

Mutual Understanding Content Policy - Version 3 (Globally Adaptive)

Purpose

Create a consistent yet adaptable approach to content moderation that supports trust, legal compliance, and cultural sensitivity across different societies.

Adaptive Hybrid Review

AI core review with a globally shared base rule set. Local human review interprets flagged content per local laws and language. Cross-border cases escalate to a Global Safety Board.

Tiered Content Handling

Tier 1: Immediate takedown for criminal or dangerous content. Tier 2: Contextual display with notices. Tier 3: Advisory labels for legal but sensitive content.

Flexible Compliance

Jurisdictions may customize notice wording, escalation steps, and advisory categories. Core principles-appeals, transparency reporting, and audit-remain mandatory.

Accountability & Oversight

Local Compliance Officer and annual regional audit by a neutral partner. A global, aggregated report provides anonymized trends.

User Engagement

Local appeals, right to explanation, and education materials tailored to local context.

Gradual Rollout

Phase 1: Core AI + local compliance. Phase 2: Public transparency and education. Phase 3: Cross-border audits and public Q&A.