

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

**TWENTIETH CONGRESS**  
First Regular Session

**HOUSE BILL NO. 3472**



---

Introduced by Representative Jan Rurik D. Padiernos, GP (Galing sa Pusong) Party-List

---

### **EXPLANATORY NOTE**

In today's digital landscape, information is disseminated instantly through social media platforms, messaging apps, and algorithm-driven content feeds. These technologies have made it easier access information, but they have also made it easier for "fake news" or misleading content to spread quickly. This harms public discourse, weakens institutions, and damages the integrity of elections and public trust.

Although there are existing laws, such as Article 154 of the Revised Penal Code (unlawful publication) and Republic Act No. 10175 (Cybercrime Prevention Act of 2012, that cover some forms of online misconduct, they are not sufficient to counter the proliferation of disinformation tactics. These include the organized use of troll farms, coordinated inauthentic behavior, fake accounts, and algorithmic manipulation, often timed to influence elections, sow confusion, or exploit public fears.

AI-generated content (audio or videos) can now create the illusion that an individual is advocating for political beliefs or promoting falsehoods when in reality, they are not. This has a detrimental impact on reputations and confuses the public, particularly during significant events such as national emergencies or elections.

Other jurisdictions have taken legal steps to address these issues. Singapore's Protection from Online Falsehoods and Manipulation Act (POFMA), the European Union's Digital Services Act (DSA), and policy framework from technology platforms like Meta and academic research from the Stanford Internet Observatory all recognize the need for enforceable, rights-respecting mechanisms to counter disinformation.

This bill seeks to respond to that challenge by:

- Providing clear definitions of online falsehoods, including synthetic and AI-altered media, and requiring evidence of serious and verifiable harm;
- Establishing the Cyber Disinformation and Digital Truth Office (CDDTO) under the Department of Information and Communications Technology (DICT) to monitor, verify, and recommend administrative enforcement through Correction, Stop Communication, or Takedown Orders;
- Clarifying that criminal prosecution for serious disinformation acts shall be conducted by the Department of Justice (DOJ) and its deputized agencies, consistent with due process under the Rules of Court;

- Creating a public reporting mechanism for individuals and institutions to alert authorities to emerging falsehoods;
- Promoting nationwide digital literacy and public education programs in coordination with Department of Education (DepEd), Philippine Information Agency(PIA), and other stakeholders;
- Protecting freedom of expression through clear exemptions for opinion, satire, parody, and fair comment; and
- Providing fair and accessible appeal mechanisms through the DICT and the judiciary, consistent with administrative due process and constitutional rights.

This legislation is not intended to suppress dissent or curtail legitimate expression. Rather, it fills a legal gap through aiming to curtail malicious disinformation that endangers public interest—while upholding constitutional freedoms and democratic values in the online space.

In view of the foregoing, the immediate passage of this bill is earnestly sought.

  
**ATTY. JAN RURIK D. PADIERNOS**  
GP (Galing sa Puso) Party-List

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

**TWENTIETH CONGRESS**  
First Regular Session

**HOUSE BILL NO. 3472**

---

Introduced by Representative Jan Rurik D. Padiernos, GP (Galing sa Pusong) Party-List

---

**AN ACT**  
**ENSURING THE INTEGRITY AND RESILIENCE OF ONLINE INFORMATION**  
**BY DETERRING ALL FORMS OF ONLINE FALSEHOODS AND**  
**DISINFORMATION**

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

**SECTION 1. *Short Title.*** - This Act shall be known as the "***Online Falsehoods and Disinformation Prevention and Accountability Act.***"

**SEC. 2. *Declaration of Policy.*** - It is the policy of the State to protect the public from online falsehoods and disinformation that endanger public health, electoral integrity, national security, and trust in institutions. The State shall ensure that such protection respects the constitutional right to freedom of expression. It likewise promotes digital literacy and critical thinking to empower responsible and informed digital citizens.

**SEC. 3. *Definition of Terms.*** - For purposes of this Act, the following terms are hereby defined as follows:

- (a) *Online Falsehood* refers to any false statement of fact communicated through digital or online means that is likely to cause serious and demonstrable harm to clearly defined public interests, such as threats to public health or safety, disruption of electoral integrity, national security risks, or erosion of lawful trust in government institutions, and only when such harm is immediate, substantial, and verifiable.
- (b) *Disinformation* refers to the intentional creation or dissemination of false or misleading content, including manipulated media, designed to deceive or mislead the public.
- (c) *Coordinated inauthentic behavior* refers to the use of multiple fake, duplicate, or automated online accounts, pages, or networks to mislead users, manipulate public discourse, or artificially amplify or suppress information online, typically in a deceptive or concealed manner.
- (d) *Serious and verifiable harm to public interest* refers to substantial and demonstrable injury or risk caused by online falsehoods or disinformation to public safety, health, national security, electoral integrity, or public order, which is supported by factual, technical, or

- expert evidence.
- (e) *Deepfake* refers to synthetic media created or altered using artificial intelligence to falsely depict individuals as saying or doing things they did not actually say or do.
  - (f) *Troll Farm* refers to a coordinated group or network using fake or duplicate accounts to spread disinformation, distort public discourse, or manipulate online engagement.
  - (g) *Correction Order* refers to an order requiring the publication of a correction notice on content identified as a falsehood.
  - (h) *Stop Communication Order* to an order requiring the individual or platform to cease dissemination of the identified false content.
  - (i) *Takedown Order* to an order requiring the removal of false content from websites, applications, or social media platforms.

**SEC. 4. *Creation of the Cyber Disinformation and Digital Truth Office and Its Duties and Functions.*** — A Cyber Disinformation and Digital Truth Office (CDDTO) is hereby created under the Department of Information and Communications Technology (DICT). The Office (DICT-CDDTO) shall have the following duties and functions:

- a. Monitor and verify the existence of online falsehoods and disinformation posing serious and verifiable harm to public interest, including those disseminated through troll farms, coordinated inauthentic behavior, or deepfake technologies;
- b. Issue or impose Correction, Stop Communication, or Takedown Orders against any person;
- c. Coordinate with online platforms and service providers for enforcement;
- d. Educate the public through digital literacy programs;
- e. Facilitate reports from individuals and institutions concerning potential disinformation;
- f. Provide for a mechanism of appeal and review;
- g. Ensure transparency, regular reporting, and safeguards against abuse;
- h. Coordinate with law enforcement agencies, including the National Bureau of Investigation (NBI) and the Department of Justice (DOJ) Cybercrime Offices, for the investigation and enforcement of orders where criminal liability may arise; and
- i. Exercise quasi-judicial functions in resolving appeals from administrative enforcement orders issued under this Act, in accordance with applicable rules and procedures.

**SEC. 5. *Exemptions and Safeguards.*** —This Act shall not apply to:

- a. Opinion or commentary clearly not presented as a statement of fact;
- b. Satire or parody, including political or social commentary expressed through humor or exaggeration;
- c. Fair reporting or criticism, made in good faith on matters of public interest;
- d. Journalistic or academic work, provided that reasonable verification was undertaken and there is no intent to mislead.

These exemptions shall be interpreted in accordance with the right to freedom of speech and of the press under Article III, Section 4 of the 1987 Constitution.

**SEC. 6. *Administrative Enforcement.*** — The CDDTO shall monitor and verify potential violations of the Prohibited Acts as defined under Section \_\_\_ of this Act. Upon

verification that such prohibited acts were committed knowingly and with intent to cause serious and verifiable harm to public interest, the CDDTO shall recommend appropriate administrative enforcement actions to the Department of Information and Communications Technology (DICT).

The following administrative Orders may be issued under this Act:

- a. **Correction Order** – Directs the subject to publish a correction notice with a link to official clarification;
- b. **Stop Communication Order** – Requires the subject to stop further dissemination of the falsehood;
- c. **Takedown Order** – Mandates the removal of content from public access.

Orders shall be issued in accordance with the standards and procedures provided in Section 7 of this Act.

Entities found to have violated Section 5 may be subject to these Orders. Non-compliance shall result in administrative penalties under Section 9 of this Act.

In cases of repeated violations or those involving aggravated harm, the CDDTO shall coordinate with the Department of Justice for possible criminal prosecution in accordance with Section 8 and 9 of this Act.

**SEC. 7. Issuance of Orders.** — A Correction, Stop Communication, or Takedown Order may be issued by the Secretary of DICT, upon recommendation of the CDDTO, based on verified findings that an online falsehood or disinformation:

- a. Exists and has been published;
- b. Results in, or is likely to result in, serious and verifiable harm to public interest, as defined under this Act; and
- c. Can be remedied or mitigated through timely intervention.

Each Order shall clearly state the following:

- i. The specific nature and content of the falsehood or manipulated media;
- ii. The basis for the finding of intent and harm;
- iii. The corrective action required and compliance period; and
- iv. The legal remedies available, including the right to appeal.

The affected party shall be notified and granted the opportunity to respond within a reasonable period. Failure to comply with an issued Order may result in administrative fines by the CDDTO. In cases of continued non-compliance or aggravated acts, the matter shall be referred to the DOJ for appropriate filing of the case.

**SEC. 8. Appeal and Review.** – Any person subject to an order may appeal in writing to the DICT within ten (10) days of receipt. Final administrative decisions may be appealed to the Court of Appeals under Rule 43 of the Rules of Court.

**SEC. 9. Effects of Defying Orders under this Act.** – Any person who willfully refuses or fails to comply with a valid Correction Order, Stop Communication Order, or Takedown

Order issued under this Act shall be criminally liable. The CDDTO shall, *motu proprio*, file the case before the the Department of Justice (DOJ).

**SEC. 10. *Unlawful Acts.*** — It shall be unlawful for any person to knowingly create, publish, share, or cause the dissemination of any online falsehood or disinformation that results in, or is likely to result in, serious and verifiable harm to public interest, including but not limited to:

- a. Causing to disrupt or disrupting public order or causing panic;
- b. Causing to undermining or undermining confidence in electoral processes;
- c. Causing to damage or damaging the reputation of individuals or institutions through fabricated content;
- d. Artificial amplification of falsehoods using troll farms, bots, or fake accounts;
- e. Causing to interfere or interfering with legitimate public discourse through coordinated inauthentic behavior.

**SEC. 11. *Penalties.*** – Any person who committed any of the unlawful acts in the preceding Section shall be initially subjected to the issuance of Correction, Stop Communication or Takedown Orders issued by the CDDTO of the DICT. Such person shall likewise be liable for an administrative fine of up to Five Hundred Thousand Pesos (P500,000.00) for every violation after due notice and hearing: *Provided*, That if the such person continues to defy lawful issued orders, or if the offense involves intent to deceive, recidivism, or committed any of the aggravating circumstances in this Section, such person shall suffer an imprisonment of one (1) year to five (5) years and a fine of not less than P100,000.00 but not more than P1,000,000.00: *Provided, further*, That the maximum penalty shall be imposed when the act is committed:

- (a) during national or local election periods;
- (b) through the use of troll farms, bots, fake accounts, or other forms of coordinated inauthentic behavior;
- (c) by a public officer who abuses their authority; or
- (d) with the intent to incite panic, disorder, violence, or undermine public trust in democratic institutions.

*Provided, however*, That when the act is knowingly committed on behalf of, or for the benefit of, a juridical person by a natural person holding a leading position by virtue of:

- (a) a power of representation for the juridical person, provided the act falls within such authority;
- (b) authority to make decisions on behalf of the juridical person; or
- (c) authority to exercise control within the juridical person.

Such juridical person shall be held liable for a fine equivalent to double the applicable fines under this Section but shall not exceed Two Million Pesos (P2,000,000.00). If the act was made possible due to the lack of supervision or control by a person in such position, and the act was committed for the benefit of the juridical person by someone under its authority, the juridical person shall be liable for a fine not exceeding One Million Pesos (P1,000,000.00): *Provided finally*, That any person who directly engages, authorizes, or knowingly benefits from the services of a juridical entity that commits unlawful acts shall likewise be held liable in accordance with the provisions of this Act. The penalties under this Act shall be without prejudice to other criminal liability that may be imposed by other applicable criminal laws.

**SEC. 12. *Public Reporting Mechanism.*** —The CDDTO shall establish and maintain an accessible, secure, and user-friendly digital platform for individuals, civil society groups, academic institutions, and other stakeholders to report suspected online falsehoods and disinformation.

The platform shall allow for the submission of relevant content, sources, and context, and shall include safeguards against abuse or misuse. All verified reports shall be reviewed in a timely manner for possible issuance of enforcement orders under this Act.

**SEC. 13. *Public Education and Digital Literacy.*** – The DICT, in coordination with the Department of Education (DepEd), the Philippine Information Agency (PIA), and the CDDTO, shall implement a nationwide public education programs to:

- a. raise awareness on online disinformation and its consequences;
- b. promote critical media consumption and fact-checking skills;
- c. encourage responsible digital citizenship and ethics; and
- d. integrate digital literacy into formal and informal education systems.

**SEC. 14. *Creation and Use of the Disinformation Control Fund.*** –There is hereby created a Disinformation Control Fund, to be administered by the DICT, sourced from administrative fines collected under this Act.

The Fund shall be used exclusively for the implementation of this Act, including but not limited to:

- a. Operational support for the Cyber Disinformation and Digital Truth Office (CDDTO);
- b. Monitoring and verification of online falsehoods and disinformation;
- c. Public education and digital literacy programs;
- d. Capacity-building activities, including training of enforcement personnel; and
- e. Development and maintenance of reporting and transparency platforms.

Disbursement from the Fund shall be subject to existing government accounting, auditing, and budgeting rules and regulations, and may require coordination and approval from the Department of Budget and Management (DBM), as appropriate.

All administrative fines collected under this Act shall be deposited into the Disinformation Control Fund and shall be exempt from reversion to the General Fund unless otherwise repealed by subsequent legislation.

**SEC. 15. *Appropriations.*** – The amount necessary for the initial implementation of this Act shall be charged against the current year’s appropriations of the DICT. Thereafter, such sums as may be necessary for its continued implementation shall be included in the annual General Appropriations Act (GAA).

In addition to the regular appropriations, the Disinformation Control Fund created under this Act shall serve as a supplemental source of funding for the operational and enforcement activities of the CDDTO.

**SEC. 16. *Implementing Rules and Regulations.*** —Within sixty (60) days from the effectivity of this Act, the DICT, in consultation with the Department of Justice (DOJ), the Department of Education (DepEd), the Philippine Information Agency (PIA), media organizations, civil society groups, and other relevant stakeholders, shall jointly promulgate the necessary rules and regulations for the effective implementation of this act.

**SEC. 17. *Separability Clause.*** —If any provision of this Act is declared unconstitutional or invalid, the other provisions not affected shall remain in full force and effect.

**SEC. 18. *Repealing Clause.*** — All laws, decrees, executive orders, rules, and regulations or parts thereof inconsistent with the provisions of this Act are hereby repealed or modified accordingly.

**SEC. 19. *Effectivity.*** — This Act shall take effect fifteen (15) days after its publication in the Official Gazette or in a newspaper of general circulation.

*Approved,*