

ROLE REFERENCE SHEET

Trust and Safety Manager

Trust and Safety Manager · Platform Policy & Integrity

Owns the policies, processes, and enforcement that keep a platform safe, from content moderation to abuse response and regulatory obligations.

WHAT THE ROLE OWNS

- Content & conduct policy
- Moderation operations
- Enforcement & appeals
- Platform integrity & abuse
- Regulatory & transparency reporting
- Responsible AI in moderation
- Safety metrics
- Moderator training & wellbeing

FRAMEWORKS

- Platform policy standards
- Content moderation frameworks
- Digital Services Act (DSA)
- Online safety regulation
- Responsible AI principles

TYPICAL CERTS

- Trust & safety experience
- AIGP a plus for AI-heavy platforms

6–10 yrs

Experience

DSA

Regulatory driver

Policy

Core focus

**Scan for the free JD template**ai-governance-jobs.com/templates

AGJ, the AI governance job board
Every role hand reviewed since 1999