Science & Technology · GS - III

Evaluate the role of artificial intelligence in content moderation, with particular reference to child safety and freedom of expression on digital platforms.

EXAM YEAR	PAPER	SUBJECT
DAILY PRACTICE 27/11/2026	GS - III	Science & Technology

MARKS	WORD LIMIT
15	~250 words

Answer Hint

Artificial intelligence, content moderation, child online safety, intermediary liability, freedom of expression, algorithmic accountability, digital governance, grievance redressal

Model Answer

The removal of millions of videos from India by YouTube during April–June 2026 highlights the growing role of automated systems in regulating user-generated content. AI-based moderation has become essential for managing the scale and speed of digital communication.

Contribution to digital safety

1. **Large-scale detection:** Machine-learning systems can scan enormous volumes of videos, images and text much faster than manual moderation.
2. **Child protection:** Automated detection can help identify child sexual abuse material, exploitative content and attempts at online grooming.
3. **Rapid intervention:** Automated systems can flag harmful material before it gains wider circulation.
4. **Consistency:** Standardised detection mechanisms can improve enforcement of platform policies across large user communities.

Limitations and ethical concerns

1. **Contextual errors:** Satire, educational material, journalism and regional-language expressions may be incorrectly classified.

2. **Algorithmic bias:** Systems trained on limited datasets may perform unevenly across languages and cultural contexts.
3. **Lack of transparency:** Users may not understand why content was removed or how automated decisions were reached.
4. **Over-removal:** Excessive moderation can restrict legitimate expression and affect independent creators.

Towards accountable moderation

India should encourage transparent platform reporting, accessible appeals, independent audits and appropriate human review of consequential decisions. Platforms must also strengthen privacy safeguards while responding to child-safety threats.

The Information Technology framework should be implemented in a manner that combines intermediary accountability with constitutional protections for free expression.

Ultimately, AI should function as an aid to responsible content governance rather than an unquestionable substitute for human judgement. Its legitimacy depends on accuracy, proportionality and effective grievance redressal.

BECOME IAS
AI POWERED UPSC PREP