESENTATIVES
Quezon City, Metro Manila

TWENTIETH CONGRESS
Second Regular Session

HOUSE BILL NO. 11397

Introduced by Representative DE CARLO “OYO” L. UY

AN ACT
ESTABLISHING MEASURES TO PREVENT THE DISSEMINATION, AMPLIFICATION,
AND MONETIZATION OF ACTUAL EXTREME VIOLENCE ON ONLINE PLATFORMS,
PROTECTING CHILDREN AND THE PUBLIC FROM EXPOSURE TO SUCH CONTENT,
REQUIRING PLATFORM ACCOUNTABILITY AND EMERGENCY COORDINATION,
AND FOR OTHER PURPOSES

Be enacted by the Senate and House of Representatives of the Philippines in Congress assembled:

CHAPTER I.
GENERAL PROVISIONS

SECTION 1. Title. – This Act shall be known as the “Safe Screens Act of 2026.”

SECTION 2. Declaration of Policy. – It is hereby declared the policy of the State to promote a safe and responsible digital environment and protect children and the public from the harmful effects of the rapid dissemination, repeated circulation, algorithmic amplification, and monetization of actual extreme violence online.

The State recognizes the importance of social media and online platforms in facilitating communication, access to information, journalism, education, research, and public participation. At the same time, the State recognizes that the design and operation of online platforms may contribute to the rapid and repeated exposure of users, particularly children, to graphic depictions of actual violence.

Accordingly, the State shall establish proportionate and risk-based responsibilities for covered online platforms to prevent or mitigate the unnecessary dissemination and amplification of actual extreme-violence content, facilitate timely coordination during emergencies, preserve relevant digital evidence for lawful investigations, and protect children from exposure to such content.

Nothing in this Act shall be construed to restrict lawful freedom of expression, freedom of the press, access to information, academic freedom, or legitimate journalism, education, research, law enforcement, documentation, or other public-interest activities.

SECTION 3. Objectives. – This Act shall pursue the following objectives:

- a) Prevent or minimize the livestreaming and rapid dissemination of actual extreme violence;
- b) Reduce the repeated circulation and algorithmic amplification of identified violent content;
- c) Protect children from being proactively exposed or recommended such content;
- d) Prevent the monetization of actual extreme-violence content, subject to public-interest safeguards;
- e) Establish mechanisms for timely coordination between online platforms and Philippine authorities during emergencies involving threats to life and public safety;
- f) Facilitate the lawful preservation of digital evidence relevant to criminal investigations and proceedings; and
- g) Establish proportionate accountability measures for serious or repeated violations by covered platforms.

SECTION 4. *Definition of Terms.* – For the purposes of this Act, the following terms shall be used:

- a) *“Abhorrent violent material”* refers to audio, visual, or audio-visual material that records or streams abhorrent violent conduct, where a reasonable person would regard the material as offensive in all the circumstances, and which is produced by a person who committed, conspired to commit, aided, abetted, counselled, procured, or was otherwise knowingly concerned in, or attempted to commit, the abhorrent violent conduct.
- b) *“Algorithmic Amplification”* refers to the automated recommendation, ranking, promotion, or distribution of content by an online platform in a manner that materially increases its visibility or reach beyond users who have intentionally sought, followed, or otherwise accessed such content;
- c) *“Covered Online Platform”* refers to an internet intermediary or participative network platform provider that enables users in the Philippines to create, upload, share, livestream, recommend, or otherwise distribute user-generated content, subject to the coverage thresholds established under this Act and its implementing rules and regulations;
- d) *“Emergency Request”* refers to a request transmitted by a competent government authority to a covered online platform for immediate action concerning content or information reasonably believed to involve an ongoing or imminent threat to life or public safety.
- e) *“Graphic Violent Content”* refers to other authentic depictions of severe real-world violence, subject to the safeguards and proportionate measures provided in this Act.
- f) *“Identified Violent Content”* refers to content determined, through the content governance processes of a covered online platform or by a competent government authority pursuant to law, to depict actual extreme violence covered by this Act;
- g) *“Monetization”* refers to the generation of revenue or other financial benefit from content through advertising, subscriptions, donations, paid interactions, revenue-sharing arrangements, or other comparable means;
- h) *“Prohibited Extreme-Violence Livestream”* refers to a livestream generated, directed, or intentionally disseminated by a perpetrator or accomplice in connection with an ongoing killing, mass-casualty attack, terrorist act causing serious physical harm or death, torture, hostage violence, or other comparable act of extreme violence, or a livestream whose continued dissemination creates a specific and imminent danger to life or materially facilitates the ongoing offense.

- 105 i) “Public-Interest Content” refers to content involving legitimate journalism, education,
106 research, human rights documentation, law enforcement, judicial proceedings, or other lawful
107 activities undertaken for a legitimate public-interest purpose; and
108
109 j) For purposes of this Act, the terms “child,” “computer,” “computer data,” “content data,”
110 “information and communications technology,” “internet intermediary,” “participative
111 network platform provider,” “streaming,” “live-streaming,” “subscriber’s information,” and
112 “registration information” shall have the meanings provided under Republic Act No. 11930,
113 otherwise known as the “Anti-Online Sexual Abuse or Exploitation of Children and Anti-Child
114 Sexual Abuse or Exploitation Materials Act,” and other applicable laws.
115

116
117 **CHAPTER II.**
118 **PREVENTION AND PLATFORM RESPONSIBILITIES**
119

120 **SECTION 5. *Prevention of Livestreaming of Actual Extreme Violence.*** - *A covered online platform*
121 *shall take reasonable and proportionate measures to promptly suspend, interrupt, or restrict a*
122 *livestream when it has reasonable grounds to believe that:*
123

- 124 a) *the livestream is generated or intentionally disseminated by the perpetrator*
125 *or an accomplice in connection with an ongoing act of extreme violence; or*
126 b) *continued livestreaming creates a specific and imminent danger to life or*
127 *materially facilitates the ongoing offense.*

128
129 *In other cases involving graphic violent content, the platform shall employ the least restrictive*
130 *reasonably available measure consistent with public safety and the public-interest protections under*
131 *this Act, such as, but not limited to:*
132

- 133 a) Promptly suspend, interrupt, or restrict the livestream;
134 b) Prevent or minimize further dissemination of the livestream while the incident is being
135 assessed;
136 c) Where appropriate, notify the user that the livestream has been restricted and provide an
137 accessible mechanism for review; and
138 d) Where the circumstances indicate an imminent threat to life or public safety, activate the
139 emergency coordination mechanism under this Act.
140

141 Platforms shall not be required to achieve absolute or error-free detection. Compliance shall be assessed
142 based on the reasonable measures adopted and the platform's efforts to implement them effectively.
143

144 **SECTION 6. *Measures Against Repeated Circulation.*** – Covered online platforms shall adopt age-
145 appropriate measures for accounts reasonably identified as belonging to children, which may include
146 disabling autoplay, warning interstitials, blurred graphic previews, and restrictions on proactive
147 recommendation of identified graphic violent content.
148

149 Such measures may include the use of content-matching technologies, hashes, digital fingerprints,
150 metadata, or other appropriate technical measures.
151

152 Measures under this section shall be designed to minimize erroneous removal or restriction of
153 substantially different content and shall provide appropriate safeguards for legitimate public-interest
154 content.
155

156 **SECTION 7. *Restriction on Algorithmic Amplification.*** - Covered online platforms shall not
157 intentionally or knowingly use recommendation, ranking, or other automated distribution systems to

algorithmically amplify identified violent content in a manner that materially increases its exposure to users who have not intentionally sought or accessed such content.

Nothing in this section shall prohibit the ordinary availability of lawful content that discusses, reports on, analyzes, or documents an incident involving actual extreme violence, provided that the content is not itself prohibited under this Act.

SECTION 8. *Protection of Children from Recommended Violent Content.* - Covered online platforms shall establish reasonable measures to prevent identified actual extreme-violence content from being proactively recommended, automatically played, or otherwise algorithmically presented to accounts reasonably identified as belonging to children.

Platforms shall adopt age-appropriate safety settings and other reasonable measures consistent with applicable laws on privacy and data protection.

SECTION 9. *Prohibition on Monetization of Actual Extreme Violence.* - A covered online platform shall not knowingly enable monetization where the uploader or platform principally exploits the graphic depiction of actual extreme violence itself for commercial gain.

This prohibition shall not apply to bona fide journalism or citizen journalism, documentary reporting, education, research, human-rights documentation, historical archives, judicial or law-enforcement use, or other legitimate public-interest content.

SECTION 10. *Platform Content Governance Measures.* - Covered online platforms shall establish and maintain publicly available policies governing the handling of actual extreme-violence content.

Such policies shall specify, at a minimum:

- (a) The types of content covered;
- (b) Procedures for reporting and reviewing such content;
- (c) Measures for interrupting or restricting livestreams;
- (d) Measures for preventing repeated circulation and algorithmic amplification;
- (e) Measures for protecting children;
- (f) Procedures for appeals and review; and
- (g) Measures for compliance with lawful requests from Philippine authorities.

CHAPTER III. EMERGENCY COORDINATION AND DIGITAL EVIDENCE

SECTION 11. *Twenty-Four-Hour Emergency Coordination Mechanism.* - Covered online platforms shall maintain a dedicated emergency coordination channel available twenty-four (24) hours a day, seven (7) days a week for authorized Philippine government authorities in cases involving:

- a) An ongoing or imminent violent attack;
- b) A livestream depicting an ongoing act of extreme violence;
- c) An imminent threat to the life or physical safety of one or more persons;
- d) A school attack or other mass-casualty incident; or
- e) Other circumstances presenting an immediate and substantial threat to public safety.

The designated government authority shall establish protocols identifying authorized government personnel and the circumstances under which the emergency channel may be activated.

No government authority may compel the removal, blocking, or restriction of expressive content under this Act except pursuant to authority expressly conferred by law and in accordance with applicable constitutional and judicial-process requirements.

Emergency requests shall be limited to actions and information reasonably necessary to address a specific and imminent threat to life or public safety.

Nothing in this Section shall independently authorize access to private communications, content data, traffic data, or other protected information except in accordance with the 1987 Constitution and the legal process applicable to the particular category of information requested.

SECTION 12. *Emergency Requests.* - Upon receipt of a valid emergency request from an authorized government authority, a covered online platform shall, consistent with applicable law:

- a) Take reasonable steps to restrict or interrupt the further dissemination of content directly associated with an ongoing violent incident;
- b) Preserve relevant information and records in accordance with Section 13; and
- c) Provide available information necessary to address the imminent threat, subject to applicable legal requirements.

Emergency requests shall be limited to information and actions reasonably necessary to address the identified emergency.

SECTION 13. *Preservation of Digital Evidence.* – Covered online platforms shall establish procedures for the preservation of digital information reasonably believed to be relevant to an ongoing or reasonably anticipated criminal investigation involving actual extreme violence. A preservation request shall identify with reasonable particularity the account, content, record, or data to be preserved and shall not, by itself, authorize access or disclosure.

Upon receipt of a valid preservation request from a competent Philippine authority, the platform shall preserve the specified information for a period prescribed by the implementing rules and regulations, or until otherwise ordered by a competent authority. The information shall be preserved for the prescribed period, subject to extension pursuant to lawful authority.

Preservation under this section shall not, by itself, authorize disclosure or access to the preserved information. Disclosure shall remain subject to the 1987 Constitution and the specific legal process applicable to the information requested, including laws governing privacy, electronic evidence, search and seizure, and lawful access to communications and data.

SECTION 14. *Chain of Custody and Integrity of Digital Evidence.* – Covered online platforms shall adopt reasonable procedures to preserve the integrity, authenticity, and relevant metadata of digital information retained pursuant to this Act.

The implementing rules and regulations shall establish appropriate procedures for the lawful transfer and authentication of preserved digital evidence.

CHAPTER IV. REPORTING, REVIEW, AND SAFEGUARDS

SECTION 15. *User Reporting Mechanisms.* - Covered online platforms shall provide accessible mechanisms through which users may report content that may constitute actual extreme violence or

otherwise fall within the prohibited acts under this Act. Reports involving an ongoing or imminent threat to life shall be prioritized.

SECTION 16. *Notice and Review.* - Where content is restricted or removed pursuant to this Act, the platform shall, when reasonably practicable, notify the affected user of:

- a) The action taken;
- b) The general basis for the action; and
- c) The available mechanism for review or appeal.

Nothing in this section shall require a platform to disclose information that would compromise an ongoing criminal investigation, reveal confidential detection methods, or create an immediate risk to public safety.

SECTION 17. *Protection of Legitimate Public-Interest Content.* - Nothing in this Act shall be construed to require the removal or restriction of content solely because it:

- a) Reports on an act of violence;
- b) Documents an event for journalistic, historical, or human rights purposes;
- c) Is used for legitimate educational or academic purposes;
- d) Is used by law enforcement or other competent authorities;
- e) Is necessary for judicial or investigative proceedings; or
- f) Serves another legitimate public-interest purpose.

In implementing this Act, covered online platforms shall distinguish, to the extent reasonably practicable, between the graphic depiction of actual violence for purposes of gratuitous dissemination or monetization and its legitimate documentation or discussion for public-interest purposes.

SECTION 18. *Protection of Freedom of Expression and Privacy.* - The implementation of this Act shall be consistent with the 1987 Constitution and applicable laws on freedom of expression, freedom of the press, privacy, data protection, access to information, and due process.

Nothing in this Act shall be construed as authorizing the general surveillance of users, indiscriminate monitoring of private communications, or access to private information without lawful authority.

The duties and restrictions under this Act shall not require the removal or restriction of material whose publication, accessibility, preservation, or use is reasonably necessary for bona fide journalism or citizen journalism, documentary reporting, education, scientific or academic research, historical archiving, human-rights documentation, law enforcement, emergency response, judicial proceedings, evidentiary preservation, or another legitimate public-interest purpose, provided that the manner of presentation is reasonably proportionate to such purpose.

CHAPTER V. PLATFORM ACCOUNTABILITY

SECTION 19. *Risk-Based Compliance.* - The designated government authority shall adopt a risk-based framework for implementation of this Act, taking into account the size, reach, functionality, user base, and potential impact of covered online platforms.

Small or low-risk platforms may be subject to proportionately different compliance requirements, as may be provided in the implementing rules and regulations.

SECTION 20. *Compliance Reports.* - Covered online platforms shall submit periodic reports to the designated government authority containing aggregated and anonymized information on:

- 320
321 a) The number of reports involving actual extreme-violence content;
322 b) The number of livestreams interrupted or restricted;
323 c) The number of identified violent-content uploads restricted from repeated circulation;
324 d) Measures undertaken to protect children;
325 e) The number of valid emergency requests received and acted upon; and
326 f) Other information reasonably necessary to assess compliance.

327
328 The reports shall not require the disclosure of personal information except when authorized under
329 applicable law.

330
331 **SECTION 21. *Government Monitoring and Audit.*** - The designated government authority may
332 conduct compliance assessments of covered online platforms consistent with applicable laws and due
333 process.

334
335 Any inspection, audit, or request for information shall be limited to matters reasonably necessary to
336 determine compliance with this Act and shall protect confidential business information and personal
337 data.

338
339 **SECTION 22. *Administrative Penalties.*** - After notice and opportunity to be heard, a covered online
340 platform that willfully or repeatedly fails to comply with a mandatory obligation under this Act may be
341 subject, as appropriate, to:

- 342
343 a) a warning or compliance order;
344 b) a corrective action order;
345 c) an administrative fine, subject to the limits prescribed under this Act; or
346 d) such other administrative measure expressly authorized by this Act.

347
348 In determining the appropriate sanction, the implementing authority shall consider the gravity and
349 duration of the violation, the number of users affected, the platform's size and reach in the Philippines,
350 whether the violation was willful or repeated, the platform's prior compliance history, the measures
351 taken to prevent or correct the violation, and the proportionality of the sanction to the harm involved.

352
353 No administrative penalty shall be imposed solely because of an isolated good-faith moderation error
354 where the platform has adopted and implemented reasonable compliance measures.

355
356 **SECTION 23. *Enhanced Administrative Fine.*** — Repeated or willful violations committed despite a
357 prior final compliance or corrective action order may be subject to an enhanced administrative fine
358 within the maximum limit prescribed by this Act.

359
360 The implementing rules and regulations may prescribe the manner of determining the appropriate fine
361 within such statutory limit, taking into account the factors stated in Section 22.

362
363 The implementing rules and regulations shall prescribe the procedures for notice, investigation, hearing,
364 determination of liability, computation of administrative fines within the limits and standards
365 established by this Act, reconsideration or appeal, and enforcement of final orders.

366
367 **SECTION 24. *Limitation on Platform Liability.*** - Nothing in this Act shall be construed to impose
368 strict liability upon a covered online platform solely because unlawful or harmful user-generated
369 content appears on its service despite reasonable compliance measures under this Act.

370
371 This Section shall be construed consistently with Section 30 of Republic Act No. 8792, without
372 prejudice to liability for the platform's own unlawful acts or its willful or grossly negligent failure to
373 perform a specific obligation imposed by this Act.

375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418

CHAPTER VI.
INSTITUTIONAL ARRANGEMENTS

SECTION 25. *Designated Government Authority.* - The Department of Information and Communications Technology (DICT) shall serve as the lead agency in coordinating the implementation of this Act.

The DICT shall coordinate with the National Telecommunications Commission (NTC), National Privacy Commission (NPC), Department of Justice (DOJ), Philippine National Police (PNP), Department of Education (DepEd), Commission on Higher Education (CHED), Department of Social Welfare and Development (DSWD), and other concerned agencies in the development and implementation of policies under this Act.

SECTION 26. *Duties of the DICT.* - The DICT shall:

- a) Establish the national framework for platform compliance;
- b) Establish and maintain the emergency coordination mechanism under this Act;
- c) Develop standards for platform reporting and compliance monitoring;
- d) Coordinate with law enforcement and other government agencies during emergencies;
- e) Develop public awareness programs on responsible digital citizenship and online violence; and
- f) Submit an annual report to Congress on the implementation of this Act.

SECTION 27. *Duties of the National Privacy Commission.* - The NPC shall provide guidance on the processing, retention, disclosure, and protection of personal information in connection with the implementation of this Act, consistent with Republic Act No. 10173, otherwise known as the "Data Privacy Act of 2012."

SECTION 28. *Duties of the Department of Education.* - The DepEd shall develop age-appropriate digital safety and media literacy programs addressing exposure to violent online content, responsible reporting of harmful content, and appropriate responses to online threats and violent incidents.

SECTION 29. *Inter-Agency Coordination.* - The DICT shall establish protocols for coordination among concerned government agencies to ensure timely and proportionate responses to online content involving actual extreme violence and imminent threats to life or public safety.

The protocols shall clearly identify the responsibilities of each concerned agency and shall include safeguards against duplication, unauthorized access to information, and unnecessary disclosure of personal data.

415
416
417
418
419
420
421
422
423
424
425
426
427
428
429

CHAPTER VII.
FUNDING AND OVERSIGHT

SECTION 30. *Appropriations.* - The amount necessary for the implementation of this Act shall be charged against the appropriations of the DICT and other concerned agencies under the current General Appropriations Act.

Thereafter, such sums as may be necessary for the continued implementation of this Act shall be included in the annual General Appropriations Act.

Nothing in this Act shall be construed as authorizing the creation or use of confidential or intelligence funds.

SECTION 31. *Congressional Oversight and Annual Report.* - The appropriate committees of the Senate and the House of Representatives shall exercise oversight over the implementation of this Act.

The DICT shall submit an annual report containing an assessment of implementation, compliance trends, enforcement actions, emerging risks, and recommendations for legislative or policy improvements.

CHAPTER VIII. FINAL PROVISIONS

SECTION 32. *Implementing Rules and Regulations.* – Within one hundred eighty (180) days from the effectivity of this Act, the DICT, in coordination with the NTC, NPC, DOJ, PNP, DepEd, CHED, and other concerned agencies, shall promulgate the rules and regulations necessary for its implementation.

The implementing rules and regulations shall establish, among others:

- a) The thresholds for covered online platforms;
- b) Minimum technical and organizational measures;
- c) Emergency coordination procedures;
- d) Digital evidence preservation procedures;
- e) Reporting and compliance requirements;
- f) Administrative penalties and enforcement procedures; and
- g) Safeguards for freedom of expression, privacy, journalism, education, research, and other legitimate public-interest activities.

SECTION 33. *Separability Clause.* – If any provision or portion of this Act is declared unconstitutional, the remainder or any provision of this Act not thereby affected shall remain in full force and effect.

SECTION 34. *Repealing Clause.* – All laws, decrees, orders, rules and regulations, other issuances, or parts thereof inconsistent with the provisions of this Act are hereby repealed or modified accordingly.

SECTION 35. *Effectivity.* This Act shall take effect fifteen (15) days after its publication in the Official Gazette or in a newspaper of general circulation.

Approved,