# Everest Group®

# Key Tech Away: Technology Providers in the Trust and Safety Landscape

April 2023: Complimentary Abstract / Table of Contents

# Our research offerings

This report is included in the following research program(s):
Trust and Safety

- Amazon Web Services (AWS)
- Application Services
- Artificial Intelligence (AI)
- Asset and Wealth Management
- Banking and Financial Services Business Process
- Banking and Financial Services Information Technology
- Catalyst™
- Clinical Development Technology
- Cloud and Infrastructure
- Contingent Staffing
- Contingent Workforce Management
- Customer Experience Management Services
- CX Excellence
- CXM Technology
- Cybersecurity
- Data and Analytics
- Digital Adoption Platforms
- Digital Engineering Services
- Digital Services
- Digital Workplace
- Employee Experience Management (EXM) Platforms
- Employer of Record (EOR)
- Engineering Services
- Enterprise Platform Services
- Exponential Technologies

- Finance and Accounting
- Financial Services Technology (FinTech)
- GBS Talent Excellence
- Global Business Services
- Google Cloud
- Healthcare Business Process
- Healthcare Information Technology
- HealthTech
- Human Resources
- Insurance Business Process
- Insurance Information Technology
- Insurance Technology (InsurTech)
- Insurance Third-Party Administration (TPA) Services
- Intelligent Document Processing
- Interactive Experience (IX) Services
- IT Services Excellence
- IT Services Executive Insights™
- IT Talent Excellence
- Life Sciences Business Process
- Life Sciences Commercial Technologies
- Life Sciences Information Technology
- Locations Insider™
- Marketing Services
- Market Vista™
- Microsoft Azure

- Modern Application Development (MAD)
- Mortgage Operations
- Multi-country Payroll
- Network Services and 5G
- Oracle Services
- Outsourcing Excellence
- Pricing Analytics as a Service
- Process Mining
- Process Orchestration
- Procurement and Supply Chain
- Recruitment
- Retail and CPG Information Technology
- Retirement Technologies
- Revenue Cycle Management
- Rewards and Recognition
- SAP Services
- Service Optimization Technologies
- Software Product Engineering Services
- Supply Chain Management (SCM) Services
- Sustainability Technology and Services
- Talent Genius™
- Technology Skills and Talent
- Trust and Safety
- Value and Quality Assurance (VQA)

If you want to learn whether your organization has a membership agreement or request information on pricing and membership options, please contact us at **info@everestgrp.com**

Learn more about our
**custom research capabilities**

Benchmarking

Contract assessment

Peer analysis

Market intelligence

Tracking: providers, locations, risk, technologies

Locations: costs, skills, sustainability, portfolios

# Contents

For more information on this and other research published by Everest Group, please contact us:

**Manu Aggarwal,** Partner

**Abhijnan Dasgupta,** Practice Director

**Aseem Nousher,** Senior Analyst

**Shubhali Jain,** Analyst

# Contents

# Contents

# Introduction

As the volume and variety of content on the internet continue to surge, the universe of toxic content is also expanding rapidly. Enterprises are compelled to moderate content in real-time and provide adequate safety features in platforms to provide a safe experience to users. An effective tech-based solution that can fast-track the moderation process and minimize the requirement for a moderator going through the toxic content is the need of the hour. However, the current third-party technology for T&S is still not advanced, and hence T&S technology only provides a first layer of filtering while a significant portion of content moderation is still being handled by human moderators.

The role of technology solutions in T&S is bound to grow in the long run. The T&S technology market is undergoing an evolutionary phase, with certain technology providers now offering innovative solutions to detect evolving risks. The industry is gradually moving toward a holistic, end-to-end platform-based solution that can safeguard the digital space from a wide range of threats. To achieve this future state, closer collaboration among all stakeholders, including technology providers, service providers, and enterprises, is imminent.

In this research, we present detailed profiles of 19-technology providers. Each provider profile presents a snapshot of its leadership, funding and growth, technology capabilities, market recognition, partnership ecosystem, and innovation.

**In this research, we focus on:**

- Current state of the T&S technology market
- Evolving trends in the market
- Future state of T&S technology market
- Technology provider landscape
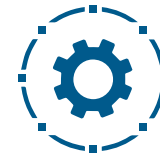- Profiles of technology providers in the T&S space

**Scope of this report**

**Geography**
Global

**Industry**
All industry verticals

**Profiles**
19 T&S technology providers

# Overview and abbreviated summary of key messages

This report examines the T&S technology landscape and provides a detailed analysis of the current T&S technology market state, evolving market trends, the future state of the T&S technology market, and the technology provider landscape along with the profiles of 19 technology providers

**Some of the findings in this report, among others, are:**

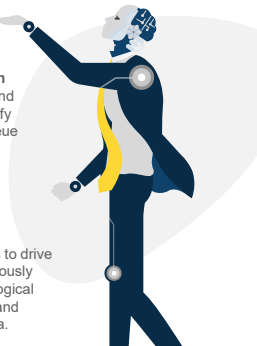| | |
|---|---|
| **Current T&S solutions are largely human-focused** | • Explosive growth in content requires scalable T&S solutions<br>• Enterprises are compelled to reach out to multiple providers to meet all their T&S technology needs<br>• Technology investments lag growth in emerging content types such as the metaverse and AI-generated content<br>• Technology development needs to keep pace with the evolving regulatory landscape<br>• Human intervention continues to be prevalent at each step of moderation |
| **Light at the end of the tunnel** | • An evolution is underway in the market, marked by the emergence of new technologies such as contextual API, increasing collaboration between service providers and technology providers to deliver a combined solution, and increasing integration of multiple content moderation solutions to create an end-to-end solution |
| **It is time to R.A.I.S.E. the game** | • In the future, a holistic, platform-based T&S solution that caters to multiple work types and content formats will become a minimum requirement, in which generative AI has the transformative power to become the cornerstone technology in T&S platforms<br>• It will require increased collaboration and co-investment from service providers, technology providers, and enterprises<br>• This will give rise to an ecosystem expansion in which The demand will shift from content moderation to the safety of the digital ecosystem |
| **Technology provider landscape** | • Technology providers belong to seven major work types: content and ad moderation, data annotation, platform safety, language translation and transcription, risk intelligence, search relevance, and wellness<br>• This report features two technology provider categories: the largest T&S technology providers and a selection of niche providers that have unique value propositions |

# This study offers four distinct chapters providing a deep dive into key aspects of T&S technology market; below are four charts to illustrate the depth of the report

## Technologies shaping T&S solutions

**Natural Language Processing (NLP)** gives computers the ability to read, comprehend, and interpret human languages. This includes text processing, retrieval, summarization, translation, and sentiment analysis.

**Artificial Intelligence (AI) / Machine Learning (ML)** use algorithms and statistical models to automatically identify and remove inappropriate or harmful content and analyze large amount of content to detect threats.

**Optical Character Recognition (OCR)** combines text detection and transcription capabilities to identify text embedded in images and queue it for automated moderation.

**Computer Vision** leverages AI to enable automatic extraction, analysis, and understanding of useful information from digital images and visual inputs.

**Digital Hash** creates digital signature for content to detect illegal/harmful images such as Child Sexual Abuse Material (CSAM) and copyright infringement when reshared.

**Deep Learning** uses neural networks to drive human-like conclusions by continuously evaluating data with a predefined logical structure to recognize, classify, and characterize objects within data.

**Generative AI** is a type of artificial intelligence that can create new and diverse data forms, including images, audio, text, and videos. Generative AI in trust and safety is a double-edged sword: on one hand, it introduces potential risks; on the other hand, it holds promising potential to improve and strengthen trust and safety solutions.

## Current state of the T&S technology market

**Global monthly consumer internet data traffic**
(2016-22; in exabytes/month) and the prominent content types that constitute this volume

Legend: Internet video | Web, email, and data | Online gaming

Traffic (does not include file sharing):
| Year | Internet video | Web, email, and data | Online gaming |
|---|---|---|---|
| 2016 | 42 | 9 | 1 |
| 2017 | 56 | 12 | 1 |
| 2021E | 184 | 27 | 11 |
| 2022E | 240 | 31 | 15 |

*ILLUSTRATIVE*

To handle this sheer volume of content, a T&S solution is required for ensuring platform safety, maintaining community standards, and preventing the spread of harmful threats, and which:
- Understands local and cultural nuances to handle the complexity of content
- Is scalable to cater to a range of content formats, including emerging formats, to effectively handle harmful content
- Works on a contextual approach to accurately identify and deal with vast amounts of content
- Has the ability to recognize and counteract tactics used by users to bypass moderation filters

**Generative AI is poised to exponentially expand content production. According to experts, by 2025, over 90% of the digital content could be generated by AI.[2]**

### The majority of scalable solutions are owned by enterprises

**amazon** — Amazon Rekognition Content Moderation tool uses ML to automate and streamline moderation, detect inappropriate content, and keep users safe with customizable API.

**Microsoft** — Microsoft Azure Content Moderator provides AI-based, customizable, scalable, real-time moderation tools along with pre-built workflows for a comprehensive solution.

**Alibaba Cloud** — Alibaba's content moderation tool uses AI and big data to monitor content. It is highly scalable and customizable to quickly recognize and respond to adverse content.

**Meta** — The Hasher-Matcher-Actioner system stops the spread of illegal and harmful content. It builds on Meta's previous open-source image and video matching software.

## Emerging T&S technologies

**Contextual AI** examines the context in which user-generated content is presented in order to identify and respond to prohibited activities in real-time

**Understand need**
Identify intent(s) and entities in the request using NLU models, automatic speech recognition, and deep learning-based approaches

**Detect plagiarism**
Streamline the plagiarism detection process, improve accuracy, reduce false positives, and provide enhanced reporting capabilities

**Improve user experience**
Platforms can now provide their users with an option to filter out AI-generated content, increasing the credibility of the platform

**Demographic API** uses a deep-learning model to identify the gender, age, and multicultural appearance of people's faces
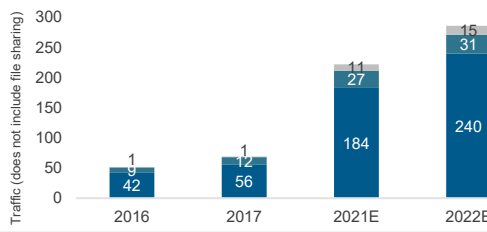
**Age gate enforcement**
Recognize underage users on dating or social media platforms to enforce age-based restrictions

**User authentication**
Scan user photos to verify users based on age and gender on social media, dating apps, or chats

**Remove race bias**
The inclusion of datasets of multi-cultural appearances results in no bias moderation toward any culture

**AI-generated content detection tool** harness the power of AI to classify between AI-generated and human-written text and images

**First level of screening**
Detect AI-generated content that is not easily recognizable to human eyes, reducing moderators' time and increasing the accuracy of content moderation

**Detect plagiarism**
Streamline the plagiarism detection process, improve accuracy, reduce false positives, and provide enhanced reporting capabilities

**Improve user experience**
Platforms can now provide their users with an option to filter out AI-generated content, increasing the credibility of the platform

## Technology provider landscape

**T&S landscape**

Focus of this research

**Integrated solution providers**

**Primary role:**
Provide integrated human + tech T&S solutions. They also provide value-added services such as policy advisory

**Value proposition:**
- Scalable solutions across industries
- Global reach

**Product companies**

**Primary role:**
Leveraged primarily for their technology solutions, available as a stand-alone product/solution for independent licensing

**Value proposition:**
- Technology expertise
- Industry- and content-specific solutions
- Some of them also offer stand-alone human moderation support

**Seven work types**

Content and ad moderation | Platform safety | Language translations and transcription | Risk intelligence | Search relevance | Wellness | Data annotation

Everest Group®

# The Key Tech Away: Technology Providers in the Trust and Safety Landscape report features 19 technology provider profiles

Key Tech Away: Technology Providers in the Trust and Safety Landscape

## Technology Provider 1 | company profile (page 1 of 2)

**Company overview**

Technology Provider 1 is a global content moderation technology provider that helps business and organizations keep their online platforms safe and secure. moderation solutions possible, helping business and organization maintain safe and engaging online communities. Over XYZ, serves clients in more than XY

**Fact sheet**

| Headquarters | Number of employees | Estimated annual revenue (US$ million) |
|---|---|---|
| XYZ | XYZ | XYZ |

| Key executives | Key investors |
|---|---|
| • XYZ, Chief Executive Officer<br>• XYZ, Chief Technologist<br>• XYZ, Chief Operating Officer | XYZ |

| Dimension | Everest Group's perspective |
|---|---|
| Vision and strategy | Future oriented vision and proactive strategy are key to their success as a leading content moderation |
| Trust and safety focus areas | Content moderation provider covering all content types and threats |
| Technology capability | Leverages advanced machine learning and natural language processing capabilities to flag and rem |
| Innovation and investments | Continued investment in product development, and technology, as well as notable partnerships with |
| Client spread / industry coverage | Caters to small, midsize, and large enterprises spanning social media, streaming and gaming, eCom |
| Market success | Ability to provide highly effective, scalable and efficient content moderation solutions that meet the e |

**Awards and recognitions**

• 2022: named as a "Top 10 AI Solution intelligence in content moderation
• 2022: received the "Innovator of the Ye AI powered solutions

**Recent platform functionality enha**

• 2023: added new functionality that allow inappropriate or harmful content based
• 2022: made significant updates to their platform's various features
• 2022: launched new capabilities allowi or harmful images in real- time

Everest Group®  Proprietary & Confidential. © 2023, Everest Global, Inc. | EGR-2023-67-R-5934

---

Key Tech Away: Technology Providers in the Trust and Safety Landscape

## Technology Provider 1 | solutions, partnerships, and strategy (page 2 of 2)

**Capability overview**

• **XYZ**
  It is an automated content moderation solution using machine learning and natural language processing
• **XYZ**
  It is a multilingual content moderation solution for global businesses and organizations
• **XYZ**
  It is a comprehensive suite of content moderation tools to ensure that businesses and organizations can maintain a safe and engaging online environment for all users
• **XYZ**
  It can detect and remove inappropriate or harmful images in real-time
• **XYZ**
  It can detect and remove inappropriate content in real-time, ensuring that live streams are safe and appropriate for all viewers
• **XYZ**
  It can detect and remove content that infringes on a brand's intellectual property rights, ensuring that businesses and organizations can protect their brand identity and reputation
• **XYZ**
  It monitors user behavior and engagement data alongside content moderation efforts

**Future strategy**

• Expand its presence in emerging markets such as Southeast Asia, Africa, and Latin America
• Continue to invest heavily in research and development to stay ahead of the curve in the rapidly evolving field of content moderation technology
• Establish strategic partnerships with other leading technology companies to expand its reach and provide integrated content moderation solutions

**Client portfolio**

XYZ

**Key partnerships**

XYZ

**Case studies**

**XYZ helped a social media platform, XYZ, to moderate in real time**
By removing inappropriate and offensive content in real-time XYZ helped XYZ to achieve a higher engagement rates, and improved brand reputation

**XYZ helped a eCommerce platform, XYZ, increase its ratings**
XYZ helped XYZ improve the quality and credibility of their product listings by removing fake reviews and fraudulent content

Everest Group®  Proprietary & Confidential. © 2023, Everest Global, Inc. | EGR-2023-67-R-5934

# Research calendar
## Trust and Safety

| Reports title | Release date |
|---|---|
| Taming the Hydra: Trust and Safety (T&S) in the Metaverse | July 2022 |
| Wisdom of the Crowd: The Gig Delivery Model in Trust and Safety (T&S) | August 2022 |
| Braving New Frontiers: Trust and Safety State of the Market Report 2022 | August 2022 |
| What's the Deal? Evolving Sourcing Priorities in Trust and Safety | November 2022 |
| Content Moderation Catch 22: Walking the Regulatory Tightrope | December 2022 |
| A Beautiful Mind: Ensuring Moderator Well-being | December 2022 |
| Trust and Safety Services PEAK Matrix® Assessment 2023 | March 2023 |
| **Key Tech Away: Technology Providers in the Trust and Safety Landscape** | **April 2023** |
| Press Play on Player Experience: Level up Beyond Game Design | Q2 2023 |
| Trust and Safety Service Provider Compendium 2023 | Q2 2023 |
| Trust and Safety Services State of the Market 2023 | Q2 2023 |
| Content Moderator of the Future | Q3 2023 |

Note: Click to see a list of all of our published Trust and Safety reports

# Everest Group®
## With you on the journey

Everest Group is a leading research firm helping business leaders make confident decisions. We guide clients through today's market challenges and strengthen their strategies by applying contextualized problem-solving to their unique situations. This drives maximized operational and financial performance and transformative experiences. Our deep expertise and tenacious research focused on technology, business processes, and engineering through the lenses of talent, sustainability, and sourcing delivers precise and action-oriented guidance. Find further details and in-depth content at **www.everestgrp.com**.

## Stay connected

**Dallas (Headquarters)**
info@everestgrp.com
+1-214-451-3000

**Bangalore**
india@everestgrp.com
+91-80-61463500

**Delhi**
india@everestgrp.com
+91-124-496-1000

**London**
unitedkingdom@everestgrp.com
+44-207-129-1318

**Toronto**
canada@everestgrp.com
+1-647-557-3475

**Website**
everestgrp.com

**Social Media**
🐦 @EverestGroup
in @Everest Group
f @Everest Group
▶ @Everest Group

**Blog**
everestgrp.com/blog