



**2025 is Indonesia's
year of AI-ing
dangerously**





2025 is Indonesia's year of AI-ing dangerously

Authors

Siti R.A. Desyana
Marina Nasution

Layouters

Docallisme Studio

Published in December 2025



Creative Commons Attribution
NonCommercial-ShareAlike 4.0 International
License



Executive Summary



Indonesia is currently facing [mounting pressure](#) to promptly adopt AI regulations and systems from major global powers, including the European Union, the United States, and China. In 2020, Indonesia launched the [Indonesian National Strategy on Artificial Intelligence](#) (Stranas KA). Subsequently, according to Circular Letter Number 9 of 2023, the Ministry of Communication and Informatics (Kemenkominfo), now Kemkomdigi, is responsible for establishing ethical guidelines for artificial intelligence. Hence, the [formulation](#) of the National Task Force on AI Roadmap and the drafting of its [presidential regulation](#).

Local companies have also [grown to a regional level](#) and have historically played a crucial role in developing and deploying these technologies. In the current emerging tech race, they, alongside the government, have [adapted AI systems](#) into their everyday workflow and frameworks, particularly those that concern public interests. These developments, however, have made way for new emerging tech-related unprecedented incidents, which then cause scholars and policymakers to struggle in categorising and regulating the new technologies and, in turn, create regulations that may not match the real-world needs and circumstances surrounding the technologies.

Thus, we believe that an evidence-based approach is crucial to properly regulate AI, and that a **Media Monitoring Repository on AI Incidents in Indonesia is the best-suited tool** to understand the nuances, depth, and variety of harms caused by AI today.

This monitoring project documents incidents that arise from the alleged direct or indirect involvement of AI (hereinafter referred to as “alleged incident”) which have been reported to the public domain, whether by the victim of the incident themselves or a reliable source of publication. We focus on collecting incidents that are allegedly caused and/or contributed to by the development, deployment, or application of AI that are publicly available on the internet, such

as those printed in online media or reported by individuals on social media. This method allows us to authentically address the need to document public grievances. Following up on our previous [Monitoring Report](#), which highlighted findings of alleged AI-related incidents from 2022 to 2024, this iteration addresses some shortcomings in the categorisation and focuses on dissecting the findings from alleged AI-related incidents in 2025.

We face challenges in fully documenting alleged AI-related incidents due to three main factors. Firstly, **case finding/detection** are difficult because AI taxonomy varies across the public, making it challenging to identify incidents using specific keywords. Investigation requires context and knowledge of communal languages and terminology. Secondly, **technology grouping** is challenging because of the lack of transparency from both the private and public sectors in disclosing which tools are implemented and used across various services and applications, hindering accurate identification of the technologies involved. Lastly, **affiliation mapping** is complicated because victims are usually only aware of the final harmful product and lack knowledge of the machines' identities and their proxy companies, making it difficult to construct a complete picture of the situation.

With these limitations in mind, below is our analysis of the overall trend in incident documentation.



Findings

The monitoring project documented a significant increase in AI-related incidents: 8 cases occurred before the issuance of the AI Ethics Circular Letter, skyrocketing to 96 after the Circular Letter's implementation, and reaching 74 cases in 2025 alone.

There have been some successful remedies in the documented cases. For example:

- The Business Competition Supervisory Commission (KPPU) and the Central Jakarta District Court made [a decision](#) about the Google Play billing system.
- The government issued a [suspension order](#) against WorldApp for their unethical collection of individual biometric data.

- Two independent investigations are currently underway (the [SNPMB committees](#) are looking into the AI-powered cheating incident).
- Three cases in which the perpetrators were caught by the police (all related to deepfake fraud involving public figures' faces).
- Two incidents have been reported to authorities (deepfakes involving [dr. Tony Setiabudi](#) & [Mahfud MD](#))
- One incident reached a non-litigative resolution (Udayana [suspended](#) its student for procurement and distribution of NCI photos).

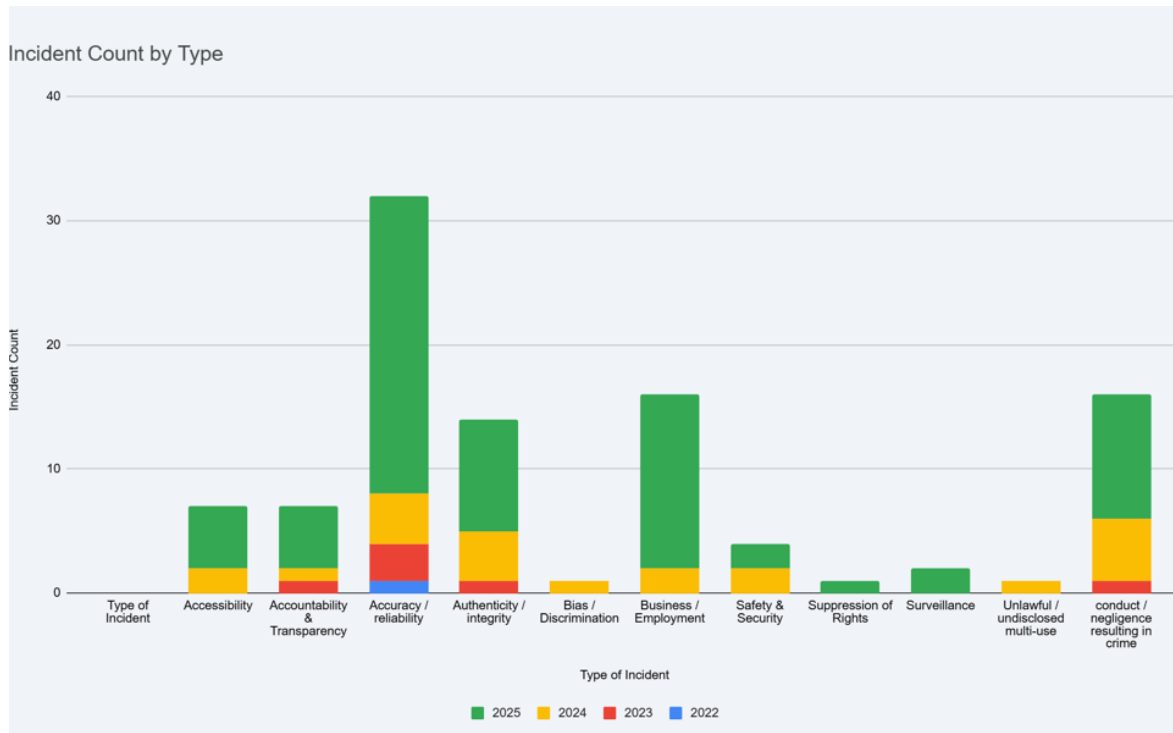
This evidence demonstrates the possibility of utilising both legal and extralegal remedies for AI-related incidents and underscores the need for greater efforts to apply existing regulations to address these emerging issues.

Numbers don't lie: Statistics of the year's findings

How many incidents occurred this year?

The analysis of incident counts across the four years (2022-2025) reveals a substantial and accelerating increase in reported issues related to emerging technology, particularly in 2025. The most significant concern is **accuracy/reliability**, which jumped from 4 incidents in 2024 to 24 in 2025, indicating that the fundamental functionality and trustworthiness of these platforms are becoming a major point of failure.

Other areas showing marked growth include **business/employment** (15 incidents in 2025) and **conduct/negligence resulting in crime** (10 incidents in 2025), suggesting that harms are increasingly shifting from technical flaws to real-world social and economic impacts. The rise in **accountability & transparency** incidents (5 in 2025) also points to growing frustrations regarding the opacity of platform operations.



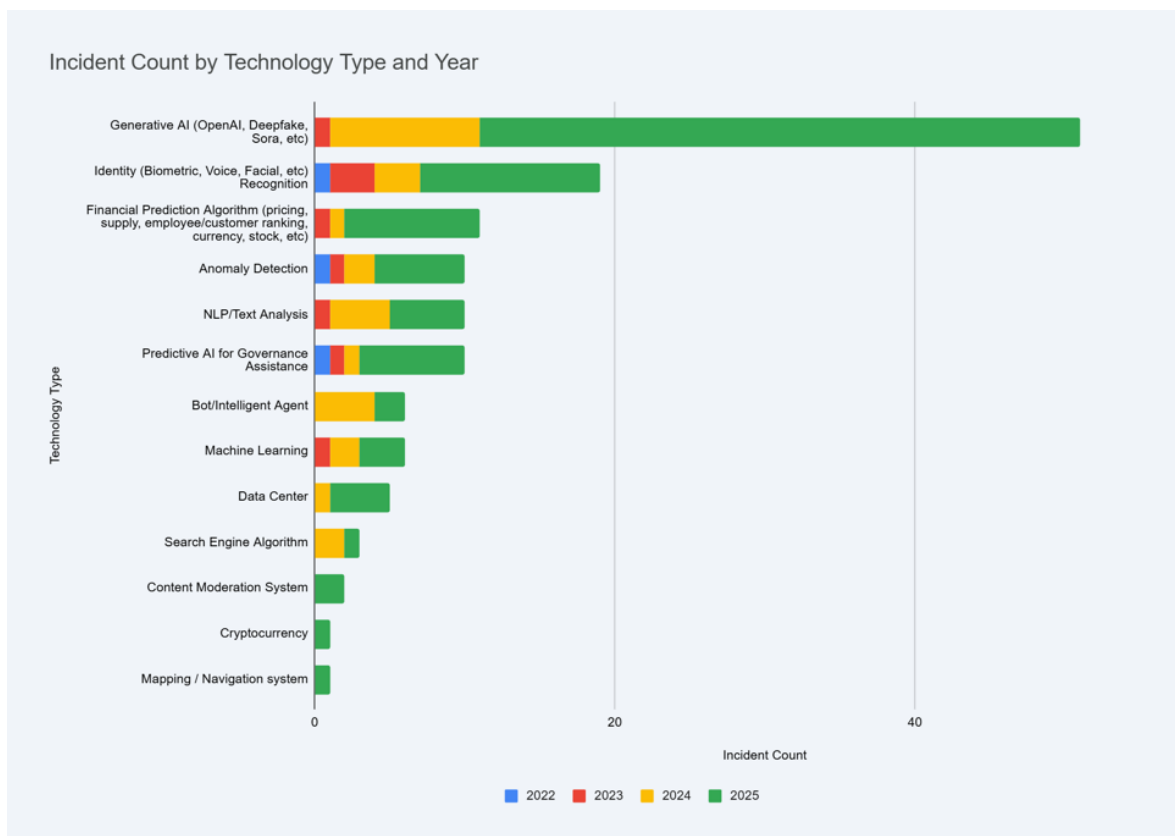
| Type of incident | 2022 | 2023 | 2024 | 2025 |
|---------------------------------------|------|------|------|------|
| Accessibility | 0 | 0 | 2 | 5 |
| Accountability & Transparency | 0 | 1 | 1 | 5 |
| Accuracy/Reliability | 1 | 3 | 4 | 24 |
| Authenticity/Integrity | 0 | 1 | 4 | 10 |
| Bias/Discrimination | 0 | 0 | 1 | 0 |
| Business/Employment | 0 | 0 | 2 | 15 |
| Safety & Security | 0 | 0 | 2 | 2 |
| Suppression of Rights | 0 | 0 | 0 | 1 |
| Surveillance | 0 | 0 | 0 | 2 |
| Unlawful/Undisclosed multi-use | 0 | 0 | 1 | 0 |
| Conduct/Negligence resulting in crime | 0 | 1 | 5 | 10 |

Which technologies were involved in the incidents?

The data on the types of technology involved in incidents identifies **generative AI** (including large language models, deepfakes, etc.) as the most volatile and problematic emerging technology. Incidents involving generative AI skyrocketed from 10 cases in 2024 to an alarming 41 in 2025, underscoring the need for stronger governance and safety measures for this technology.

Other technologies with high involvement in the documented alleged incidents include **identity recognition** (12 incidents in 2025), reflecting rising issues with biometric and facial recognition systems, and **financial prediction algorithms** (10 incidents in 2025), which often impact economic decision-making and fairness.

The continued involvement of **anomaly detection** systems (6 incidents in 2025) suggests persistent problems with automated monitoring and flagging systems. Lastly, cases involving **data centres**, including governmental database leaks and performance failures, as well as excessive energy consumption, show a rise in 2024-2025 (6 cases total, 4 in 2025 alone).

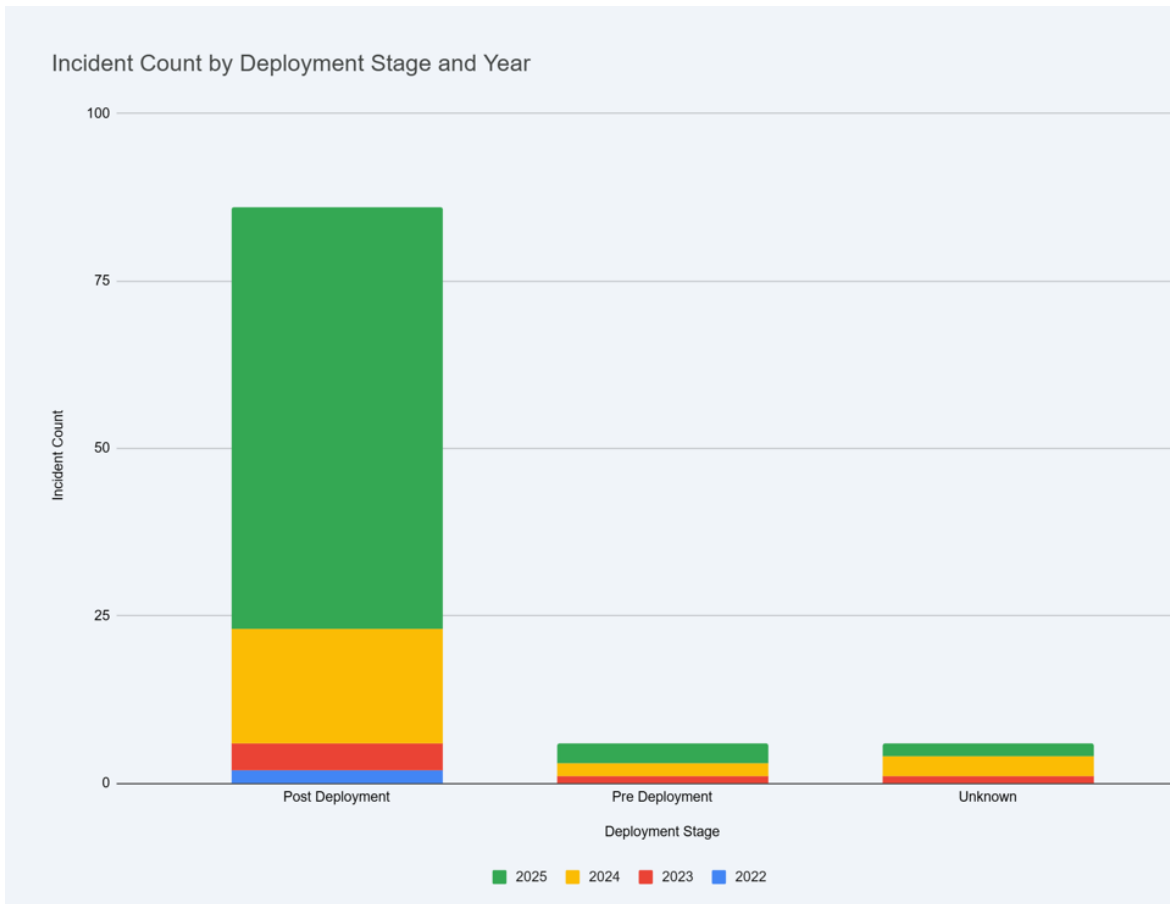


| Type of technology involved | 2022 | 2023 | 2024 | 2025 |
|---|------|------|------|------|
| Machine Learning | 0 | 1 | 2 | 3 |
| Identity (Biometric, Voice, Facial, etc) Recognition | 1 | 3 | 3 | 12 |
| Financial Prediction Algorithm (pricing, supply, employee/customer ranking, currency, stock, etc) | 0 | 1 | 1 | 10 |
| Speech Recognition | 0 | 0 | 0 | 0 |
| Search Engine Algorithm | 0 | 0 | 2 | 1 |
| Content Moderation System | 0 | 0 | 0 | 3 |
| Autonomous Driving/Flight System | 0 | 0 | 0 | 0 |
| NLP/Text Analysis | 0 | 1 | 4 | 5 |
| Generative AI (OpenAI, Deepfake, Sora, etc) | 0 | 1 | 10 | 41 |
| Bot/Intelligent Agent | 0 | 0 | 4 | 2 |
| Anomaly Detection | 1 | 1 | 2 | 6 |
| Robotics | 0 | 0 | 0 | 0 |
| Cryptocurrency | 0 | 0 | 0 | 1 |
| Mapping/Navigation system | 0 | 0 | 0 | 1 |
| Predictive AI for Governance Assistance | 1 | 1 | 1 | 7 |
| Data Centre | 0 | 1 | 1 | 4 |

When are the incidents most likely to harm the public?

The timeline data overwhelmingly indicate that the vast majority of documented harms occur after the technologies' public deployment. Post-deployment incidents accounted for 67 of the 74 total recorded incidents in 2025, an increase from 17 in 2024. This pattern strongly suggests that current pre-deployment testing, risk assessments, and ethical review processes are fundamentally inadequate for catching real-world vulnerabilities and negative externalities.

The relatively low numbers for pre-deployment (5 in 2025) and unknown (2 in 2025) incidents reinforce the conclusion that accountability and mitigation efforts must focus heavily on continuous monitoring, rapid response, and robust redress mechanisms once a system is live.



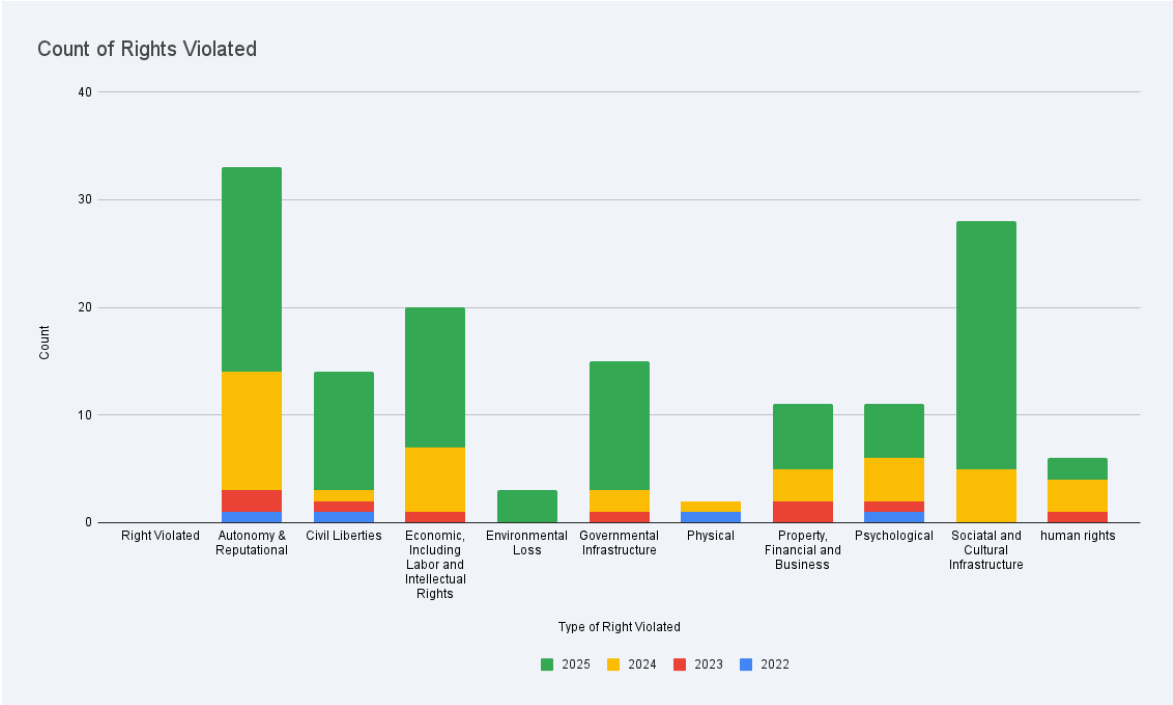
| Stage of the AI lifecycle when the incident occurred | 2022 | 2023 | 2024 | 2025 |
|--|------|------|------|------|
| Pre-deployment | 2 | 4 | 17 | 67 |
| Post-deployment | 0 | 1 | 2 | 5 |
| Unknown | 0 | 1 | 3 | 2 |

What harms are borne out of these incidents?

To determine which rights are being violated, we use classifications and definitions adapted from the [AIAAIC repository](#). The rights violation data paints a picture of harms shifting from individual issues to broader societal disruption. In 2025, the leading category was **societal and cultural infrastructure** with 23 violations, jumping significantly from 5 in 2024. This indicates that new technologies are increasingly disrupting collective systems, norms, and public spaces, especially in the degradation of information.

Violations of **autonomy & reputational** rights remained high (19 cases in 2025), reflecting continuing issues with disinformation and identity-

related harms. Violations related to **economics, including labour and intellectual rights** (15 cases in 2025), highlight the growing impact of algorithmic systems on employment, creative works, and financial fairness. Furthermore, the 11 documented violations of **civil liberties** demonstrate an escalating threat to fundamental freedoms.



| Type of right violated | 2022 | 2023 | 2024 | 2025 |
|---|------|------|------|------|
| Autonomy & Reputational | 1 | 2 | 11 | 20 |
| Civil Liberties | 1 | 1 | 1 | 11 |
| Economics, Including Labour and Intellectual Rights | 0 | 1 | 6 | 15 |
| Environmental Loss | 0 | 0 | 0 | 3 |
| Governmental Infrastructure | 0 | 1 | 2 | 12 |
| Physical | 1 | 0 | 1 | 1 |
| Property, Financial and Business | 0 | 2 | 3 | 7 |
| Psychological | 1 | 1 | 4 | 5 |
| Societal and Cultural Infrastructure | 0 | 0 | 5 | 23 |
| Human Rights | 0 | 1 | 3 | 2 |

How severe is the harm?

Meanwhile, the scale of harm from incidents involving this emerging technology has worsened drastically, with an increase in the number of documented incidents and the emergence of new incidents categorised as serious and even catastrophic. We categorise the severity of incident impacts using OECD groupings, combined with elements of the [MIT AI Incident Tracker](#) scale, which we adapt to suit the Indonesian context. Here is the scale we use:

| Classification | | AI Incident | Serious AI Incident | AI Disaster |
|---|------------------------------------|--|--|---|
| Individual and/or community health (physical and psychological) | | Minor to moderate injuries/illnesses/disorders, no fatalities. | Death or serious injury, including severe injuries, mental disorders, and/or chronic illnesses. | Small-scale health crises (1-99) to mass casualties (national level) resulting in large-scale loss of life and/or mass health disruption. |
| Disruption to critical infrastructure | Governmental infrastructure | Brief, localised, and recoverable disruption to critical infrastructure. | Serious local to regional disruptions or breaches of critical infrastructure that compromise or cause service failure. | Widespread disruptions across multiple regions to the national level that are irreversible and/or cause infrastructure collapse. |
| | Societal & cultural infrastructure | Slightly to moderately offensive/misleading content; minor information muddle; visible bias. | Harmful disinformation and harassment; systematic removal of critical information; systematic discrimination affecting specific communities. | Institutionalised discrimination; targeted incitement to violence leading to mass radicalisation. |

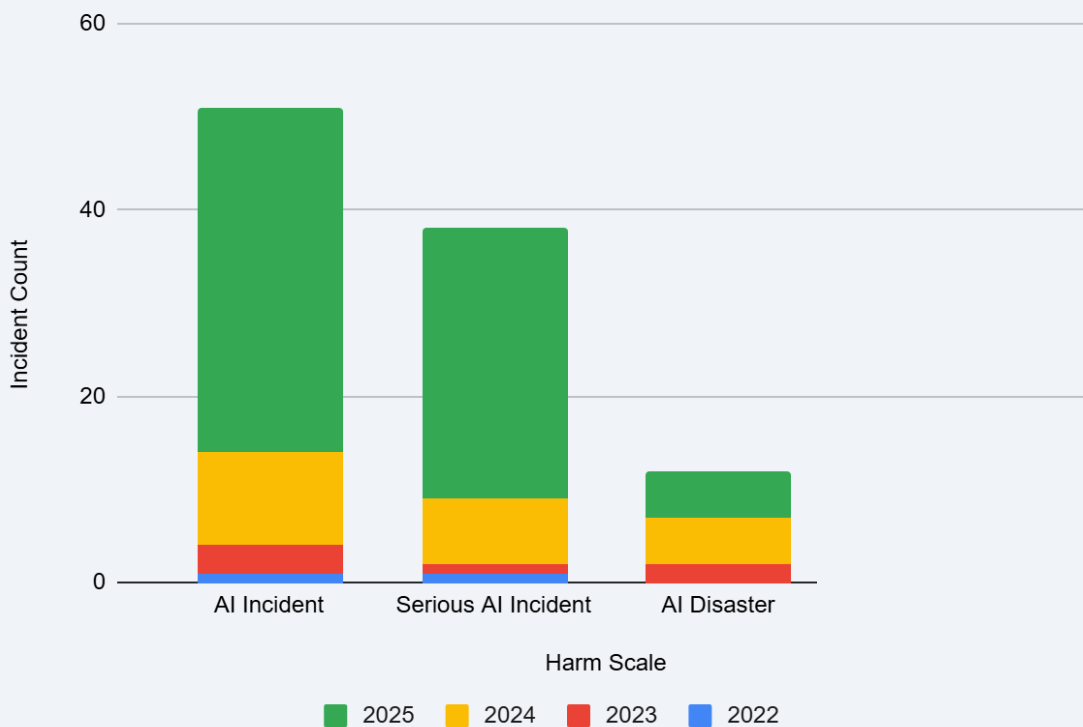
| Classification | | AI Incident | Serious AI Incident | AI Disaster |
|---|--|--|--|--|
| Violations of human rights, intellectual property rights, and workers' rights | Civil liberties | Minor violations in the implementation of the democratic process. | Systematic voter manipulation, automatic moderation of free speech, targeted attacks on government critics. | The collapse of democratic institutions and the takeover of power by technology-enabled authoritarianism. |
| | Human rights | Limited restrictions causing moderate inconvenience. | Significant restrictions affect thousands of people. | Widespread and systemic suppression of rights affecting communities at the multi-regional and/or national level. |
| | Economic, including labour and intellectual rights | Small to moderate material and/or financial losses disrupting individual livelihoods (IDR1 million-IDR50 million). | Moderate to significant property damage and/or financial losses causing an individual to lose the ability to meet basic needs (IDR50-500 million). | Damage extends to disasters or extreme direct financial loss (IDR500 million and above). |

| Classification | | AI Incident | Serious AI Incident | AI Disaster |
|--|----------------------------------|--|--|--|
| Damage to property, communities, and the environment | Property, financial and business | Small to moderate material and/or financial losses that disrupt company operations (<USD10,000). | Moderate to significant property damage and/or financial losses that require the business to cease operations (USD10,000-USD1 million). | Damage extends to catastrophic or extreme direct financial losses (USD1 million and above). |
| | Autonomy & reputational | Small to moderate basic data breaches/misuse that can be limited. | Serious privacy breaches and/or misuses involving identifiable data, affecting up to 1,000 individuals. | Societal-level privacy collapses with widespread surveillance. |
| | Environmental loss | Minor to moderate damage impacting local communities and is recoverable. | Long-term impacts include major to severe damage that is partially reversible and impacts large-scale communities (e.g., entire cities or provinces) | Widespread ecosystem collapse leading to irreversible ecological destruction (e.g., species extinction, loss of water resources in the area, etc.) |

The increase in the incident count has also contributed to the growing variation in the scale of harm caused. Our analysis of the documented alleged incidents shows a rise in the classification of cases with higher impact severity. The total number of recorded incidents was 74 in 2025. While cases categorised as standard AI incidents increased significantly to 38 (up from 10 in 2024), the most concerning trend is the leap in cases categorised as serious AI incidents, which quadrupled from 7 in 2024 to 30 in 2025. This suggests that the impact of emerging tech failures is becoming more profound, systemic, and difficult to mitigate.

The number of alleged incidents classified as AI disasters increased from 5 to 6 counts, indicating that while catastrophic failures are less frequent than serious ones, they still pose a persistent, high-risk threat. This data strongly suggests that there must be robust risk management and preemptive safety measures.

Incident Count by Harm Scale and Year (Excluding AI Hazard)



| Harm scale | 2022 | 2023 | 2024 | 2025 |
|---------------------|------|------|------|------|
| AI Incident | 1 | 3 | 10 | 38 |
| Serious AI Incident | 1 | 1 | 7 | 30 |
| AI Disaster | 0 | 2 | 5 | 6 |

Who can be held accountable?

For our 2025 findings, we recorded the involvement of various entities in the alleged incidents documented and grouped them into two major classifications. In the **private sector** classification, the big tech company [Alphabet](#)—which owns the Google ecosystem, including its Workplace feature, its LLM and generative AI tool Gemini, Veo, and Nano Banana, as well as YouTube—leads the count with 7 alleged incidents. The rank is followed up by TikTok (5, mostly as enablers of deepfake and AI-powered disinformation for dissemination); Meta and Twitter (2); Grab and Gojek (2 each, for the same cases as they employ similar techniques); Tokopedia (2); and Shopee (2).

Among **public services**, the Indonesian National Police had the most involvement in the alleged incidents (6, all of which are related to E-TLE implementation), followed by the Directorate General of Taxation (2, relating to Coretax implementation) and Civil Registry Service Office (1, in the case of the Digital ID App).

| Identified entities involved in the alleged incidents | 2025 |
|---|------|
| Perusahaan Swasta | |
| Google | 7 |
| Meta | 6 |
| TikTok | 5 |
| Twitter | 2 |
| Grab | 2 |
| Gojek | 2 |
| Tokopedia | 2 |
| Shopee | 2 |
| Layanan Publik | |
| Indonesian National Police | 6 |
| Directorate General of Taxation (DJP) | 2 |
| Civil Registry Service Office (Dispendukcapil) | 1 |

They keep popping up: highlighted themes in the documented incidents

Deepfakes for illegal activities

used for non-consensual intimate imagery (NCII)

TW: This content discusses non-consensual intimate imagery (NCII) and technology-facilitated sexual abuse, which may disturb or trigger some readers/viewers. Please read carefully and seek professional help if required.

The monitoring project documented 8 NCII cases involving the use of deepfakes and generative AI between 2022 and 2024. We categorised the incidents into three major groups:

1. The **business of mass-produced NCII** categories includes the case of a college student under investigation [for creating and selling approximately 4000 NCII images](#) generated from stolen photos.
2. **Individually targeted NCII** cases feature the use of AI-generated nudes as a threat in [retaliation](#) for online loans and by an [anonymous account](#) to harass a woman cosplayer, as well as cases [targeting minors](#).
3. **NCII of famous individuals** includes the non-consensual use of male athletes' likenesses for the [polaroid trend](#) and a viral TikTok account generating [sensual AI-edited images](#) of various hijabi and queer-presenting influencers and celebrities, including children.

Indonesian law addresses technology-facilitated sexual violence and indecent content through several regulations. **Law No. 12/2022 on Sexual Violence Crimes**, specifically Article 14(1)(b), penalises anyone who, without authorisation, transmits electronic information or documents with sexual content against the recipient's will and directed at sexual desire, with a maximum imprisonment of 4 years and/or a fine of up to Rp200,000,000.

Additionally, **Law No. 1/2024 on Electronic Information and Transactions**, Article 27(1) in conjunction with Article 45(1), punishes anyone who knowingly and without authorisation broadcasts, exhibits, distributes, transmits, and/or makes accessible electronic information or documents containing indecent content for public knowledge with a maximum imprisonment of 6 years and/or a maximum fine of Rp1 billion.

These regulations aim to cover various forms of electronic-based sexual offences and the spread of inappropriate content. However, law enforcement has yet to successfully implement these existing frameworks.

Used for fraud & online gambling

The monitoring effort discovered several instances of the use of deepfakes and generative AI to promote illegal activities, such as fraud and online gambling. In a notable case, police arrested a conman who [used deepfaked videos of high-profile political figures](#), including President Prabowo Subianto and Minister of Finance Sri Mulyani, to fraudulently advertise governmental financial assistance and defraud billions of rupiah. Furthermore, deepfake videos manipulating Indonesian celebrities and prominent journalists, such as Najwa Shihab and Raffi Ahmad, have surfaced on social media to [promote online gambling sites](#), where original video content about vaccines was deceptively edited to endorse illegal slot sites.

Law No. 1 of 2024 concerning Electronic Information and Transactions (ITE) is a common regulation used by law enforcement officials (APH) when dealing with online content. As the primary regulation governing misleading information that causes consumer harm and gambling content, it can be commonly assumed that, should the authorities choose to prosecute these cases, these instruments will be implemented to address the above cases.

Beyond this regulation, there is also **Law No. 1 of 2023 concerning the New Criminal Code** (which addresses the spread of false information that impacts commodity prices) and **Government Regulation No. 71 of 2019**, which requires the government and Electronic System Organisers (PSEs) to prevent the spread of illegal content, including through access restrictions. **Minister of Communication and Information Technology Regulation No. 5 of 2020** further requires private PSEs to ensure their systems do not contain prohibited content.

Therefore, in handling cases of incitement/fraud and gambling through deepfakes, law enforcement agencies should direct accountability demands toward platform managers who “facilitate” the distribution of related content, in accordance with the mandate of the relevant regulations. However, in the following subsection, it is seen that this law is instead used not to suppress the aforementioned cases but rather for content [moderation](#), content [removal](#), and even the [criminalisation](#) of several activists.

AI-generated content facilitates information muddle

The monitoring effort found extensive use of generative AI, including deepfakes and LLMs, in the dissemination of disinformation. We categorise it into 2 major groups:

- 1. The use of AI for disinformation in critical conditions**, such as during the 2024 Indonesian election, saw instances like the Prabowo-Gibran campaign [utilising AI](#) for image rebranding to downplay Prabowo's past human rights violations and a debunked, [suspected AI-generated audio clip](#) of then-presidential candidate Anies Baswedan being scolded. Beyond politics, AI-generated hoaxes included a [non-factual image of mass deforestation](#) in Raja Ampat, [a smear campaign video targeting an activist](#) from the "Reset Indonesia 2025" protest, and a fake video of a [CNN journalist's arrest](#) following the Sumatra Flood reporting.
- 2. Day-to-day information muddle**, involved a [fake video](#) of Minister Sri Mulyani announcing a fraudulent government investment programme and a specially appointed staff member of the Ministry for Digital Information and Communication [disseminating an incorrect quote from a law](#), likely sourced via AI, highlighting the pervasive impact of AI misinformation on public trust and information accuracy.

Law enforcement officials (APH) frequently employ Indonesia's current regulatory framework to "combat misinformation", which includes disinformation and information muddle. For example, in the case of [Daniel Tangkilisan](#), or the Bandung Institute of Technology (ITB) student arrested for [generating an image using AI](#) depicting figures resembling Prabowo Subianto and Joko Widodo kissing, APH implemented **the Electronic Information and Transactions (ITE) Law**, namely Law No. 11 of 2008, amended by Law No. 19 of 2016, and most recently revised by Law No. 1 of 2024.

With the enactment of **Law No. 1 of 2023 concerning the Criminal Code (New Criminal Code)**, it would not be surprising if APH implemented Article 263(2), which prohibits the dissemination of false news that can cause public unrest, as well as the dissemination of uncertain, exaggerated, or incomplete news known to cause public unrest (Article 264). The ITE Law sanctions the deliberate

dissemination of information/electronic documents containing false notices that cause public unrest (Article 28(3)).

Human rights defenders often face intimidation due to this lax regulation. The “defamation” clause is commonly applied in criminal cases, especially when criticising government policies, as seen in the aforementioned cases, which contradicts the intended spirit of this regulation. According to international law, as outlined in the [Rabat Plan](#), criminal provisions on freedom of expression can only be used as a last resort, when the expression in question is hate speech that leads to real-world crimes, such as the [violence against Rohingya refugees](#) in Aceh following [the rise of hate speech](#) and [disinformation](#) on social media targeting them.

The government must act immediately when AI-enabled disinformation, both in its creation and transmission, spreads rapidly and causes real-world unrest due to a poor content moderation system and algorithmic amplification that ignores societal norms. One legal framework for platform accountability that can be used and updated is the **Minister of Communication and Informatics Regulation (Permenkominfo) Number 5 of 2020 concerning Private Electronic System Organisers (Private PSE Law)**. Given the significance of this regulation in ensuring the fulfilment of human rights and the public interest, it should be included in legislation, and its implementation should ensure transparent and accountable legal procedures while prioritising victim protection and the public interest.

Corporate Use

We recorded several instances of AI incidents involving big tech platforms in Indonesia, which focus on the accountability of their use of emerging tech in decision-making. It highlights how **algorithmically driven gamification systems** can [burden](#) e-hailing drivers with smaller compensation and longer hours, forcing them to pay for subscriptions to access more customers, while the system takes a significant cut of the fare. It also shows how such systems can [unfairly penalise users and merchants](#) with low ratings on Gojek without providing transparency to the rating system.

Furthermore, the use of **automated decision-making** on platforms like Shopee is criticised for leading to [automatic suspensions](#) of live hosts based on seemingly trivial reasons (e.g., being quiet or sitting down) with no clear right to appeal, suggesting monitoring by bots. Lastly, the issue of **biometric use**, such as the [incident of large-scale iris data collection](#) by WorldApp with the lure of financial rewards, is considered unethical because the consent of the data subjects was obtained through deception.

We also document cases of **misinformation from embedded generative AI services** in big tech companies, some of which resulted in privacy breaches. For example, Google Gemini [incorrectly](#) attributed a company's customer service number to a JKT48 member's personal number, leading to spam and harassment for the company and prompting it to issue a public disclaimer. Similarly, Meta AI's Llama 4 model has [falsely attributed](#) personal phone numbers to customer service lines of various telecommunication companies (Telkomsel, Axis, Indihome, etc.), resulting in a data breach for the individuals involved. Victims of these incidents currently lack effective recourse mechanisms for these privacy violations.

The existing regulatory framework includes **Law No. 27 of 2022 concerning Personal Data Protection (PDP)**, Article 10(1), which grants data subjects the right to object to decisions based solely on automatic processing, including profiling, that have legal consequences or a significant impact on them. Additionally, the Regulation of the **Minister of Communication and Informatics regarding Private Electronic System Organisers (Private PSE Law)**, Article 9(1), holds these operators accountable for the reliable, secure, and responsible operation of electronic systems and the management of electronic information and/or electronic documents within them.

However, significant regulatory gaps persist: e-hailing drivers are [trapped in partnerships with companies](#), thereby depriving them of their rights as workers, while their burden of responsibilities is heavier than that of partners. This partnership deliberately ignores the provisions of **Law No. 13 of 2003 concerning manpower**, which appear to provide companies with a loophole to evade their obligations

to provide drivers with basic protections, such as minimum wages, social security, annual and sick leave, and humane working hours. Weak enforcement of existing laws, including Law No. 39 concerning human rights, and the lack of regulations on transparency in the ranking system will only slow down the fulfilment of basic rights and livelihoods for these drivers (gig workers).

Tech-powered mass surveillance

The monitoring effort identified at least two incidents of large-scale surveillance in Indonesia. First, there are allegations of the Indonesian government [conducting mass surveillance](#) during critical political periods, potentially using tools like IMSI catchers or Cellebrite UEFD. During the #DarkIndonesia March 2025 protest, [participants and bystanders](#) near the protest area reported being logged out of or having their social media and WhatsApp accounts suspended, a pattern also observed in past protests against the Penal Code revision. Second, an Indonesian company in Jakarta, First Wap, has been implicated in the global spyware business, deploying surveillance technology to [track activists and other high-profile individuals](#), including Italian journalist Gianluigi Nuzzi. Investigation revealed that First Wap [had preyed on approximately 14,000 phone numbers](#) through its surveillance practices.

Indonesia's regulatory framework offers several layers of protection for privacy and communication rights, providing a foundation for platform accountability. **The 1945 Constitution** (Article 28G(1)) affirms the right to personal protection and security against threats. The right is more thoroughly elaborated in **Law No. 39 of 1999 on Human Rights**, which guarantees the right to safeguard oneself, family, and reputation (Article 29(1)) and protects the freedom and confidentiality of communications, allowing interference only upon a judicial order (Article 32).

Furthermore, the **Personal Data Protection Law (Law No. 27 of 2022)** grants data subjects rights such as the right to information, the right to revoke consent, and the right to object to automated decision-making (Articles 5-11), while establishing criminal consequences for unauthorised data acts. Lastly, the **Telecommunications Law (Law No. 36 of 1999)** generally prohibits the unauthorised interception of information (Article 40).

Despite these protections, a significant regulatory gap exists, as Indonesia lacks sufficient specific regulation concerning mass spyware, which often allows such activities to operate in “grey areas”.

Data centres and their violations of environmental rights

Our monitoring effort identified at least two incidents that pose **a risk of excessive water and electricity consumption**. The [rapid growth of data centres](#), particularly in Indonesia’s hotter regions, presents significant risks regarding excessive water and electricity consumption. High temperatures increase the burden on cooling systems and reduce power efficiency, contributing to 1.5% of the global electricity consumed by data centres in 2024. A critical local impact is seen in Batam, where the construction of 18 new data centres has been linked to increased droughts and water supply shortages for citizens. By 2032, it is estimated that the Batam Industrial Complex’s data centres will [consume about 8% of the island’s total water supply](#), raising concerns about resource allocation and climate resilience.

For individual analysis of all the alleged incidents, please check the breakdown in the next section.

