ESAFETY COMMISSIONER
Senate Legislation Committee
2021-22 Supplementary Budget Estimates Hearings
25 – 28 October 2021

BACK POCKET BRIEF

28. ANONYMITY, DIGITAL IDENTITY AND FACIAL RECOGNITION

KEY MESSAGES

- Identity verification online can raise tension between a number of rights – to privacy, freedom of expression and freedom from abuse. eSafety acknowledges this tension, and recognises that a balance is needed, where the misuse of anonymity or identity shielding is restricted without removing any of the legitimate benefits they can bring.
- Stronger and more transparent identity-related policies are needed on most digital platforms – and industry should embed safety and risk management into their products, including:
  - identifying, addressing and responding to fake, imposter or impersonator accounts
  - preventing known techniques used by perpetrators to target and abuse others
  - authenticating and validating user identities – particularly when transgressions occur.
- A platform or service does not need to know or control identifying information on its user base to address harmful behaviours.
  - Information such as IP addresses, device identifiers and analysis of user behaviour, metadata and account details are often enough to take action and limit the ability for users to re-register or create multiple or fake accounts.
- Solutions that seek to address the protection of users online must be evidence-based, accurate, proportionate, non-discriminatory and subject to independent and objective scrutiny. Technical measures need to address safety, privacy and security concerns before deployment.
- Online conflict, harassment and targeted abuse are inherently social and cultural problems – not technological.

KEY STATISTICS AND RESEARCH FINDINGS

- We know that anonymous, fake, imposter and impersonator accounts are contributing factors to negative online experiences. In 2020, 16% of our cyberbullying complaints included the categories 'fake' or 'impersonation' accounts.
  - The Online Safety Act provides us with stronger information gathering powers, particularly for obtaining information about end users, which will assist in the resolution of complaints and administration of end-user notices where required.
- Twitter stated it can identify 99 per cent of users whose accounts were permanently suspended due to abuse targeted towards English players during the Euro 2020 Final.
- New Danish research suggests that, "[Online] aggression is not an accident triggered by unfortunate circumstances, but a strategy [hostile people] employ to get what they want including a feeling of status and dominance in online networks."[i]

Contact Officer/s
Name: [redacted]
Phone: [redacted]
Email: [redacted]

INDUSTRY DEVELOPMENTS

- Industry is taking proactive steps and enhancing user safety through investment and innovation in new technologies that ensure their users are identifiable. Such measures ensure that age-appropriate safeguards can be implemented, and appropriate enforcement mechanisms employed.
- User verification methods include:
    - **Verified chat through validated email/phone number**– Twitch creators and moderators can now enable 'verified chat', requiring users to validate their phone and/or email before they can send messages.
    - **Government issued ID or passport** - Instagram (but not Facebook), TikTok and Tinder (Match Group) allow users to verify their identity by uploading a government identification document (e.g., drivers licence)
    - **Age estimation through facial analysis** – Yubo uses YOTI's facial analysis tool rather than facial recognition to verify the age of its users.
    - **Multi-factor verification using ID and real time verification** - Roblox requires players scan some form of legally recognised ID (e.g. passport or driver's licence) onto its app. Facial recognition then matches a live selfie against the ID. Users must undertake verification for early access to Roblox's new voice chat features.
- Many platforms and services are addressing privacy and safety concerns through age-appropriate safety features:
    - Tik Tok has set all 13-15yo accounts to private by default and tightened controls related to how u18s can interact with other users and content
    - Instagram now prompting u18s to turn accounts to private on sign-up and have tightened controls on how u18s can interact with other users.
    - Tinder has introduced photo verification which allows users to pose in live selfies to verify their profile photos.
- These recent developments support our Safety by Design principles by providing more robust online safety safeguards to users as the digital ecosystem evolves.

BACKGROUND

ANONYMITY

- Anonymity offers privacy, data security, protection to individuals to explore religion, sexuality or support services and avenues for whistleblowing. It supports democratic participation and freedoms and can prevent possible physical dangers from occurring.
- Unfortunately, it is also used to incite hate and violence against others, to commit illegal activities, spread disinformation and disrupt social cohesion, groom children, and/or stalk women.
- Online anonymity has been canvassed in the media on several occasions, including:
    - racial hate speech being directed at athletes
    - coordinated pile-on attacks against prominent figures or celebrities (e.g., Magda Szubanksi's involvement in government COVID-19 awareness campaigns)
    - far-right sympathisers using anonymous or encrypted platforms to recruit and communicate.
- There is evidence of individuals perpetrating abuse behind pseudonyms and fake profiles, although others will use identifiable accounts. Revealing personal information

Contact Officers
Name: [redacted]
Phone: [redacted]
Email: [redacted]

can increase levels of toxicity in certain forums and can fuel feelings of credibility and persuasiveness. There is no absolute deterrent to nefarious online actors.

- Industry has a responsibility to identify possible misuse and enforce their terms of use. Industry should proactively educate users on positive online behaviours and the consequences of harmful activity. eSafety's Safety by Design initiative can guide and enable industry to take these actions.

IDENTITY AUTHENTICATION ONLINE

- Companies already verify and authenticate users for content moderation purposes, advertising purposes, or to gain market traction using facial recognition, other biometric data, online behavioural signals, third-party verification and identity validation processes.
- Two-factor and multi-factor authentication are being used more frequently. Blockchain-based identity management systems, digital signatures, multi-layered trust platforms and differential traceability also offer digital identity solutions which eSafety – and the Government more broadly - are monitoring.
- The *Know Your Customer (KYC)* approach is an effective solution in the financial and gambling sectors as it reflects legislated regulatory requirements. A range of technical and practical issues must be considered to apply such a system transparently, legally and equitably across digital platforms.

FACIAL RECOGNITION TECHNOLOGY

- Research into the application of facial recognition has shown high rates of inaccuracy, error and bias – particularly for women and people from culturally and linguistically diverse backgrounds. Continued investment is refining of these technologies.
- Recent concerns about algorithmic bias, software inaccuracy and questionable use practices - have led the likes of IBM, Amazon and Microsoft to seek to withdraw their facial recognition services from use within law enforcement settings in the United States.
- Community concerns include unease over inherent biases in algorithmic systems and security and privacy concerns. Research by Monash University in May 2020 highlighted that:
    - Almost half of those surveyed (49%) felt that use in public spaces constituted an invasion of privacy
    - More than a third of the respondents felt that the risks outweighed its benefits
    - More than 60% felt that they should have the right to opt out of any facial recognition database.
- Using facial recognition technology as part of national digital identity schemes (nationally and internationally) has raised technical, policy and human rights concerns.
- eSafety acknowledges the recommendations and actions from broader government on the use of age assurance and verification technologies, including considerations by the Digital Transformation Agency on their Digital Identity and the Department of Home Affairs Enterprise Biometric Identification Services tools.
- eSafety is actively considering what role facial recognition might play, if any, in an age verification system for online pornography.

---

[i] BOR, A., & PETERSEN, M. (2021). The Psychology of Online Political Hostility: A Comprehensive, Cross-National Test of the Mismatch Hypothesis. American Political Science Review, 1-18.

Contact Officers
Name: [redacted]
Phone [redacted]
Email: [redacted]