

Íntegra do posicionamento

TikTok

Tamanho das equipes e condições de trabalho

- Em março, reestruturamos uma equipe específica dentro da organização global de Confiança e Segurança, o que nos permitiu aprimorar ainda mais os processos de garantia de qualidade. Nenhum moderador de conteúdo foi afetado nesta ocasião. Seguimos, desde então, contratando globalmente para nossa equipe de Segurança.
- No mundo, são mais de 40.000 profissionais de Confiança e Segurança trabalhando para proteger a comunidade TikTok, que conta com moderadores brasileiros que entendem o contexto local. Temos a expectativa de investir, apenas neste ano, mais de US\$ 2 bilhões em esforços de Confiança e Segurança.
- Nós nos esforçamos para promover um ambiente de trabalho atencioso para os funcionários, especialmente para os profissionais de confiança e segurança. Usamos uma abordagem baseada em evidências para desenvolver programas e recursos que contribuam para o bem-estar psicológico dos moderadores.
- Enfatizamos principalmente medidas de cuidado preventivas para aumentar o bem-estar por meio de ferramentas, apoio e treinamento oportunos, desde o recrutamento e integração até o término do contrato de trabalho, que ajudam a promover a resiliência e aumentar o bem-estar.
- Isso pode incluir ferramentas e recursos que visam possibilitar que os funcionários controlem a exposição a conteúdos visuais ao avaliar ou moderar conteúdos, como aplicação de escala de cinza, mudo e desfoque; treinamento para gerentes para ajudá-los a identificar quando um membro da equipe pode precisar de maior apoio ao seu bem-estar; e apoio clínico e terapêutico.
- Também fornecemos aos nossos funcionários da área de confiança e segurança a possibilidade de se associar à Trust and Safety Professional Association (TSPA). Essa associação permite que eles acessem recursos para desenvolvimento profissional, participem de workshops e eventos e conectem-se a uma rede de pares no setor. A nossa prioridade é resguardar o bem-estar de nossos colaboradores.

Transparência

- Queremos que todos saibam quais são as nossas regras e padrões e como os aplicamos. Procuramos tornar fácil o acesso às nossas políticas e práticas, aplicando-as de forma consistente e igualitária. Nossos esforços estão registrados no [Centro de Transparência](#).
- O TikTok publica trimestralmente [Relatórios de Aplicação das Diretrizes da Comunidade](#), tornando transparente nossos esforços de moderação.
- Nestes relatórios é possível ver volume vídeos e contas removidas, com recortes por políticas dentro das nossas Diretrizes. Além da taxa de respostas após a denúncia de um usuário.
- Também estão expostos volume de remoções por mercados e distribuição de idioma entre nossa equipe de moderadores.

- Aqueles que desejarem podem também fazer download de todos os dados incluídos nestes relatórios para fazerem suas próprias análises.
- Recentemente lançamos um relatório exclusivo sobre nossos esforços de [combate a operações de influência disfarçadas](#), que inclui redes criadas com o intuito de manipular nossa comunidade.

Inteligência artificial na moderação

- Nossa [abordagem de moderação de conteúdo](#) combina tecnologia avançada com revisão humana para identificar e remover rapidamente material que possa violar as diretrizes da comunidade. A equipe global conta com mais de 40 mil pessoas dedicadas à segurança.
- Os vídeos carregados no TikTok primeiro são avaliados por nossa tecnologia de moderação automatizada, que visa identificar conteúdo que viola as Diretrizes da Comunidade. Esses sistemas examinam uma variedade de sinais em um conteúdo, como palavras-chave, imagens, títulos, descrições e áudio.
- Se não for identificada nenhuma violação, ele será disponibilizado para visualização na plataforma.
- Se for encontrada uma possível violação, o sistema de moderação automatizado o encaminhará às nossas equipes de segurança para avaliação adicional ou o removerá automaticamente se houver um alto nível de confiança de que o conteúdo viola as Diretrizes da Comunidade.
- Essa remoção automatizada é aplicada quando as violações são mais evidentes, como nudez ou segurança de jovens.
- Sempre investimos na melhoria da precisão de nossos sistemas de moderação automatizados para que possamos remover com maior eficácia e em grande escala conteúdo violador, bem como reduzir o número de remoções incorretas; se os usuários acreditarem que cometemos um erro, eles poderão recorrer da remoção do respectivo conteúdo.
- Nosso [Relatório de Aplicação das Diretrizes da Comunidade](#) traz informações sobre qual o volume de conteúdo foi moderado por automação. Esse volume acompanha o volume total de remoções trimestralmente.