Regulatory Institute

Model Provisions for the Online Protection of Minors

These model provisions were designed to establish a comprehensive framework for the protection of minors in digital environments. As children and adolescents increasingly engage with online platforms, social media, gaming services, and other digital technologies, they face unique vulnerabilities that existing regulatory frameworks often fail to adequately address. These model provisions aim to fill that gap by offering policy-makers and drafters a wide range of potential policy and regulatory options on this subject, without prejudging or prescribing any particular policy or legislative choice.

The model provisions will, of course, need to be adapted to the legal and policy context in which they are used. In particular, obligations and powers may need to be assigned to authorities and persons other than those set out in the model provisions.

The model provisions were primarily drafted by one large language model, Claude Sonnet, with input from several large language models during the references collection and ideation phase. They have undergone substantive human editing. The model provisions do not constitute a single, integrated, or internally consistent document. To avoid prejudging any decisions, we have deliberately left some variations and overlaps.

To make use of the model provisions effectively, the following four steps are recommended:

- Select the provisions deemed relevant;
- Adapt them to the relevant context;
- Merge and reorganise them as necessary;

These model provisions may be further supplemented, particularly regarding enforcement mechanisms. Regulators are invited to refer to the Institute's broader corpus of resources, including:

- The Cross-Sectoral Standard Model Provisions;
- The List of Sanctions and Accompanying Measures;
- The List of Powers and Obligations;
- The Model Provisions on Vulnerable Persons and Persons with Special Needs; and
- The Model Law on Cross-Border Internet Activities and Virtual Worlds.

Taken together, these instruments may serve to reinforce, operationalise, and expand the present provisions, thereby ensuring a more coherent and effective legislative regime for the protection of minors online.

For questions regarding this version of the model provisions, or to request or contribute to further sets, please do not hesitate to contact us: manager@regulatoryinstitute.org.

Table of Contents with Page Numbers

(Table of Contents with links: see below)

Table of Contents with Links

PART I: FOUNDATIONAL FRAMEWORK

Chapter 1: General Provisions

1. Title and Purpose

This Act shall be known as the **" Online Protection of Minors Act."**

This Act aims at establishing a comprehensive framework to protect minors in digital environments by preventing exposure to harmful content, safeguarding personal data, empowering parents and guardians with effective oversight tools, and ensuring robust enforcement mechanisms.

2. Scope and Application

2.1 This Act applies to anybody who is in online contact with a minor residing, living or being in … (jurisdiction). This Act also applies to anybody residing, living or being in … (jurisdiction) who is in online contact with a minor anywhere in the world. “Online contact” includes cases in which minors are seen or heard or see and hear against their will and without active or voluntary contribution.

2.2 This Act applies to minors contacted in accordance with the previous paragraph.

2.3 This Act also applies to all digital service providers, their commercial partners and both their staff that:

1) process personal data of minors within the jurisdiction,
2) facilitate user interaction or content creation accessible to minors covered by the previous paragraph, or
3) employ algorithmic systems affecting minors, or target marketing or content toward minors. This includes social media platforms, video-sharing and livestreaming platforms, gaming platforms, messaging applications, content recommendation services, educational technology platforms, and any digital service accessible to persons under 18 years.

2.4 This Act applies regardless of: the location of incorporation or headquarters of the Digital Service Provider, the location of servers or technical infrastructure, whether the service is provided free of charge or for payment, and the technology or medium used to provide the service.

3. Definitions

For the purposes of this Act:

"**Minor**" means any natural person under eighteen (18) years of age, subcategorised as:

1) "Young Child": a person under seven (7) years of age;

2) "Child": a person who has attained the age of seven (7) years but is under thirteen (13) years of age;
3) "Adolescent": a person who has attained the age of thirteen (13) years but is under eighteen (18) years of age.

"**Digital Service Provider**" or "Provider" means any natural or legal person who:

1) Provides an information society service;
2) Operates an online platform, website, or application;
3) Offers software as a service accessible via the internet; or
4) Provides online gaming, social networking, or content sharing services.

"**Harmful Content**" means material that:

1) Depicts or promotes, including with puppets or other representations, illegal activity including:

- Sexual contacts of and with minors in any form;

- Violence, sexual abuse or sexual exploitation against humans or animals in any form;

- Terrorist activities or violent extremism;

- Illegal substance use or distribution; or

- Criminal activity instructions.

2) Threatens minor wellbeing through:

- Self-harm, suicide, or dangerous challenge promotion;

- Eating disorder encouragement or extreme dieting promotion;

- Dangerous activities likely to cause physical injury; or

- Cyberbullying, harassment, or targeted abuse.

3) Contains age-inappropriate material including:

- Sexual content not suitable for minor developmental stage;

- Graphic violence or disturbing imagery;

- Expression or promotion of hate or discrimination;

- Promotion of gambling; or

- Predatory commercial content.

4) Impairs healthy development by promoting unhealthy relationships with technology or social media, including with algorithms causing or enhancing dependency.

"**Disturbing Imagery**" includes any visual content, whether static, animated, or video, that is likely to cause significant emotional distress, psychological harm, or trauma to minors due to its graphic, violent, grotesque, or otherwise deeply unsettling nature.

"**Predatory Commercial Content**" means any marketing, advertising, or promotional material, including text, images, audio, video, interactive media, or algorithmically targeted content, designed to exploit minors' cognitive, emotional, or developmental vulnerabilities for financial gain or commercial advantage.

"**Processing**" means any operation performed on personal data, including:

1) Collection, recording, organisation, structuring, storage;
2) Adaptation, alteration, retrieval, consultation, use;
3) Disclosure by transmission, dissemination, or otherwise making available; and
4) Alignment, combination, restriction, erasure, or destruction.

"**Age Assurance**" means any system or process used to determine or verify that a user has attained a certain age, including but not limited to:

1) Self-declaration with verification mechanisms;
2) Identity document checks;
3) Biometric age estimation;
4) Third-party age verification services; and
5) Behavioral analysis systems.

"Parental Control System" includes technical tools enabling parental activity monitoring, content filtering, time management, communication oversight, and purchase controls.

"Privacy-Preserving Technology" includes zero-knowledge proofs, homomorphic encryption, differential privacy, and secure multi-party computation.

4. Guiding Principles

All provisions of this Act must be interpreted and applied according to the following principles:

Best Interests of the Minor: Any ambiguity in the interpretation of this Act must be resolved in favor of the interpretation that best protects the interests and welfare of minors.

Proportionality: Measures implemented under this Act must be proportionate to the risk of harm and the age of the minor.

Privacy by Design: Technical and organisational measures must incorporate privacy protection from the outset of system design.

Technological Neutrality: Requirements must be interpreted to apply regardless of the specific technology employed.

Rights-Based Approach: Interpretation and implementation of this Act must respect fundamental rights and the rights of minors established in this act.

Evidence-Based Policy: Measures taken in accordance with this Act must be grounded in research and best practices.

N.B.: Please select further interpretation rules from Chapter A. of the Cross-sectoral Standard Provisions.

5. Rights of Minors

5.1 Every minor has the right to:

1) Protection from all forms of online harm, exploitation, and abuse;
2) Privacy and protection of personal data with enhanced safeguards;
3) Access age-appropriate information and services;
4) Receive digital literacy education and online safety information;
5) Participate safely in digital environments appropriate to their age;
6) Have their views considered in matters affecting them online; and
7) Seek remedy for violations of their rights under this Act.

Chapter 2: Fundamental Obligations of Digital Service Providers

6. General Duty of Care

6.1 Every Digital Service Provider owes a duty of care to minor users to:

1) Take reasonable steps to prevent foreseeable harm;
2) Design services with minor safety as a paramount consideration;
3) Regularly assess and mitigate risks to minors;
4) Respond promptly **OR** within 24 hours to identified safety concerns; and
5) Maintain transparency about safety measures and policies.

6.2 The duty of care requires Digital Service Providers to:

1) Conduct Minor Safety Impact Assessments in accordance with a scheme provided by provided by the [REGULATORY AUTHORITY] before launching new features;
2) Implement age-appropriate design standards;
3) Establish clear policies and procedures for minor protection;
4) Train personnel on child safety requirements;
5) Maintain various communication channels for reporting of harmful content or illegal communication, and publicise these communication channels in their services;

6) Take down harmful content or block illegal communication within 24 hours after being notified thereof; and
7) Cooperate with law enforcement and child protection agencies[, including those of other jurisdictions].

6.3 Digital Service Providers must document compliance with the duty of care through:

1) Written safety policies and procedures;
2) Regular internal audits and assessments;
3) Incident logs and response records;
4) Training records and certifications for staff;
5) Use of harmful content or illegal communication detection software or artificial intelligence provided by the [REGULATORY AUTHORITY], where available;
6) Use of other harmful content or illegal communication detection software or artificial intelligence, where the [REGULATORY AUTHORITY] has not made available any such means; and
7) Annual compliance reports to the [REGULATORY AUTHORITY], following the format set out by the [REGULATORY AUTHORITY].

7. Prohibited Practices

7.1 Digital Service Providers may not:

1) Use manipulative design practices to exploit minors' developmental vulnerabilities;
2) Process minor personal data for targeted advertising without explicit parental consent;
3) Enable contact between adults and minors without appropriate safeguards or enable any contact between adults and children;
4) Amplify or recommend harmful content to minor users;
5) Implement features designed to maximize screen time at the expense of minor wellbeing;
6) Use dark patterns to manipulate minors into sharing personal information; or
7) Create artificial social pressure through engagement metrics visible to minors.

7.2 Violations of prohibited practices shall result in immediate cessation orders, penalties as specified in Chapters 11 and 14, public disclosure of violations, and enhanced monitoring requirements.

PART II: TECHNICAL AND SERVICE REQUIREMENTS

Chapter 3: Age Verification and Access Management

8. Mandatory Age Assurance Requirements

8.1 Every Digital Service Provider must implement age assurance systems that:

1) Accurately determine whether a user is a minor;
2) Categorise minors into appropriate age bands;
3) Apply corresponding protections based on age determination;
4) Regularly reverify age status; and
5) Maintain audit logs of age verification processes.

8.2 Age assurance systems must:

1) Achieve a minimum accuracy rate of … (e.g. 95) percent for determining if a user is under eighteen (18);
2) Achieve a minimum accuracy rate of … (e.g. 98) percent for determining if a user is under thirteen (13);
3) Undergo annual **OR** regular third-party audits to verify accuracy rates; and
4) Offer clear processes for users to correct age misclassification, with particular scrutiny where users claim to be older.

8.3 Age assurance methods must be proportionate to risk:

1) High-risk services necessarily include social media, dating platforms, platforms with user-generated content, messenger services, virtual worlds and games with communication channels accessible to gamers. High-risk services must use highly reliable methods;
2) Medium-risk services include selling and other commercial transaction platforms. Medium-risk services must use at least reasonably reliable methods;
3) Low-risk services must use at least self-declaration with plausibility checks or similar safeguards;
4) The [REGULATORY AUTHORITY] shall [publish guidance] **OR** [set out binding rules by means of … (act category)] on risk categorisation and corresponding age assurance methods. It may also, by means of … (act category) increase the minimum accuracy rates for high-risk and medium-risk services [in the light of technical progress].

9. Technical Standards for Age Verification

9.1 High-Risk Services must implement at least two of the following:

1) Government-issued identification verification with:

(A) Document authenticity checks;
(B) Liveness detection to prevent spoofing;
(C) Secure deletion of document images after verification; and
(D) Tokenisation of age verification status;

2) Facial age estimation technology that:

(A) Achieves accuracy within one (1) year for ages 13-14;
(B) Does not store biometric data beyond verification;
(C) Includes anti-spoofing measures; and
(D) Is certified by an accredited testing laboratory;

3) Credit card verification with:

(A) Authorisation and immediate reversal;
(B) Verification of cardholder age with issuing bank;
(C) No storage of financial information; and
(D) Alternative methods for users without credit cards.

9.2 Medium-Risk Services must implement at least one of the following:

1) Enhanced self-declaration with:

(A) Consistency checks across multiple data points;
(B) Behavioral analysis for age estimation;
(C) Periodic reverification requirements; and
(D) Parental notification for users claiming to be minors;

2) Third-party age verification services that:

(A) Are certified by the [REGULATORY AUTHORITY];
(B) Use privacy-preserving protocols;
(C) Provide verifiable age tokens; and
(D) Maintain appropriate insurance coverage;

3) Educational email domain verification combined with additional checks.

9.3 Low-Risk Services must implement the following:

Simple self-declaration with:

(A) Clear age requirements stated;
(B) Regular reminders about age restrictions;
(C) Monitoring for obvious misrepresentation; and

(D) Easy reporting mechanisms for age fraud.

10. Privacy-Preserving Age Verification

10.1 All age verification systems must:

1) Collect the minimum data necessary to determine age;

2) Delete verification data immediately after age determination, retaining only:

(A) Age bracket classification;
(B) Verification timestamp;
(C) Method used for verification;

3) Not share verification data with third parties except:

(A) Certified age verification services;
(B) As required by law enforcement with proper legal process;

4) Implement zero-knowledge proof systems where technically feasible; and

5) Provide users with transparency about data handling.

10.2 Prohibited practices in age verification include:

1) Creating persistent biometric profiles;
2) Using age verification data for any purpose beyond age determination;
3) Combining age verification data with other user data for profiling;
4) Retaining identity documents or biometric data; and
5) Discriminating based on the method of age verification chosen.

11. Age-Based Access Tiers

11.1 Digital service providers must implement the following access tiers:

1) Ages 0-6 (Young Child Tier): No independent account creation permitted; access only through parental accounts with curated educational content; no social features or external communication; mandatory time limits not exceeding one hour daily; no data collection beyond service provision.
2) Ages 7-12 (Child Tier): Account creation only with verified parental consent; content filtered to age-appropriate material; communication limited to parent-approved contacts; no public profile visibility; educational and creative tools prioritized; comprehensive parental oversight dashboard required.

3) Ages 13-15 (Young Teen Tier): Account creation with parental notification required; enhanced privacy settings enabled by default; content recommendations limited to age-appropriate material; social features restricted to mutual connections; time management tools activated by default; targeted advertising prohibited.
4) Ages 16-17 (Older Teen Tier): Near-adult privileges with targeted protections; access to most platform features except age-restricted content; enhanced privacy controls and data rights; transition planning tools for approaching majority; optional parental oversight available.

11.2 Access tiers must include regular age verification updates, smooth transitions between tiers, parental override capabilities, special provisions for emancipated minors, and cultural sensitivity adjustments.

12. Account Creation and Parental Consent

12.1 For users under thirteen (13) years of age, Digital Service Providers must:

1) Obtain verifiable parental consent before creating accounts;

2) Provide clear information to parents about:

(A) Data collection and use practices;
(B) Available parental controls;
(C) Content types accessible;
(D) Communication features enabled; and
(E) Rights to review and delete child data;

3) Implement robust consent mechanisms that:

(A) Verify the identity of the consenting adult;
(B) Confirm their relationship to the child;
(C) Record consent details and limitations;
(D) Allow granular consent for specific features; and
(E) Enable easy consent withdrawal.

12.2 Paragraph 12.1 also applies to any minors seeking access to services presenting (a) sexual or sadistic scenes or (b) ill-treatment of animals, unless the scenes are part of a documentary or of an artistic undertaking.

12.3 Acceptable methods for obtaining verifiable parental consent are **OR** include:

1) Calling a toll-free number staffed by trained personnel;
2) Videoconferencing with trained personnel;
3) Providing a digital signature that complies with legal standards;
4) Uploading government-issued identification for comparison; and
5) Using federated identity systems approved by the [REGULATORY AUTHORITY].

12.4 Providers must:

1) Retain records of parental consent for the duration of the account;
2) Reverify consent annually;
3) Notify parents of any material changes to services; and
4) Honor parental requests to review or delete child data within thirty (30) days.

Chapter 4: Privacy and Data Protection

13. Enhanced Privacy Requirements for Minors

13.1 Digital Service Providers must configure all accounts of minors with the highest privacy settings by default, disable features that could compromise privacy unless explicitly enabled, implement technical measures to prevent unauthorised access to data of minors, conduct Privacy Impact Assessments for any feature affecting minors, and document privacy protection measures and their effectiveness.

13.2 Default privacy settings for minors must include profile visibility limited to approved connections, location tracking and sharing disabled, contact restricted to existing connections, search engine indexing disabled, third-party data sharing prohibited, behavioral advertising disabled, and recommendation algorithms limited to age-appropriate content.

13.3 Changes to privacy settings for minors under sixteen (16) require: clear explanation of implications in age-appropriate language, cooling-off period of twenty-four (24) hours for significant changes, parental notification for users under thirteen (13), ability to easily revert to default settings, and audit log of all privacy setting changes.

14. Data Minimisation and Purpose Limitation

14.1 Providers must implement data minimisation by:

1) Collecting only data strictly necessary for providing the requested service;
2) Not processing data of minors for secondary purposes without explicit consent;
3) Implementing technical measures to enforce data minimisation;
4) Regularly reviewing and deleting unnecessary data; and
5) Documenting the necessity for each type of data collected.

14.2 The following data processing activities for minors are prohibited:

1) Creating personality profiles for commercial purposes;
2) Behavioral advertising based on profiling;
3) Sale or rental of personal data to third parties;
4) Cross-service tracking without explicit consent;

5) Processing special categories of data except where strictly necessary;
6) Using minor data to train AI systems without explicit consent; and
7) Emotion recognition or mood analysis for users under sixteen (16).

14.3 The following data retention limitations apply:

1) Active user data: Maximum ninety (90) days after last use;
2) Inactive account data: Maximum one (1) year;
3) Backup data: Maximum six (6) months;
4) Log data: Maximum three (3) months;
5) Deleted content: Maximum thirty (30) days; and
6) Evidence of legal violations: As required by law.

14.4 The obligations and limitations of this Section 14 do not apply where the Digital Service Providers collect, save and store data to fulfil obligations under this Act to track penalised actions of users or otherwise improve the protection of minors.

15. Special Categories of Data

15.1 Special categories of data are **OR** include data on:

1) Racial or ethnic origin;
2) Political opinions or religious beliefs;
3) Health or genetic data;
4) Biometric data for identification;
5) Data concerning sexuality or sexual orientation; and
6) Location data beyond general geographic region.

15.2 Processing of special categories of minor data is prohibited except:

1) With explicit consent from parents (for under 13) or the minor (for 13+);
2) When necessary for vital interests of the minor;
3) For reasons of substantial public interest authorized by law; and
4) When manifestly made public by the minor with parental awareness.

15.3 When processing special categories is permitted, Digital Service Providers must:

1) Implement enhanced security measures;
2) Limit access on a need-to-know basis;
3) Conduct specific impact assessments;

4) Provide enhanced transparency; and
5) Enable granular control over such data.

16. Data Subject Rights for Minors

16.1 Minors have the following enhanced rights:

1) Right to Access: Obtain their data in an understandable format within fifteen days.
2) Right to Rectification: Correct any inaccurate data within 24 hours **OR** seven days of the request.
3) Right to Erasure: Delete their data within seven days of the request.
4) Right to Portability: Export data in machine-readable format.
5) Right to Restriction: Limit processing while disputes are resolved.
6) Right to Object: Opt-out of any non-essential processing.

16.2 For minors under thirteen, these rights may be exercised by parents. For minors aged thirteen to fifteen, both the minor and parent may exercise these rights. For minors aged sixteen to seventeen, the minor exercises these rights independently, with parental access upon request.

17. Data Breach Notification

17.1 In case of a data breach affecting minors, providers must notify the [Regulatory Authority] within twelve hours of discovery, notify affected parents within twenty-four hours, notify affected minors in age-appropriate language within forty-eight hours, provide credit monitoring for affected minors for two years, conduct forensic analysis and provide detailed reports.

17.3 Breach notifications must detail the nature and extent of the breach, the types of data affected, the potential consequences, the mitigation measures taken, the steps to be taken by the affected individuals, and the contact details for addressing questions and other follow-up.

NB - This Chapter can be supplemented by the provisions of Chapter N of the Cross-sectoral Standard Provisions.

Chapter 5: Content Safety and Moderation

18. Content Classification System

18.1 All content accessible to minors must be classified into the following categories:

1) G (General): Suitable for all ages, including young children.

2) T (Teen): Suitable for ages thirteen (13) and above.
3) M (Mature): Suitable for ages sixteen (16) and above.
4) R (Restricted): Not suitable for minors.

18.2 Classification criteria shall consider violence and its graphic nature, sexual content and nudity, language and profanity, substance use depictions, frightening or intense scenes, dangerous or imitable behavior, and discriminatory content.

18.3 Providers must apply classifications before content is made accessible, display classifications prominently, enforce age-based access restrictions, allow parental override controls, and maintain a classification accuracy of at least ninety percent.

18.4 The [REGULATORY AUTHORITY] shall [publish guidance] **OR** [set out binding rules by means of … (act category)] on content classification criteria and the application of these criteria, including by establishing reference cases.

19. Prohibited Content

19.1 The following content may not be accessible[to any minor], even when collected or provided for scientific or documentary purposes:

1) Child Sexual Abuse Material (CSAM) in any form;
2) Content promoting or instructing self-harm or suicide;
3) Content promoting eating disorders or dangerous weight loss;
4) Terrorist content or content promoting violent extremism;
5) Instructional content for creating and using weapons or explosives;.
6) Content dealing with child trafficking or exploitation, except content warning against these phenomena in a non-shocking, educative way;
7) Content promoting dangerous challenges likely to cause harm;
8) Content promoting or presenting sexual acts on animals; or
9) Content promoting or presenting other ill-treatment of animals.

19.2 For minors under thirteen, also the following is prohibited:

1) Content depicting realistic violence;
2) Any sexual content beyond age-appropriate education;
3) Content designed to frighten;
4) Gambling or gambling-like mechanics; or
5) Direct marketing or commercial pressure.

19.3 Digital Service Providers must implement systems to proactively detect prohibited content using automated tools, remove detected content within two hours, prevent re-uploading using hash databases, report CSAM to authorities within fifteen minutes, and preserve evidence of any prohibited content for law enforcement.

20. Content Moderation Requirements

20.1 Providers must implement content moderation systems including: automated detection technology updated quarterly; human review teams trained in child safety; clear moderation guidelines that are publicly available; response time targets based on severity; and appeal processes for content decisions.

20.2 Automated detection must include: image and video analysis using perceptual hashing; natural language processing for text content; audio analysis for harmful speech; behavioral analysis for grooming patterns; and network analysis for coordinated harmful behavior.

20.3 Human moderation requirements:

1) Moderators must have completed forty (40) hours of initial training;
2) Moderators must undergo ongoing training of eight (8) hours quarterly;
3) Moderators must have psychological support and counseling available;
4) The daily exposure to harmful content may not exceed four (4) hours; and
5) Moderators must have the option to periodically rotate to other tasks to prevent burnout.

21. Rapid Response Protocols

21.1 Digital Service Providers must set up and apply tiered Rapid Response Protocols. The Rapid Response Protocols must cover the cases and fulfil the requirements set out in paragraphs 21.2 to 21.4.

21.2 An Imminent Harm Response must be initiated within one (1) hour at least in the following cases:

1) Credible threats of violence;
2) Active self-harm or suicide content;
3) Live streaming of harmful acts;
4) Child exploitation in progress; and
5) Emergency situations involving minors.

21.3 A High Priority Response must be initiated within six (6) hours) at least in the following cases:

1) Cyberbullying and harassment;
2) Non-consensual intimate images;
3) Hate speech targeting minors;
4) Dangerous challenge content; and
5) Grooming attempts.

22.4 A Standard Response must be initiated within twenty-four (24) hours) at least in the following cases:

1) Age-inappropriate content;
2) Mild policy violations;

3) Reported concerns requiring investigation; and
4) Content classification errors.

22. Anti-Grooming Measures

22.1 Providers must implement systems to detect and prevent grooming by monitoring for:

1) Adults initiating excessive private contact with minors;
2) Requests to move communication to less-monitored platforms or communication channels;
3) Requests for personal information or images;
4) Offers of gifts, money, or other inducements;
5) Attempts to isolate minors from parents or friends; and
6) Sexual or romantic language directed at minors.

22.2 When potential grooming is detected, providers must immediately suspend suspicious communications, alert trust and safety teams within thirty minutes, preserve evidence for law enforcement, notify parents of affected minors under sixteen, report to law enforcement within two hours, and provide support resources to potential victims.

22.3 Prevention measures must include:

1) Default settings preventing adult-minor contact;
2) Warnings when minors attempt to share personal information;
3) Age-inappropriate communication filters;
4) Regular safety education for minor users; and
5) Clear reporting mechanisms accessible to minors.

NB - This Chapter can be supplemented by the provisions of Chapters G and H of the Cross-sectoral Standard Provisions and Sections 18, 19 and 22 of our Model Law on Cross-border Internet Activities and Virtual Worlds.

Chapter 6: Requirements for Algorithms

23. Algorithmic Design

23.1 Providers using algorithmic systems affecting minors must conduct Algorithmic Impact Assessments before deployment, prioritise minor safety and wellbeing in algorithm design, avoid optimising for engagement at the expense of wellbeing, implement safeguards against harmful content amplification, and enable user control over algorithmic recommendations.

23.2 Prohibited algorithmic practices for minors include: creating filter bubbles that isolate minors in harmful content, using persuasive design patterns to maximise screen time, implementing variable reward schedules mimicking gambling, personalizing content to exploit emotional vulnerabilities, and using dark patterns to manipulate behavior.

23.3 Required algorithmic features include user-controlled chronological content feeds, content diversity requirements to prevent echo chambers, circuit breakers for excessive use patterns, positive content promotion during vulnerable periods, and educational content prioritisation options.

24. Transparency Requirements

24.1 Providers using algorithmic systems affecting minors must disclose in age-appropriate language: how recommendations are generated, what factors influence content visibility, how user behavior affects future recommendations, options for modifying algorithmic behavior, and data inputs used for personalisation.

24.2 Providers using algorithmic systems affecting minors must publish annual transparency reports. Transparency reports must include:

1) Statistics on content recommended to different age groups;
2) Average time spent by minors on the platform;
3) Most common pathways to harmful content;
4) Effectiveness of safety interventions; and
5) Algorithm changes affecting minor users.

24.3. Providers using algorithmic systems affecting minors must:

1) Initiate [bi-]annual algorithm audits by an accredited certification body approved by the [REGULATORY AUTHORITY];
2) Provide the certification body and the [REGULATORY AUTHORITY] permanent access to data and systems necessary for the assessment of the algorithmic systems;
3) Publish audit summaries;
4) Adopt and execute corrective action plans; and
5) Ensure follow-up verification.

25. Addictive Design Prohibitions

25.1 Providers using algorithmic systems affecting minors may not implement:

1) Infinite scroll mechanisms;
2) Autoplay features for video content;
3) Push notifications encouraging immediate return;
4) Streak features creating usage pressure;
5) Social validation metrics (likes, views) visible to minors;
6) Loot boxes or gambling-like mechanics; and
7) Fear of missing out (FOMO) inducing features.

25.2 Providers using algorithmic systems affecting minors must implement the following wellness features: time spent reminders every thirty minutes, natural stopping points in content feeds, bedtime mode disabling notifications, weekend and vacation modes, and positive activity suggestions after extended use.

25.3 Providers using algorithmic systems affecting minors must demonstrate that algorithms and connected features:

1) Do not exploit psychological vulnerabilities;
2) Support balanced platform use;
3) Encourage offline activities;
4) Respect natural sleep patterns; and
5) Foster healthy social connections.

26. Safety Features and Tools

26.1 Providers using algorithmic systems affecting minors must implement the following safety features:

1) Communication Safety: Unknown contact filtering, image blurring for unsolicited photos, keyword filtering for harmful language, voice modulation for anonymous interactions, and screen recording prevention in private chats.
2) Crisis Intervention Tools: Prominent help buttons on all screens, automated crisis resource deployment, direct hotline connections, location-appropriate service referrals, and follow-up check mechanisms.

26.2 Providers using algorithmic systems affecting minors must also implement reporting and response systems including one-click reporting from any content, anonymous reporting options, detailed category selection, evidence preservation automation, and status tracking for reporters. Response standards include acknowledgment within 1 hour, initial review within 6 hours, decision communication within 24 hours, appeals process within 48 hours, and aggregate reporting to regulators monthly.

Chapter 7: Mental Health and Wellbeing

27. Mental Health Protections

Providers using algorithmic systems affecting minors must implement features to protect minor mental health:

1) Detection systems for content indicating:

    (A) Suicidal ideation or self-harm;
    (B) Eating disorders or body dysmorphia;
    (C) Severe anxiety or depression;
    (D) Substance abuse; and

(E) Social isolation or withdrawal;

2) Intervention protocols:

(A) Immediate resource display for crisis situations;
(B) Gentle check-ins for concerning patterns;
(C) Connection to professional help;
(D) Peer support group recommendations; and
(E) Parental notification for severe risks (with minor awareness).

28. Crisis Response Systems

28.1 Providers using algorithmic systems affecting minors must maintain 24/7 crisis response capabilities:

1) Trained crisis counselors available via chat;
2) Immediate connection to national suicide prevention hotlines;
3) Warm handoff protocols to local emergency services;
4) Follow-up check-ins after crisis events; and
5) Coordinated response with mental health professionals.

28.2 The following must trigger the crisis response:

1) Explicit statements of self-harm intent;
2) Goodbye messages or final statements;
3) Specific method searches or discussions;
4) Sudden behavior changes indicating risk; or
5) Reports from concerned peers or adults.

28.3 A crisis response must include:

1) Human review within fifteen (15) minutes of detection;
2) Automated support material deployment within five (5) minutes;
3) Professional contact within thirty (30) minutes;
4) Documentation for continuity of care; and
5) Post-crisis safety planning support.

29. Positive Digital Wellbeing

29.1 Providers using algorithmic systems affecting minors must actively promote digital wellbeing through:

1) Features encouraging balanced use;
2) Rewards for taking breaks;
3) Celebrations of offline achievements;
4) Social features promoting quality over quantity; and
5) Educational content about healthy habits.

29.2 Metrics and gamification must:

1) Avoid creating social pressure;
2) Reward positive behaviors, not just engagement;
3) Include wellbeing achievements;
4) De-emphasize vanity metrics; and
5) Promote authentic self-expression.

Chapter 8: Parental Rights and Family Support

30. Mandatory Parental Control Features

30.1 Providers using algorithmic systems affecting minors must offer parents/guardians comprehensive control dashboards including:

1) Activity Monitoring Features: Real-time usage statistics with visualizations, content categories accessed, accounts interacted with, searches performed (with privacy balance for teens), content uploaded or shared, and virtual goods purchased or acquired.
2) Control Mechanisms: Set daily, weekly, and monthly time limits, schedule access hours and blackout periods, block specific features or content categories, approve or deny contact requests, review and remove content posted by their child, and pause account access immediately.
3) Notification Systems: Alert parents to account creation or significant setting changes, attempts to access blocked content, new contact requests or friendships, potential safety concerns detected, unusual activity patterns, and policy violations by their child.

30.2 Parental controls must be easy to access and configure, available in multiple languages, functional across all devices, resistant to circumvention by minors, and respectful of age-appropriate privacy.

31. Progressive Autonomy Framework

31.1 Parental control levels shall be based on minor's age:

1) Age 0-6: Full parental control with no independent access;
2) Age 7-12: High parental oversight with limited independence;
3) Age 13-15: Balanced oversight with increasing privacy;

4) Age 16-17: Limited oversight with substantial autonomy.

31.2 Transition mechanisms must include gradual reduction of controls as minors age, notification to minors about parental monitoring levels, negotiation tools for families to agree on appropriate controls, clear sunset dates for specific restrictions, and documentation of autonomy milestones reached.

32. Family Digital Agreements

32.1 Providers must offer tools for creating family digital agreements that set mutual expectations for platform use, define consequences for misuse, establish communication protocols, create shared understanding of safety rules, and allow periodic review and updates.

32.2 Agreement templates must include time limit negotiations, content access permissions, privacy boundaries, emergency communication procedures, and digital wellness commitments.

33. Parental Education and Support

33.1 Providers must offer comprehensive parental education including platform safety feature tutorials, risk awareness training, digital parenting best practices, age-appropriate guidance materials, and crisis response resources.

33.2 Education delivery methods include interactive online courses, downloadable guides in multiple languages, video tutorials and webinars, community forums moderated by experts, and one-on-one support for complex situations.

33.3 Required topics include recognising signs of cyberbullying, identifying grooming behaviors, supporting healthy screen time habits, having difficult conversations about online safety, and balancing protection with independence.

PART III: IMPLEMENTATION AND ENFORCEMENT

Chapter 9: Auditing of Service Providers

Select provisions from Section “E. Certification by Private Conformity Assessment Bodies” of our Cross-sectoral Standard Provisions.

Set up provisions on what has to be audited in which frequency. Set-up rules encompassing subsidiaries and service providers.

Chapter 10: Public Ensurance of Compliance

Select provisions from Chapters E and J of our Cross-sectoral Standard Provisions. Section 42 of our Model Law on Cross-Border Internet Activities and Virtual Worlds.

Chapter 11: State Enforcement Powers

Select provisions from Chapter K of our Cross-sectoral Standard Provisions, from Section 41 of our Model Law on Cross-Border Internet Activities and Virtual Worlds and from our List of Powers and Obligations.

Chapter 12: Private Enforcement Support

Select provisions from Section L of our Cross-sectoral Standard Provisions and Chapter 4 of our Model Law on Cross-Border Internet Activities and Virtual Worlds,

Chapter 13: Dispute Resolution

Select provisions from Section M of our Cross-sectoral Standard Provisions.

Chapter 14: Sanctions

Select provisions from Section 65 of our Cross-sectoral Standard Provisions and our List of Sanctions and Accompanying Measures.

Chapter 15: International Cooperation

Select provisions from Chapter F of our Cross-sectoral Standard Provisions (namely for the recognition of foreign auditing) and Sections 43 and 44 of our Model Law on Cross-Border Internet Activities and Virtual Worlds.

Chapter 16: Final Provisions / Miscellaneous

Select provisions from Chapter Q of our Cross-sectoral Standard Provisions.

Further chapters or even an entire Part IV could establish protection and support systems for minors e.g. on the basis of our Model Provisions on Vulnerable Persons and Persons with Special Needs.