

# App Ratings & Access Control Framework Proposal

Protecting  
Kiwi Kids  
Online



Developed and presented by Our Kids Online  
November 2025

# Contents

Executive Summary	02
Core Pillars	03
Alignment with the NZ Child Internet Safety Act	04
Classification Criteria	05
Legislative Addendum (FVPCA 1993 - Part 9A)	06
Implementation Road Map	06
Summary of Benefits	07

# Executive Summary

Aotearoa New Zealand will have strong protections for children with the proposed NZ Child Internet Safety Act (NZCISA), but a critical gap remains: messaging, gaming and social media apps are not subject to a mandatory age-based rating system. This proposal establishes a national app-classification framework—mirroring the familiar R18, R16, PG, G schema—to keep under-18s from encountering illegal and harmful content and engagement mechanics.

What makes this version unique is that enforcement does not rely on platforms or global app stores. Instead, it uses filtered SIM cards, device-level filtering, and public network controls mandated under NZCISA. This local infrastructure-based model ensures child safety is fully enforceable under New Zealand jurisdiction.

# Core Pillars

## App Classification System

- Administered by a specialised team with clear age-based classifications, reducing risk of illegal and R18 content being accessed by those under 18 years of age.
- Triggers based on both content and design (e.g., dark patterns, unmoderated chat, infinite scroll).
- Includes algorithmic risk and addictive design as formal criteria.

## Enforcement through NZ Infrastructure (Not App Stores)

- Filtered SIM cards: Required for all minors under 18. These block traffic to R18-rated app endpoints, making them unusable.
- Device-based filtering: Pre-installed and non-removable filters on all under-18 devices, blocking installation or launch of R18 apps.
- Private Wi-Fi filtering: Implementation of DNS/IP filtering on all ISP-provided routers for residential users.
- Public Wi-Fi & school filtering: Enforced DNS/IP filtering to prevent R18 app use.
- Parental responsibility: Guardians must take all reasonable steps to ensure filtered SIMs/devices are used and take all reasonable steps to prevent circumvention.

## Transparent Compliance System

Existing age labels in global app stores are often misaligned with NZ standards.

# Alignment with the NZ Child Internet Safety Act

This framework operationalises the NZCISA by addressing gaps in app regulation that filtering alone cannot close. Together, they form a layered, proactive model of protection:

NZCISA Provision	App Ratings Framework Complement
Filters ISP/device level access	Adds pre-market classification & app visibility controls
Focus on illegal/harmful content	Covers Illegal and R18 content on apps
SIM filtering mandates	Leverages SIM filtering to enforce app access control

# Classification Criteria

The development of the classification criteria will be led by a Digital App Safety Team (DAST), whose members will be formally appointed as Inspectors of Publications under relevant legislation. This initial team must include individuals with proven experience in assessing or managing exposure to objectionable digital content. DAST will consult closely with the Office of Film and Literature Classification (OFLC) to ensure alignment with existing classification standards and legal frameworks.

In determining age classifications for apps accessible to individuals under 18, DAST may also engage with a range of expert stakeholders, including child safety NGOs, digital psychologists specialising in child and adolescent media use, DIA Digital Safety Teams (including those focused on violent extremism and child sexual exploitation), the New Zealand Principals' Federation, and youth representatives. The CVoD labelling system per Films, Videos, Publications and Classifications Act 1993 may be referred to as a guide. 'Dip-testing' ensures mahi is manageable.

The table below indicates factors that must be taken into account.

## Content & Systemic Risk Criteria

Risk Category	Description
1. Addictive Design	Features like infinite scroll, reward systems, or dark patterns that are designed to exploit attention and encourage compulsive use.
2. Algorithmic Harm	Exposure to harmful or age-inappropriate content driven by recommendation algorithms or curated feeds.
3. Explicit Material	Presence of nudity, pornography, or graphic violence — including user-generated content.
4. Cyberbullying	Lack of effective moderation tools, reporting mechanisms, or safeguards against peer abuse.
5. Predators & Grooming Risk	Features such as unmonitored chat, anonymous direct messaging, or absence of age verification that increase the risk of exploitation.
6. Unmoderated User-Generated Content	User-generated content that bypasses platform monitoring or filtering systems during upload.

# Legislative Addendum (FVPCA 1993)

Consideration to be given to any necessary addendum to the Films, Videos and Publications Classification Act 1993 that would be required or deemed appropriate to facilitate the implementation of the classification of apps.

## Implementation Roadmap

Phase	Key Outputs
Planning & funding	Outline, initial budget
Team formation & onboarding	Establishment of Digital App Safety Team
Consultation Process	Consultation with OFLC and others
Classifications	Criteria agreed and implemented
Portal MVP & standards	Pilot rating system
Legislation introduced	FVPCA amendment bill tabled if required
Enforcement tools live	SIM/device filtering in market
Full launch	Ratings enforced nationwide

# Summary of Benefits

- Child Protection: Blocks explicit, violent, objectionable, addictive, and algorithmically harmful content at the infrastructure level.
- National Sovereignty: Does not depend on foreign tech firms to enforce NZ law.
- Clarity for Caregivers: SIM and device-based controls create a clean, enforceable barrier.
- Regulatory Leadership: Builds a precedent for systemic digital harm prevention.