

SECURING CHILDHOOD

Beyond the Age Limit: A Safety-by-Design Framework for Protecting Filipino
Children Online

By Albert Pangilinan Dela Cruz
Cybersecurity Practitioner and Digital Governance Advocate

October 2026

Version 5.2, 9 October 2026. Companion document: Securing Childhood, 12-Month Action Roadmap (v2.2).

Status

Status as of 9 October 2026: the Senate has approved SB 2424 (the SMART KIDS Act). The House has not yet passed its version, and the bill is not yet law. Facts are as reported up to that date.

About this paper

This paper was written independently by the author in his personal capacity as a private citizen. It is not produced on behalf of, funded by, or directed by any company, employer, client, agency or organization. The author also works as a cybersecurity consultant and has business interests in technology. This paper was prepared independently and was not commissioned or funded by any client or company.

A note on evidence

Three school shootings in 2026 have produced strong claims in public debate. This paper states each claim only as strongly as the evidence allows. Where officials have linked an attack to online activity, the paper reports what they said and notes what has not been established. Appendix B sets out the wording the author considers safe to use.

Version history

Version	Date	Summary of changes
1.0	27 June 2026	First release after the Tacloban City shooting
2.0	21 August 2026	Updated after the Ateneo de Zamboanga shooting
3.0	22 September 2026	Updated after the Banga shooting; citation corrections
5.0	26 September 2026	Repositioned around making the coming law work; PhilSys safeguards; ten amendments
5.1	8 October 2026	Lessons from the September Discord block; new references
5.2	9 October 2026	Senate approval of SB 2424 (under-18 rule for high-risk platforms); parent penalties; retargeted at the bicameral conference; source links

Abstract

Between June and September 2026, three fatal school shootings in the Philippines drew urgent attention to how children encounter violent content, harmful communities and predatory contact online. Government has responded quickly: platforms have been blocked or pressed into commitments. On 8 October 2026 the Senate approved a bill barring anyone under 18 from "high-risk" social media and online game platforms. A House committee has approved a different version, and the two must now be reconciled.

This paper argues that an age limit is a reasonable floor but not a complete system. On its own, it leaves children on the services it does not cover with adult-grade settings, does not reach exempt messaging apps and games, can create new privacy and security risks through identity checks that every adult user must also pass, and depends on enforcement by ultimatum. The September 2026 Discord block, which cut off legitimate users on some networks but not others and was followed by a rise in VPN use, shows the cost of that approach.

The paper proposes a five-part safety-by-design framework to make the coming law work: safe defaults for every minor, privacy-safe age checks and device setup, safer networks, capable adults with a clear path to early help, and fair, fast enforcement with public accountability. It closes with ten amendments for the bicameral conference and a phased set of actions for government, Congress, industry, schools and families. The approach draws on UNICEF and OECD guidance and on regulation in the European Union, the United Kingdom, Australia, Brazil and Singapore.

Executive Summary

The Philippines is right to set age limits for social media. But an age limit decides only who may enter. It does nothing about what happens to the teenagers who are allowed in, the children who get past it, or the harms that take place on games and messaging apps outside its scope. Congress, DICT and the enforcement agencies have acted quickly and in good faith. The task now is to turn urgent measures into a system that lasts.

What government has already done

DICT and CICC blocked GoreBox and secured commitments from Roblox (National ID age checks and a local office) and Meta (a joint working group and threat-data sharing). In September, CICC gave Reddit and Discord 24 hours to name local representatives; Discord was blocked for about a day and restored after it made safety commitments. The Senate approved SB 2424 on 8 October by 16 votes to 1, barring under-18s from high-risk platforms, and a House committee has approved a consolidated substitute bill. DepEd, DILG and PNP are reviewing school security.

What is still missing

Gap	Why it matters
No safety rules on the services children can still use	The Senate version exempts private messaging, games not mainly used for social networking and other services. The House version allows 13 to 17 year-olds on. Either way, children remain on services with adult-grade defaults unless the law says otherwise.
Age checks that could create new risks	An under-18 rule means every user of a high-risk platform, adults included, must prove their age. A single National ID check used for this becomes a target for attackers and can exclude people who are not registered.
Enforcement by ultimatum	Whole-platform blocks act quickly but fall unevenly across networks, cut off legitimate users, push people to VPNs and invite legal challenge.
No path from warning sign to help	Group-chat warnings and known signs of distress did not reach anyone able to intervene.
No baseline to measure success	Without a national survey and published data, no one will know whether the law works.

Five pillars to make the law work

Pillar	In one sentence	Main owners
1. Safe by Default	Every minor's account starts private and restricted, including teens the law allows on.	Platforms, DICT
2. Privacy-Safe Age Checks and Device Setup	Age is checked without building an identity trail, and new devices ask at setup whether a child will use them.	DICT, PSA, NPC, device makers, app stores
3. Safer Networks	Telcos and ISPs offer free, opt-in family-safe settings.	Telcos, ISPs, NTC
4. Capable Adults and Early Help	Parents, teachers and counselors know the settings, recognize warning signs and have a clear path to help.	DepEd, DOH, DSWD, LGUs

Pillar	In one sentence	Main owners
5. Fair, Fast Enforcement and Accountability	Published rules for restricting platforms, one door for reports, safeguards on monitoring, and public results.	DICT, CICC, NBI, PNP, DOJ, NPC

Priorities for the next 90 days

1. Publish a Platform Restriction Protocol so that any future block follows clear criteria, notice and time limits.
2. Write five privacy and security safeguards into the PhilSys age-check integrations before they go live.
3. Set safeguards on the monitoring access ("green lanes") that platforms have offered to authorities.
4. Table the ten amendments in Section 6.4 in House plenary debate and the bicameral conference.
5. Issue a school distress pathway so that warning signs reach a counselor, a health professional and the family.

Table of Contents

Abstract	3
Executive Summary	4
What government has already done	4
What is still missing	4
Five pillars to make the law work	4
Priorities for the next 90 days	5
1. Introduction	8
1.1 Why this matters now	8
1.2 Scope and approach	8
1.3 How to read this paper	8
2. Current State	9
2.1 What is known about the 2026 incidents	9
2.2 The government response so far	9
2.3 The September Discord block	10
2.4 Legislation	10
2.5 Four types of online risk	11
2.6 A complementary layer: firearm storage	12
3. Analysis	13
3.1 How harmful content escalates	13
3.2 Where the fixes go: four system layers	13
3.3 Why an age limit alone is not enough	13
3.4 Lessons from enforcement by ultimatum	14
3.5 Parents cannot close these gaps alone	14
4. Proposed Framework	16
4.1 Guiding principles	16
4.2 Pillar 1: Safe by Default	16
4.3 Pillar 2: Privacy-Safe Age Checks and Device Setup	17
4.4 Pillar 3: Safer Networks	17
4.5 Pillar 4: Capable Adults and Early Help	18
4.6 Pillar 5: Fair, Fast Enforcement and Accountability	18
5. Implications	19
5.1 Who leads each pillar	19
5.2 For government and regulators	19
5.3 For Congress	19

5.4 For platforms, device makers and telcos	19
5.5 For schools and families.....	20
5.6 Alignment with global standards	20
6. Recommendations	21
6.1 Short term: the first 90 days (October to December 2026)	21
6.2 Medium term: to September 2027	21
6.3 Long term: to June 2028 and beyond	21
6.4 Ten amendments to make the law work.....	21
7. Conclusion.....	23
About the Author	24
References	25
Appendix A: Glossary	27
Appendix B: What Can and Cannot Be Claimed	28
Appendix C: Quick-Start Checklists	29
Appendix D: Revision Note	30

1. Introduction

1.1 Why this matters now

Three fatal school shootings took place between June and September 2026: in Tacloban City (22 June), at Ateneo de Zamboanga University (18 August) and at Banga National High School in South Cotabato (18 September). In the Zamboanga case, the police investigation identified prior preparation and exposure to violent online material among the circumstances of the attack.[17] In the Banga case, officials reported that the student was active in online communities that glorify violence, had shown signs of distress that were known at home months earlier, and warned friends in a group chat before the attack.[18] Investigators have not established that any single platform caused any of the attacks.

The cases point to a pattern the current system is not built to catch. Harmful influences reach children through everyday apps, and warning signs that are visible to someone do not reach anyone able to act.

1.2 Scope and approach

This paper looks at the digital side of child safety: the devices, networks, platforms and people that shape what children see and who can reach them online. It treats firearm storage and school security as essential complementary measures but does not cover them in depth.

The paper takes a system view. It does not ask who is to blame. It asks where in the system a risk can be reduced, who is best placed to act, and how results can be measured. Recommendations are aligned with UNICEF and OECD guidance, the UN Committee on the Rights of the Child's General Comment 25, and current regulation in the European Union, the United Kingdom, the United States, Australia, Brazil and Singapore.

1.3 How to read this paper

Section 2 describes the current state. Section 3 analyses why an age limit on its own will not be enough. Section 4 sets out the framework, Section 5 its implications for each stakeholder, and Section 6 the recommendations and amendments. The companion Roadmap turns these into dated actions with suggested owners and targets.

2. Current State

2.1 What is known about the 2026 incidents

Incident	What officials reported	What has not been established
Tacloban City 22 June 2026	Two teenage shooters; three students killed and about 20 wounded. One suspect was reported to be a frequent player of the game GoreBox. The DOJ referred to "nihilistic violent extremism." [1]	That the game caused the attack.
Zamboanga City 18 August 2026	A student livestreamed the attack and used his father's pistols. The PNP's completed investigation found no conspiracy and no school or personal grievance, but identified prior preparation and exposure to violent online material. [17] The Interior Secretary said Roblox had likely influenced the shooter. [2]	A conclusive link to any online group. The Interior Secretary told a House budget hearing that none was found. [2]
Banga, South Cotabato 18 September 2026	A Grade 9 student (reported as 16 by most outlets) took an improperly secured firearm from home. The DILG reported three deaths including the shooter; other official tallies differ. Officials said he was active in a "true crime" online community, had shown signs of distress known to his family months earlier, and had warned friends in a group chat. [13] [18]	That any specific site or community directly prompted the attack. The Interior Secretary stopped short of saying so. [18]

Across the three cases, students with little visible offline warning were exposed to violent online material. In Banga, the warnings were real but were not acted on. The answer is not to profile children. It is to build systems that make harmful contact harder, make warning signs easier to report, and route them to people who can help.

2.2 The government response so far

Date	Action
April 2026	DICT and Roblox reach an initial agreement on child-safety measures, before any of the shootings.
June 2026	CICC blocks GoreBox; the block is later made permanent by agreement with the developer. [12] DICT urges Congress to pass a digital safety law with a minimum age of 16. [23]
August 2026	Roblox commits to National ID (PhilSys) age checks and a local office by October; Meta commits to cooperation and threat-data sharing. [19] SB 2424 is endorsed to the Senate floor. [22]
September 2026	Meta and DICT form a working group on a PhilSys age check. [19] A House committee approves a substitute bill. [21] CICC gives Reddit and Discord 24 hours to name local representatives; Discord is blocked on 23 September and restored on 24 September. [20] [25]
October 2026	The Senate approves SB 2424 on third and final reading on 8 October, by 16 votes to 1, barring under-18s from high-risk platforms. [29]

2.3 The September Discord block

The Discord case is the clearest recent example of how enforcement currently works, and it is worth setting out in some detail because it shapes the recommendations in this paper.

- On 21 September, CICC wrote to Reddit and Discord giving each 24 hours to name a resident agent or local representative, or face restriction.[20]
- CICC first said both platforms had been given extensions, then clarified that only Reddit had.[25]
- The block took effect on the evening of 23 September. It was applied network by network, so access was cut on some networks, including Globe and DITO, while it remained reachable on others, according to user reports.[26]
- Freelancers, virtual assistants and online communities that use Discord for work lost access. A former DICT Undersecretary estimated that at least 15,000 Filipino freelancers and virtual assistants were affected.[27]
- A VPN provider reported a sharp rise in sign-ups from the Philippines during the block, which stayed above its earlier level after the block was lifted.[27] This is vendor data, but it is consistent with experience elsewhere.
- The block was lifted on 24 September after Discord committed to give authorities a priority "green lane" for threats, to roll out its teen safety settings in the Philippines, and to set up a local office or in-country representatives. CICC publicly apologized for the inconvenience.[25][28]

The outcome, platform commitments and local accountability, was the right one. The process was costly. A published protocol would have reached the same outcome with less disruption and on firmer legal ground.

2.4 Legislation

On 8 October 2026 the Senate approved SB 2424 on third and final reading, by 16 votes to 1 with no abstentions. Now titled the Safe Media Access and Responsible Technology for Kids in Digital Spaces (SMART KIDS) Act, it bars anyone under 18 from creating accounts on or accessing "high-risk" social media and online game platforms.[29] A House committee substitute, consolidating ten bills, takes a different approach: it bars under-13s and restricts harmful content, short-form feeds and chat groups for users aged 13 to 17.[21] The bill becomes law only after the House passes its version, a bicameral conference reconciles the two, and the President signs it. The only vote against came from Senator Hontiveros, who argued for tiered restrictions for adolescents instead of an outright ban, and Senator Gatchalian said parental consent cannot override the restriction.[29]

Provision	Senate-approved SB 2424 (as reported)	House committee substitute (as reported)
Age rule	No accounts or access for under-18s on high-risk platforms	No social media under 13; feature limits for 13 to 17
Coverage	Social media and online games classified as high-risk. Exempt: email, private messaging, games not mainly used for social networking, educational, news and professional tools	Social media platforms

Provision	Senate-approved SB 2424 (as reported)	House committee substitute (as reported)
Age checks	Age-assurance mechanisms required; method not stated in reports. Senator Lacson had proposed the National ID as the primary check[22]	Not reported
Platform penalties	₱5 million to ₱20 million per violation, plus ₱100,000 per day of continued non-compliance	Not reported
Parents	Parents or guardians who knowingly help a minor bypass age checks: fine of up to ₱50,000, or community service and mandatory digital parenting and child safety programmes	Not reported
Enforcement	DICT implements and sets up a public reporting platform; urgent complaints escalated to platforms	Requires local registration of platforms

Senate details as reported by Tech Pilipinas, Philstar and NewsBytes.PH.[29] Reports did not state which age-assurance method the Senate adopted. DICT had earlier called for a minimum age of 16.[23]

Both are reasonable starting points, and this paper takes no position on the exact age. Its concern is what any version leaves open, which Section 3 explains. Three features of the Senate version matter in particular. First, the exemptions keep children on private messaging apps and games, where much of the 2026 harm took place, so safe defaults are still needed there. Second, an under-18 rule means every adult user of a high-risk platform must also pass an age check, which raises the stakes for privacy and security. Third, the fine on parents sits uneasily with the aim of encouraging families to ask for help.

2.5 Four types of online risk

This paper uses the OECD's "4Cs" to describe the risks children face online.[16]

Risk type	What it means	Philippine examples
Content	What children see	Violent edits, "nothing matters" content and self-harm material served by feeds
Contact	Who reaches children	Adults posing as peers in game chats;[6] recruitment into violent online communities
Conduct	How children behave and are treated	Cyberbullying; identity built around violent subcultures; livestreaming harmful acts
Contract or commercial	How the system is designed to earn money	Engagement-driven feeds; weak age checks; profiling and advertising aimed at minors

On scale, *Disrupting Harm in the Philippines* (ECPAT, INTERPOL and UNICEF, 2022) found that 20% of internet-using children aged 12 to 17 had experienced online sexual exploitation or abuse in the past year, around two million children.[15] That figure concerns sexual exploitation, not violent radicalization.

2.6 A complementary layer: firearm storage

The Interior Secretary stated that Banga was the third time this year that an improperly secured gun was used in a school attack.[13] Firearm storage and owner accountability fall outside this paper's digital focus, but a complete national response needs both.

3. Analysis

3.1 How harmful content escalates

Regulators in the European Union, the United Kingdom and Australia describe a common pattern in how recommendation systems can draw young users toward more extreme material:[3]

1. Exposure: a child sees one violent clip.
2. Recommendation: the feed serves more of the same.
3. Escalation: content and communities grow more extreme.
4. Normalization: the extreme starts to feel ordinary.

The research evidence is mixed. It also shows real benefits from children's use of the internet. That is why this framework focuses on safer design rather than on cutting children off entirely.

3.2 Where the fixes go: four system layers

Each risk moves through four layers. No single layer stops a risk on its own, and no single actor controls all four. Safety is therefore a shared responsibility.

Risk	Device	Network	Platform	People
Content	Age-rated app store; filters on for child profiles	Opt-in family filtering at no cost	Safer feeds for minors; fast removal	Children know how to report
Contact	Child profile limits contact from strangers	None at this layer	Private accounts; no direct messages or group invites from strangers	Adults recognize grooming; one place to report
Conduct	Screen-time and livestream controls	None at this layer	Block, mute and report tools; livestream safeguards	School teams act on warnings; counseling
Contract	Age-range signal to apps, never a birthdate	Clear family plans	No profiling ads to minors; engagement features off	Regulators audit and publish results

3.3 Why an age limit alone is not enough

The table below shows the gaps any age limit leaves open and the pillar of the framework in Section 4 that closes each one.

Gap	Example	Pillar
Children on services the age limit allows or exempts get adult-grade defaults unless the law requires otherwise	A 16-year-old in an exempt messaging app or game, open to contact from strangers	1. Safe by Default
Children below the limit get in anyway	False birthdates, borrowed accounts, a parent's phone, VPNs	2, 3 and 4

Gap	Example	Pillar
Harms outside "social media"	Game chats, Discord servers, messaging groups	1 (defaults apply to any service likely to be used by minors)
Age checks can create new risks	A central ID check becomes an attack target; unregistered children are excluded	2. Privacy-Safe Age Checks
Enforcement by ultimatum	Whole-platform blocks that fall unevenly and cut off legitimate users	5. Platform Restriction Protocol
Warning signs with nowhere to go	Group-chat warnings; signs of distress known at home	4. Early Help; 5. one place to report
No way to know if it works	No baseline survey, no published data	5. Measurement and transparency

3.4 Lessons from enforcement by ultimatum

The GoreBox and Discord cases show both the strengths and the limits of acting platform by platform.

- Speed works: deadlines brought platforms to the table within days and produced real commitments.
- Uneven blocking is arbitrary in practice. When a block depends on which network a user is on, protection is inconsistent and the rule is hard to explain.
- Legitimate users pay the cost. Workers and communities with no connection to the harm lose access.
- Blocks push users to VPNs, which weakens the very age limits and parental filters the new law relies on.
- Unclear terms undermine trust. Conflicting statements about extensions created confusion that written notice would have avoided.

A published restriction protocol keeps the speed and removes most of the cost. It is set out in Pillar 5 and drafted in the companion Roadmap.

3.5 Parents cannot close these gaps alone

Filipino parents care deeply about their children's safety, but safety tools are scattered, technical and switched off by default. International evidence suggests many parents need help with tasks such as these, although a Philippine survey has not yet measured the gap:

- Setting up Screen Time on iPhone or iPad, or Family Link on Android
- Understanding how VPN apps can bypass filters
- Limiting app downloads by age rating
- Turning on family filtering with their telco
- Recognizing signs of grooming, radicalization or distress

This is not a moral failing. It is a gap created by systems that were not designed with parents in mind.

4. Proposed Framework

Safety by design means the safe path is the easy path. The framework has six guiding principles and five pillars.

4.1 Guiding principles

Principle	In practice
Children's rights first	Protect children while respecting their rights to information, participation and privacy (UN General Comment 25).
Systems over symptoms	Rules on design, defaults and response work better than one-off blocks.
Privacy and security by construction	Check age without creating an identity trail. Treat any central age-check service as critical infrastructure.
Support families, do not fine them	Penalties fall on companies that fail their duties, never on children. Where a parent knowingly helps a child get around the rules, the response should be education and support, not a fine.
Fair process	Every restriction on a platform follows published criteria, notice, a chance to respond, time limits and review.
Evidence-led and reviewable	Measure against a baseline, publish results, and review the law within two years.

4.2 Pillar 1: Safe by Default

Any service likely to be used by minors, not only social media, should start minors in the safest setting. This should apply to every user under 18 on every service they can lawfully use, including services exempt from the age limit, such as private messaging apps and most online games. Main owners are platforms, game and messaging services, with DICT as regulator.

Minimum expectations

- Minors' accounts are private by default.
- Strangers cannot message minors or add them to group chats by default, consistent with the House bill's chat-group restriction for ages 13 to 17.
- Feeds for minors filter violent, self-harm and adult content, and short-form autoplay is off by default.
- Engagement features such as streaks and late-night notifications are off by default.
- No profiling-based advertising to minors.
- Block, mute and report tools are easy to find, with fast moderation.
- A named senior person is accountable for child safety in the Philippines.

Precedents: EU Digital Services Act, Article 28 and the 2025 Commission Guidelines; UK Online Safety Act and Ofcom's Children's Codes; Brazil's ECA Digital.

4.3 Pillar 2: Privacy-Safe Age Checks and Device Setup

Any age limit depends on age checks. Government is building one route through the National ID system (PhilSys), starting with Roblox and Meta.[19] Under the Senate's under-18 rule, every user of a high-risk platform, adults included, would need to show they are old enough, so the age-check system could reach most Filipino adults online. This paper supports PhilSys as one route, not the only route, with five safeguards. Main owners are DICT, PSA, NPC, device makers and app stores.

Safeguard	What it means
1. Answer the age question, not the identity question	PSA replies only "above or below the age threshold," not whether an ID matches a named person.
2. Platforms keep nothing	No retention of ID numbers or check results beyond a yes or no age flag. NPC audits compliance.
3. More than one route	Device age-range signals (already used by Apple and Google in several countries), age estimation with safety buffers, and PhilSys all count. Children not yet registered are not locked out.
4. Independent oversight	NPC acts as privacy co-regulator, and a privacy impact assessment is published before launch.
5. Security from day one	A central age-check service is a high-value target. It needs threat modelling, rate limits, logging, penetration testing and a breach plan before go-live.

Device setup: minimum expectations

- Setup asks whether a child will use the device.
- The device shares only an age range with apps (under 13, 13 to 15, 16 to 17, 18 and over).
- Safe browsing, age-rated downloads and content filters are on for child profiles.
- VPN apps need parent approval on child profiles.
- A single "family safety mode" switch turns everything on. Apple Screen Time and Google Family Link are the closest existing examples.

Precedents: California AB 1043 (device age signal, 2027); Apple Declared Age Range and Google Play Age Signals; Singapore's App Distribution Code; Australia's Age Assurance Technology Trial. RA 11930 already requires anonymous age verification for adult-content providers and a policy study on age controls by intermediaries.[24]

4.4 Pillar 3: Safer Networks

Telcos and ISPs reach almost every household and can add a simple layer of protection. Main owners are telcos, ISPs and NTC.

Minimum expectations

- An opt-in family-safe network profile at no extra cost.
- Easy to switch on or off, with parental override.
- Promoted at the point of sale, in packaging and by SMS, in Filipino and regional languages.
- No blanket network-level VPN blocking, which would disrupt banking and business.

Precedents: UK family-friendly network filters; NTC's existing quarterly RA 11930 compliance reports from intermediaries.[24]

4.5 Pillar 4: Capable Adults and Early Help

Parents, teachers and counselors need practical skills and a clear path from a warning sign to help. Main owners are DepEd, DOH, DSWD, LGUs and School Counselor Associates under RA 12080.

Minimum expectations

- A short, practical parent curriculum delivered through schools, barangays, parishes, clinics and employers.
- School teams trained to take warnings seriously, including those shared in group chats, without over-reacting to ordinary teenage remarks.
- A distress pathway: when a family or school learns of self-harm notes, depression or violent fixation, they know whom to call (school counselor, DOH or the National Center for Mental Health, DepEd TeleSafe #337333).
- Students taught that reporting a friend's warning protects that friend.
- Success measured by what parents and schools do, not by attendance.

Precedents: US Secret Service school threat-assessment model; UNICEF and OECD guidance on parental mediation.

4.6 Pillar 5: Fair, Fast Enforcement and Accountability

Enforcement must be fast enough to protect children and fair enough to survive legal challenge. Main owners are DICT, CICC, NBI, PNP, DOJ, NPC and CHR.

Minimum expectations

- A published Platform Restriction Protocol: criteria, written notice, a right to respond, the least restrictive measure first, time limits, review and a public log.
- For imminent threats, targeted action on specific content or accounts within hours, before any whole-platform block.
- Where a block is unavoidable, a single order applied consistently across all networks, with a stated end date.
- Local representatives and 24/7 law-enforcement contacts for major platforms.
- Safeguards on the "green lane" monitoring access platforms have offered: legal basis, human review, no automated profiling of children, and audit.
- One referral pathway across CICC, PNP, NBI, DepEd, DSWD and DOH.
- Annual transparency reports from platforms, building on NTC's quarterly reports, plus a national baseline survey and a public dashboard.

Precedents: Ofcom enforcement; Australia's eSafety compliance updates; Singapore's Online Safety Commission; Indonesia's tiered enforcement.

5. Implications

5.1 Who leads each pillar

Each stakeholder owns part of the solution. L marks the lead role and S a supporting role.

Stakeholder	P1	P2	P3	P4	P5
DICT	L	L	S	S	L
PSA		L			
NTC		S	L		S
NPC	S	L	S		S
CICC, NBI, PNP, DOJ	S				L
DepEd, DOH, DSWD, LGUs				L	S
Congress	S	S	S	S	S
Platforms, game and messaging services	L	S		S	S
Device makers and app stores	S	L		S	
Telcos and ISPs		S	L	S	
Schools				L	S
Parents and communities			S	L	S

P1 Safe by Default; P2 Age Checks and Device Setup; P3 Safer Networks; P4 Capable Adults and Early Help; P5 Enforcement and Accountability.

5.2 For government and regulators

The main shift is from acting case by case to acting by published rule. That protects the agencies as much as the public: a protocol-based decision is easier to defend in court and easier to explain to users. It also means treating the PhilSys age-check service as critical infrastructure, with security testing and privacy oversight before launch.

5.3 For Congress

Whichever version emerges from the bicameral conference, the age limit can be kept while adding the duties that make it work: safe defaults for every under-18 on the services they can still use, coverage of games and messaging, several approved age-check routes, fair process for restrictions, education rather than fines for parents, and a two-year review. Section 6.4 sets these out as ten amendments.

5.4 For platforms, device makers and telcos

Most of what this framework asks already exists somewhere in industry practice: teen account defaults, device age signals, family filters. The change is that these become consistent, on by default

and verifiable in the Philippines. Clear rules also reduce the risk of sudden blocks, which serves platforms and their users.

5.5 For schools and families

Schools gain a defined path for acting on warning signs, and families gain practical tools and a number to call. Neither is expected to carry the problem alone. The aim is that a warning seen by a classmate or a parent reaches someone who can help, quickly and without punishment.

5.6 Alignment with global standards

Jurisdiction	Key instrument	What the Philippines can adopt
European Union	Digital Services Act, Article 28; Guidelines (2025)	Safe defaults for minors; no profiling ads; exemptions for small local developers
United Kingdom	Online Safety Act 2023; Ofcom Children's Codes (2025)	Codes of practice; a named accountable officer; risk assessments
United States	COPPA; California AB 1043 (2027); state app-store laws	A device age-range signal at setup with minimal data
Singapore	Online Safety Codes; Online Safety Commission (2026)	Start with app stores and operating systems; victim redress
Australia	Social Media Minimum Age Act (December 2025)	A minimum age paired with "reasonable steps," platform-only penalties and a two-year review
Brazil	ECA Digital (in force March 2026)	A coordination decree; device and app-store age signals; local representation
Indonesia	PP 17/2025 (2026 enforcement)	Tiered, risk-based treatment of platforms

6. Recommendations

6.1 Short term: the first 90 days (October to December 2026)

- Publish a Platform Restriction Protocol and apply it to every future restriction (DICT, CICC, DOJ).
- Set the five PhilSys safeguards in the Roblox and Meta integrations before go-live (DICT, PSA, NPC).
- Publish safeguards for platform "green lane" monitoring access (CICC, NPC, CHR).
- Publish a register of platform commitments, including Roblox, Meta, Reddit and Discord, with deadlines (DICT, CICC).
- Issue a school distress pathway memorandum (DepEd, DOH).
- Table the ten amendments in House plenary debate and the bicameral conference (House, Senate conferees).

6.2 Medium term: to September 2027

- Draft implementing rules and a Child Safety Baseline Standard, with public and youth consultation.
- Field a national baseline survey of parents and teenagers.
- Pilot the parent curriculum, school support teams, a single referral pathway, device setup prompts and a telco family profile.
- Publish pilot results, a default-settings audit and a public dashboard.

6.3 Long term: to June 2028 and beyond

- Enforce the standard, with the first annual platform transparency reports.
- Scale successful pilots nationally.
- Commission an independent review of the law within two years of enactment.

6.4 Ten amendments to make the law work

These apply to whichever version emerges from the bicameral conference. Plain-language model text is in the companion Roadmap.

No.	Amendment	Reason
1	Require safe defaults for every user under 18 on any service they can use, including services exempt from the age limit and, under a lower House age, teens above it	Closes the largest gap an age limit leaves
2	Apply the default-safety duties to services likely to be used by minors, including messaging apps and games exempt from the age limit	The 2026 cases involved game and chat platforms, which the Senate version largely exempts

No.	Amendment	Reason
3	Allow several age-check routes (device signals, estimation, PhilSys); no single mandatory method	Inclusion, resilience and fewer workarounds
4	PhilSys checks return an age answer only; platforms may not keep ID numbers	Privacy and security
5	Make NPC a co-regulator for age assurance and require a published privacy impact assessment	Independent oversight
6	Put the Platform Restriction Protocol in law: notice, reply, least restrictive step, consistent application across networks, time limit, review, public log	Fair, consistent and challenge-resistant enforcement
7	Require local representation and 24/7 law-enforcement contacts	Already sought by CICC; make it a standing rule
8	Penalize operators, scaled to size, never children. Replace the Senate's fine on parents who knowingly help a child bypass age checks with the digital parenting and child safety programme alone	Fines may stop families from asking for help; education addresses the cause
9	Require annual transparency reports and a national baseline survey	Measure whether the law works
10	Require an independent review within two years	Learn and adjust, as Australia does

7. Conclusion

No law, age check or parental control replaces a trusted adult who is present and willing to listen. Children who feel heard are less likely to seek belonging from strangers online and more likely to speak up when something is wrong. The Banga case is a painful reminder that warning signs were present, and that the adults around the student needed a clear path to help.

The Philippines has already acted with urgency. The next step is to turn urgent measures into a lasting system: safe by default, private by design, fair in enforcement, supportive of families and honest about results.

The goal is not to build walls around children. It is to raise children who can navigate a world without walls, inside a system designed to keep them safe.

About the Author

Albert Pangilinan Dela Cruz is a cybersecurity practitioner and digital governance advocate with experience in cybersecurity operations, digital governance and public-sector technology policy. He has contributed to national initiatives including the E-Commerce Act, the Cybercrime Prevention Act, the Data Privacy Act, and the Supreme Court's Rules on Electronic Evidence and Digital Notarization.

He wrote this paper independently as a private citizen. It does not represent the views of any employer, company, client or organization. He also works as a cybersecurity consultant and has business interests in technology; this paper was not commissioned or funded by any client or company.

References

- [1] DOJ spokesperson Polo Martinez, press briefing, 23 June 2026, as reported by the Philippine News Agency and Philstar.com.
- [2] DILG Secretary Jonvic Remulla, briefing on the Ateneo de Zamboanga shooting, 19 August 2026 (Philippine Daily Inquirer; GMA News Online). The statement that there was "no conclusive link" to the online group True Crime Community/764 was made at a House of Representatives appropriations committee hearing on the DILG budget (Philippine Daily Inquirer, 27 August 2026). The Roblox-influence view is the Secretary's stated assessment, not a forensic finding.
- [3] UNICEF Innocenti and LSE, Global Kids Online; Ofcom, Children and Parents: Media Use and Attitudes Report (21 May 2026); European Commission, DSA Article 34 risk assessments; Australian eSafety Commissioner reports.
- [4] "SILG Remulla pushes total Roblox ban amid Ateneo de Zamboanga shooting," GMA News Online, 20 August 2026.
- [5] Secretary Remulla, press briefing, Camp Crame, 19 August 2026 (Philippine Daily Inquirer; The Star, Malaysia).
- [6] Education Secretary Sonny Angara, in "Roblox questioned on community safeguards amid PH scrutiny over youth safety," Rappler, 20 August 2026.
- [7] National Center for Missing & Exploited Children, CyberTipline data on online enticement, 2019 to 2020.
- [8] Winters and Jeglic, Journal of Child Sexual Abuse, 2020; Winters, Jeglic and Johnson, Child Abuse & Neglect, 2023.
- [9] "Imee Marcos to regulatory agencies: Do your jobs," The Manila Times, 21 August 2026.
- [10] "BOC relieves father of school shooting suspect," Philstar.com, 19 August 2026.
- [11] Wright and Ysart, "Beyond the Gates," Coalition of Cyber Investigators, 30 July 2026, citing US Secret Service National Threat Assessment Center (2021).
- [12] "GoreBox developer voluntarily agreed to game block, DICT," Philstar.com, 26 August 2026.
- [13] Manila Bulletin, 18 to 19 September 2026; Philstar.com, 19 September 2026; GMA News Online live updates, 18 to 21 September 2026. <https://www.gmanetwork.com/news/topstories/regions/1002870/live-updates-banga-national-high-school-shooting-kills-3-in-south-cotabato/story/>
<https://www.philstar.com/headlines/2026/09/18/2557162/deped-confirms-shooting-south-cotabato-school>
- [14] Philippine bills as reported in the press to September 2026 (SB 2066, SB 2424, HB 9665, HB 9825 and others). To be confirmed against official texts.
- [15] ECPAT, INTERPOL and UNICEF, Disrupting Harm in the Philippines, 2022.
- [16] OECD, Recommendation on Children in the Digital Environment and revised typology of risks, 2021.
- [17] "No link between violent online group, Zamboanga shooter, DILG," Philippine Daily Inquirer, 27 August 2026 (includes PNP investigation findings). <https://newsinfo.inquirer.net/2293498/no-link-between-violent-online-group-zamboanga-shooter-dilg>
- [18] GMA News Online, 19 to 20 September 2026; "Shooter visited 'deep dark web,' left notes," The Manila Times, 20 September 2026; Pilipino Star Ngayon, 20 September 2026.
<https://www.gmanetwork.com/news/topstories/regions/1002917/live-updates-banga-national-high-school-shooting-probe/story/> <https://www.manilatimes.net/2026/09/20/news/national/shooter-visited-deep-dark-web-left-notes/2428762>
- [19] "Roblox to set up local office in October, CICC," The Manila Times, 29 August 2026; "Roblox, Meta to beef up protection vs online threats," Philstar.com, 29 August 2026; ID Tech, on the Meta and DICT PhilSys working group,

September 2026. <https://www.manilatimes.net/2026/08/29/business/corporate-news/roblox-to-set-up-local-office-in-october-cicc/2414296> <https://www.philstar.com/headlines/2026/08/29/2552510/roblox-meta-beef-protection-vs-online-threats> <https://idtechwire.com/meta-and-philippine-government-form-working-group-on-national-id-verification/>

[20] Philippine Daily Inquirer, 22 to 24 September 2026, on the CICC directives to Reddit and Discord; Philstar.com, 22 and 24 September 2026. <https://globalnation.inquirer.net/339005/cicc-gives-reddit-discord-24-hours-to-set-up-ph-presence> <https://www.philstar.com/headlines/2026/09/22/2558090/cybersecurity-agency-directs-discord-reddit-set-philippine-offices> <https://www.philstar.com/headlines/2026/09/24/2558608/cybersecurity-agency-clarifies-extension-granted-only-reddit-not-discord>

[21] "House panel approves bill banning social media for children under 13," Tech Pilipinas, 8 September 2026. <https://techpilipinas.com/house-panel-bill-banning-social-media-children-under-13/>

[22] Philippine Daily Inquirer, on the Senate committee endorsement of SB 2424, August 2026; BusinessMirror and Manila Bulletin, 15 September 2026, on Senator Lacson's PhilSys amendment. <https://newsinfo.inquirer.net/2293186/senate-panel-endorses-bill-setting-16-as-minimum-age-for-social-media> <https://businessmirror.com.ph/2026/09/15/senator-proposes-use-of-national-id-for-age-verification-in-socmed/> <https://mb.com.ph/2026/09/15/national-id-eyed-as-primary-age-check-for-social-media-users>

[23] GMA News Online, explainer on the social media ban debate, 27 June 2026. <https://www.gmanetwork.com/news/topstories/nation/992931/explainer-is-a-social-media-ban-for-minors-protection-or-just-a-band-aid-solution/story/>

[24] Republic Act No. 11930 and its Implementing Rules; NTC Memorandum Order No. 006-09-2025. <https://elibrary.judiciary.gov.ph/thebookshelf/showdocs/2/97582> <https://elegal.ph/ntc-requires-internet-intermediaries-to-submit-quarterly-compliance-reports-on-oseac-and-csaem/>

[25] "Discord ban lifted, CICC" and "CICC apologizes for inconvenience caused by Discord ban," GMA News Online, 24 September 2026. <https://www.gmanetwork.com/news/topstories/nation/1003532/discord-ban-lifted-cicc/story/> <https://www.gmanetwork.com/news/topstories/nation/1003628/cicc-apologizes-for-inconvenience-caused-by-discord-ban/story/>

[26] Unbox.ph, "Globe, GOMO, DITO has blocked Discord," 23 September 2026; Philippine Daily Inquirer, 23 September 2026, on the NTC order. <https://unbox.ph/?p=267836>

[27] Proton VPN, "Philippines VPN sign-up spike," September 2026. Vendor data; includes the estimate by former DICT Undersecretary Jeffrey Ian Dy. <https://protonvpn.com/blog/philippines-spike-9-2026>

[28] "Discord, Reddit to comply with order to set up PH office: CICC," PhilSTAR Life, 23 September 2026. <https://philstarlife.com/news-and-views/933293-discord-reddit-to-set-up-ph-office-cicc>

[29] "Senate passes bill restricting social media access for minors under 18," Tech Pilipinas, 8 October 2026; "Senate OKs bill restricting socmed for under-18s," Philstar.com, 9 October 2026; "Senate passes bill restricting under-18 access to high-risk online platforms," NewsBytes.PH, 8 October 2026. <https://techpilipinas.com/senate-passes-bill-restricting-social-media-access-minors-under-18/> <https://www.philstar.com/headlines/2026/10/09/2561963/senate-oks-bill-restricting-socmed-under-18s> <https://newsbytes.ph/2026/10/08/senate-passes-bill-restricting-under-18-access-to-high-risk-online-platforms/>

Appendix A: Glossary

Term	Meaning
Age assurance	Any method of checking a user's age range, from a device signal to an ID check
Age-range signal	A message from a phone to an app saying only that the user is, for example, 13 to 15, without a birthdate
Green lane	A priority channel platforms give authorities to report or investigate threats
Grooming	When an adult builds trust with a child in order to abuse, exploit or radicalize them
Leakage	Signs of violent intent that a person reveals to others before an attack
PhilSys	The Philippine Identification System (National ID), run by PSA
Platform Restriction Protocol	A published, fair process for restricting or blocking an online service
Safety by design	Building products so that the safe choice is the easy, automatic one
VPN	An app that hides where internet traffic goes; it can bypass filters and blocks
4Cs	The OECD typology of online risk: Content, Contact, Conduct, Contract

Appendix B: What Can and Cannot Be Claimed

Claim in public debate	Evidence status	Suggested wording
A game or platform caused the shootings	Not established by investigators	"The student was reported to use X; a causal role has not been established."
Online material played a part	PNP (Zamboanga) and DILG (Banga) cite exposure to violent online material and communities	"Investigators identified exposure to violent online material among the circumstances."
Warning signs appeared before attacks	Banga: group-chat warning and earlier signs of distress reported	"Warning signs were visible to others. A system must help them reach someone who can act."
The Discord block harmed legitimate users	Reported by users and media; the 15,000 figure is one person's estimate, reported by a VPN vendor	"Legitimate users, including freelancers, lost access; one estimate put the number at 15,000 or more."
Blocks drive VPN use	Vendor data from one provider	"A VPN provider reported a sharp rise in Philippine sign-ups during the block."
Online sexual exploitation of Filipino children is widespread	Disrupting Harm (2022): 20% of internet-using 12 to 17 year-olds in the past year	State with year and definition.
Most parents do not know controls exist	No Philippine survey yet	Use "likely" until a baseline survey reports.
The Philippines has banned social media for under-18s	The Senate approved SB 2424 on 8 October 2026; the House version differs and the bill is not yet law	"The Senate has approved an under-18 rule for high-risk platforms; it still needs the House, a bicameral conference and the President's signature."
Minimum-age laws reduce harm	No outcome evidence yet; Australia's review is due by end-2027	"Adopted in Australia; results not yet evaluated."
PhilSys covers about 96% of the population	Stated by Senator Lacson citing PSA; date and basis not given	Attribute: "reported by Senator Lacson, citing PSA."

Appendix C: Quick-Start Checklists

For parents: a 10-minute setup

- Set up a child profile (Screen Time on iPhone or iPad; Family Link on Android).
- Limit downloads to your child's age rating and require approval for new apps, including VPNs.
- Turn on safe search and content filters.
- Set TikTok, YouTube, Roblox, Discord and game accounts to private, and turn off messages from strangers.
- Ask your telco about free family-safe settings.
- Talk regularly: ask who your child talks to online and which groups they are in.
- If you find signs of self-harm, depression or violent fixation, contact the school counselor or DepEd TeleSafe #33733 promptly.
- Save CICC 1326 for reporting online threats.

For schools

- Name a threat-assessment and support team (school head, counselor and a teacher, with DSWD and DOH links).
- Teach students that reporting a warning, including one seen in a group chat, protects their friend.
- Set up a distress pathway: counselor, then DOH or TeleSafe, then the family.
- Hold parent sessions and publicize TeleSafe and CICC 1326.
- Prepare a post-incident plan, including psychosocial support and a rule against sharing violent footage.

For platforms: a self-check

- Are all under-18 accounts private by default, including teens above the minimum age?
- Can strangers message minors or add them to groups?
- Do feeds for minors filter violent and self-harm content?
- Is there a local representative and a 24/7 law-enforcement contact?
- Do you keep ID numbers from age checks? You should not.

Appendix D: Revision Note

Version 5.2 (9 October 2026)

- Updated for Senate approval of SB 2424 (SMART KIDS Act) on 8 October 2026: under-18 rule for high-risk platforms, exemptions, penalties and parent provisions; added a Senate and House comparison to Section 2.4.
- Reframed Pillar 1 and amendments 1 and 2 around services exempt from the age limit; noted that an under-18 rule extends age checks to adult users.
- Replaced the principle "no punishment of families" with "support families, do not fine them" and revised amendment 8.
- Added reference [29] and an entry to Appendix B; added links to online sources; added a status line and a disclosure of the author's professional interests.

Version 5.1 (8 October 2026)

- Added Section 2.3 on the September Discord block and Section 3.4 on lessons from enforcement by ultimatum.
- Added consistent application across networks to Pillar 5 and to amendment 6.
- Added references [25] to [28] and two entries to Appendix B.

Version 5.0 (26 September 2026)

- New thesis: an age limit is the floor, not the system.
- Updated the Banga, Zamboanga, Roblox, Meta and legislative facts; corrected the venue in reference [2].
- Redesigned Pillar 2 around PhilSys as one route with five safeguards; added a distress pathway to Pillar 4.
- Added the ten amendments and new legal hooks under RA 11930 and NTC Memorandum Order 006-09-2025.
- Clarified the author's role as an independent private citizen.