Highlights from TrustCon 2023

Recent developments in trust and safety

Prepared by KPMG

Background: On July 11–13, 2023, the Trust & Safety Professional Association (TSPA) hosted the second annual TrustCon in San Francisco, California, with over 800 representatives from technology platforms, third-party services, regulatory bodies, and civil society. TrustCon is a global conference dedicated to creating and fostering a global community of practice among trust and safety professionals to explore successes, lessons learned, and the future of the field. Here is what attending members of the KPMG Risk & Compliance team identified as key themes relevant to Big Tech risk and regulatory compliance.

Key themes

1 Increasing **trust and safety regulations are largely perceived as a positive but can be enhanced by greater iteration with platforms and flexibility on developing issues** and gray areas where the community has yet to agree on best practices.

- Trust and safety regulations can **enhance safety** of the community, **increase executive focus** on trust and safety requirements, **clarify essential versus nonessential requirements**, and help to **standardize** and level the playing field across platforms for collaboration.

- Platforms are paying attention to the regulatory landscape with some concern about meeting some requirements for Year 1.

- A number of participants recommended trial periods, **iterations of reporting cycles** for feedback from regulators, and **more consistent dialogue** with regulators to enhance the effectiveness and reasonableness of regulations.

- Some platforms expressed concern with attempts to regulate emerging areas where best practices have yet to be established by the community:

  — **Example:** Implementation of **age verification requirements** can carry significant implications for privacy, data security, access to information, freedom of speech, and overall rights of the child. Data collection methods present data security and privacy concerns. Automated detection methods still have a robust margin of error. Regulatory requirements related to age verification must leave room for platforms to weigh the various risks as appropriate for their unique scenario.

2   **Trade-offs are a constant in the trust and safety world, and regulatory requirements must include flexibility to allow proportionality** of response across unique platforms.

- Trust and safety frequently require choosing between two or more bad options. Many presenters stressed that the decision between trade-offs will be largely unique between platforms depending on their purpose, user base, type of contents, etc.

- **Proportionality is key**, and each company should weigh trade-offs for society and their users on a scale, considering privacy, safety/security, freedom of expression, localization, consistency, and due process, among other elements, determining where they are in greatest conflict and where to put the proverbial thumb on the scale to guide enforcement in various scenarios.

- There are also trade-offs for regulators:

  — Increased regulations can increase safety in some respects but can also decrease proportion of time trust and safety professionals spend on policy development and enforcement in favor of time spent on compliance.

  — Speed of policy enforcement versus speed of transparency: Often, more effective policy enforcement can lead to delayed transparency reporting, due to prioritization of trust and safety efforts, as transparency data is time-consuming to select, clean, approve, and publish.

3   **Closer collaboration with regulators can help them to develop regulations with the greatest impact that will drive platform change** and ensure trust and safety has a seat at the platform table.

- Regulators from several regulatory bodies attended TrustCon 2023 to learn directly from trust and safety professionals as they prepare to release additional regulations, compacts, and guidelines.

- Overall, regulators expressed the desire to release regulations that will **drive platform change** and help trust and safety get the internal prioritization they need.

- Regulators stressed the importance of **coregulation and cooperation** between (1) regulators and regulators, (2) regulators and platforms, and (3) regulators and civil society.

4   While substantiating metrics are critical and have received significant attention, **qualitative transparency is essential for meaningful public and regulatory reporting.**

- Reporting on metrics alone has limitations considering the differences of implementation and definitions between platforms and the potential for misrepresentation or misunderstanding of issue across and between platforms.

- **Meaningful transparency reporting includes qualitative elements** and should:

  — Explain the **decision-making** process for product design and launch, policy development, and enforcement, providing detail on how trade-offs between risks and values are evaluated

  — Describe **how trust and safety fits in** to overall product design, launch, and operation

  — Use **case studies** to describe product design and approval, the content moderation lifecycle, and examples of principled deviation from standards and policies

  — Metrics should be **contextualized** with information on geopolitical, cultural, or other events that impact numbers at different times.

5 **Artificial intelligence is a double-edged sword for trust and safety,** as it enhances and accelerates trust and safety processes but represents the greatest concern for trust and safety professionals today with limited definitive solutions to date.

- Generative AI was one of the key topics of the convention, and platforms indicated they are actively working to both leverage and moderate artificial intelligence in the ways it most impacts their users and processes, including:

  — **Content moderation:** Machine learning tools accelerate the detection and removal of violative content, enhance the identification of classifiers, and enable consistent and comprehensive content analysis at scale.

  — **Transparency and adaptability:** Algorithms can introduce bias into automated processes, and limited visibility into models can lead to reduced consumer trust. Additionally, models can be difficult to adapt to different languages and contexts in the digital world.

  — **Misinformation:** Generative AI enables disinformation and deepfake content at scale, accelerating the potential spread of misinformation and raising the bar for identifying fake content.

Contact us

John Kemler
Managing Director, Advisory
Risk, Regulatory & Compliance
T: 347-754-2133
E: jkemler@kpmg.com

Casi Alexander
Director, Advisory
Risk, Regulatory & Compliance
T: 808-517-2117
E: casialexander@kpmg.com

Nicole Trawick
Director, Advisory
Risk, Regulatory & Compliance
T: 214-949 3335
E: ntrawick@kpmg.com

Some or all of the services described herein may not be permissible for KPMG audit clients and their affiliates or related entities.

Learn about us: **in** | **kpmg.com**