



Trust & Safety Assessment Tool for Edtech Companies



This tool is designed to help EdTech stakeholders assess whether their policies, systems, and practices align with best-in-class trust and safety (T&S) standards.

For each question, select: **YES, PARTIALLY, or NO.**



I. Governance & Policy Framework

Assesses the existence of clear trust and safety policies and accountability structures.

Question YES PARTIALLY NO

- Do you have a publicly available trust and safety policy?
- Who is responsible for trust and safety oversight (named officer or team)?
- Are your T&S policies reviewed at least annually?
- Are your operations compliant with COPPA, GDPR, and applicable local child safety laws?





II. User Protection Mechanisms

Evaluates features that protect users from harassment, abuse, and harm.

Question YES PARTIALLY NO

Can users (especially students) report abusive behavior directly within the platform?

Can harmful content detected and manage through real-time AI or human moderation?

Are bullying or harassment filters or detection systems in place?

Is there a defined escalation process for safety incidents?

What is the average response time for resolving user safety incidents?

(Goal: within 24–48 hrs)





III. Data Privacy and Security

Checks how user data, especially of minors, is protected.

Question YES PARTIALLY NO

Is user data encrypted in transit and at rest?

Do you have consent protocols in place, especially for underage users and their guardians?

Is only the minimum necessary data collected from users?

Are your data retention and deletion policies transparent and user-controlled?

Are systems regularly audited for security vulnerabilities?





IV. Inclusive & Equitable Design

Assesses whether the platform is inclusive and respects diverse users.

Question YES PARTIALLY NO

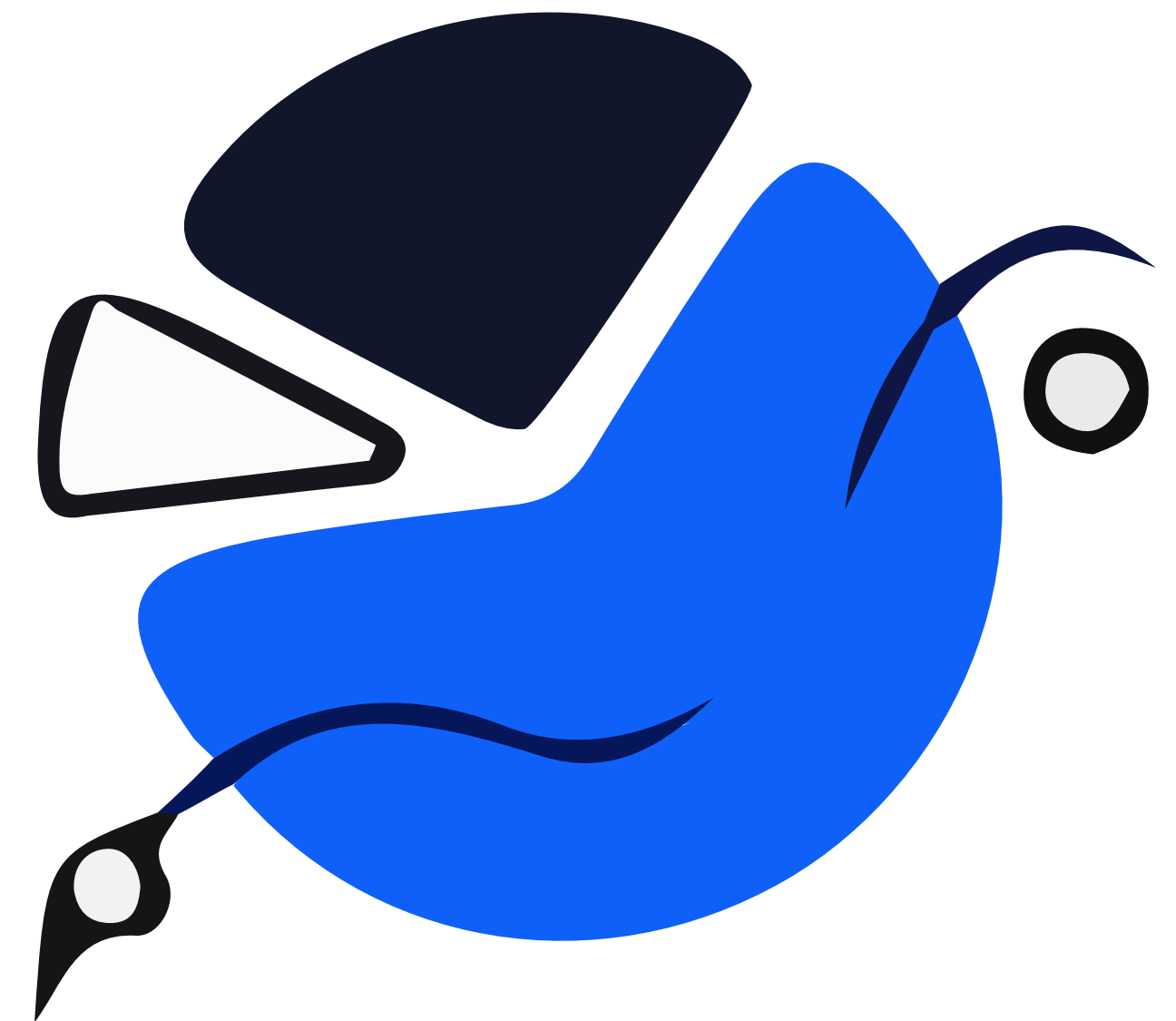
Does your platform meet accessibility standards (e.g., WCAG)?

Is your content culturally inclusive and appropriate for different age groups?

Do you provide language options for non-English-speaking users?

Are your moderation practices designed to reduce bias and promote fairness?

Do moderators receive training on implicit bias and cultural sensitivity?





V. Education & Support

Evaluates the educational support provided to users on digital safety and rights.

Question YES PARTIALLY NO

Are users onboarded with clear guidance on safety settings and privacy tools?

Do you offer online safety education resources for parents, teachers, and learners?

Are dedicated support channels (e.g., email, helpdesk, chat) available for users in distress?

Are your support staff trained to respond to online safety incidents effectively?





VI. AI and Algorithmic Transparency

Looks at the use of AI in moderation and personalization, and how biases are addressed.

Question YES PARTIALLY NO

Does your platform use AI for content moderation or personalized learning experiences?

Are users informed when AI decisions affect their content or experience (e.g., flagging, bans)?

Do you conduct regular bias and fairness audits of your algorithms?

Can users appeal or override AI-based decisions?



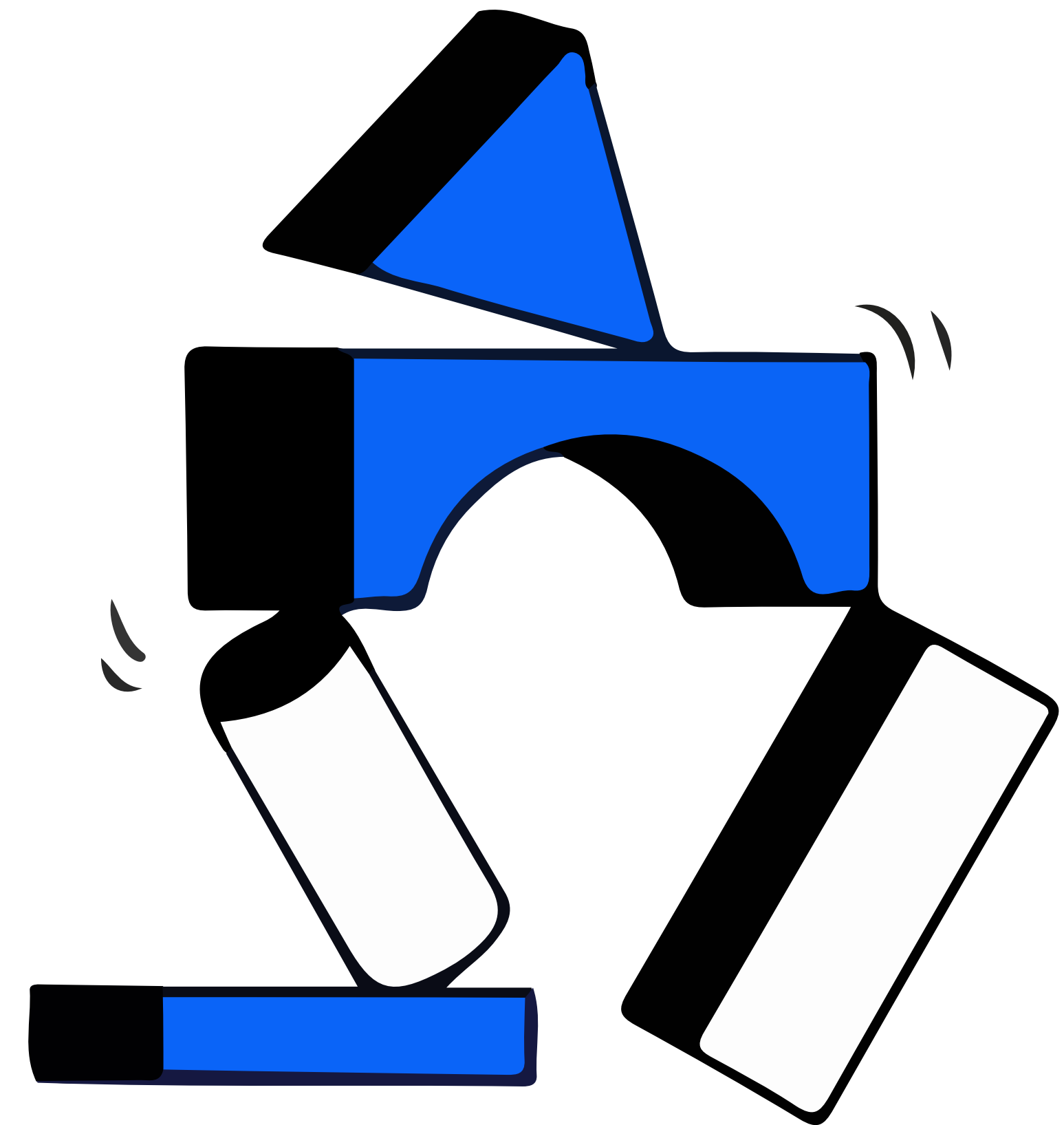


VII. Incident History and Transparency

Evaluates transparency in reporting safety incidents and institutional learning.

Question YES PARTIALLY NO

Have you published transparency or trust and safety reports in the past year?
Do you conduct internal reviews after major user safety incidents?
Can you provide examples of safety improvements made in response to user harm or complaints?



Contact:

 info@trustedtechafrica.com

 trustedtechafrica.com



Trusted Tech

Africa Ltd