

Content Moderation

Detect Harmful Content in Indian Languages —
Before It Reaches Your Users

JioCloud Content Moderation helps you spot Inappropriate content written in Indian languages or mixed formats — before it is published.

The Challenge

- People often write in more than one language — mixing Hindi, Tamil, English, or writing phonetically.
- Most tools miss Inappropriate posts when they include slang, local expressions, or informal tone.
- If this content goes unchecked, it can lead to user complaints, legal issues, or damage to your platform's image.
- Manual review doesn't scale well, and many platforms don't have tools built for Indian language patterns.



The JioCloud Solution

Struggling with diverse Indian content? JioCloud Content Moderation uniquely detects abusive or inappropriate material across Hinglish(Hindi + English), regional scripts, and mixed formats. Benefit from real-time checks, easy API integration, a dedicated Playground for pre-live testing, and empower your teams with independent, cost-effective, pay-per-review management.

Key Features

- **Multilingual offensive content detection**
Flags explicit, abusive, or hateful content in Indian scripts and mixed-language inputs.
- **Context-aware NLP(Natural Language Processing)**
Understands regional slang, sarcasm, and informal phrasing.
- **Real-time flagging**
Detects violations before content is posted or published.
- **Playground for testing and validation**
Test real samples and review labels instantly.
- **API(Application Programming Interface)-first architecture**
Add moderation to UGC(User-Generated Content) systems, chat apps, or content platforms easily.
- **Complete self-service**
Manage provisioning, usage, and credits from a unified dashboard.
- **Pay-as-you-go pricing**
Affordable and metered billing with no fixed contract.

What You Gain

- Reduce exposure to inappropriate or abusive content.
- Moderate regional and code-mixed text with cultural accuracy.
- Lower manual review effort and moderation costs.
- Protect brand safety and enable policy compliance.
- Improve trust and safety across user communities.
- Integrate quickly with APIs and test in a live Playground.



Use Cases in Action

Social platforms and messaging apps

Flag offensive posts or messages in real time to reduce hate speech and build safer communities.

eCommerce reviews and Q&A

Filter abusive or fake reviews to protect buyer experience and seller reputation.

Gaming and live chat environments

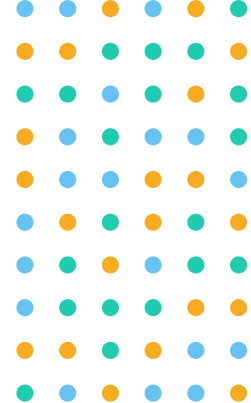
Detect harassment and toxic behavior in fast-moving chats across multiple languages.

Ideal for

- Social media and messaging platforms
- eCommerce and marketplace apps
- Gaming platforms with chat features
- Live-streaming and video portals
- Forums, communities, and UGC platforms

Why JioCloud

- **Trained for Indian languages**
Covers local scripts, dialects, and mixed-language content.
- **Live playground**
Test and fine-tune moderation before deploying.
- **Simple API integration**
Add real-time filtering with minimal development effort.
- **Full self-service control**
Provision, manage, and scale the service without support tickets.
- **Transparent billing model**
Pay only for what's analysed — with no lock-ins or surprise costs.



Ready to Power Smarter Storage?

Run faster, scale instantly, and recover with confidence - with JioCloud Block Storage.

Reach out at jpl.cloudsales@ril.com or visit (website) to get started.

