Making the Internet Safe

The internet has evolved way beyond the metaphorical 'global village square'. New platforms and networks have connected, entertained, and protected communities in ways no one could have imagined in the days of dial-up.
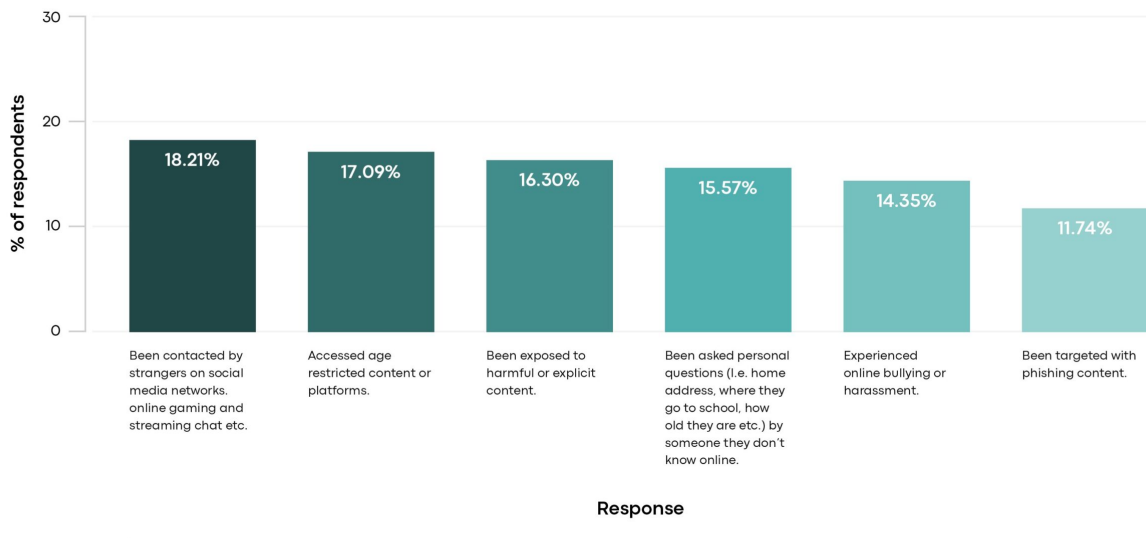
Temps de lecture : minute

1 October 2024

However, with the rapid advancement and adoption of new platforms, new, well-publicised risks have also emerged for children and society. It's important that online platforms prioritise their users' safety as a core feature in their development, but as of today, it's evident there is significant progress to be made.

<u>Our research</u> illustrates the scale of this challenge, with 74% of parents saying their kids have experienced online harms, from seeing illicit content to receiving unwanted contact with strangers. Since only one in ten (11%) children say they'd inform their parents about seeing harmful content online, the true scale of online risk is likely even bigger.

Parents are increasingly demanding more action and accountability from tech titans, with the vast majority (84%) supporting stricter regulations and laws to protect children online, even if it might limit their internet access. As is often the case, legislators globally are racing to catch up with the changing internet landscape, driven by an increasingly wary and mistrustful public.

To your knowledge, has your child encountered the following online?

% of respondents (y-axis, scale 0 to 30)

- Been contacted by strangers on social media networks, online gaming and streaming chat etc. — **18.21%**
- Accessed age restricted content or platforms. — **17.09%**
- Been exposed to harmful or explicit content. — **16.30%**
- Been asked personal questions (i.e. home address, where they go to school, how old they are etc.) by someone they don't know online. — **15.57%**
- Experienced online bullying or harassment. — **14.35%**
- Been targeted with phishing content. — **11.74%**

Response (x-axis)

For this research 3,000 parents were interviewed across the UK and US (1,500 per market) in June 2024. Research was conducted by Censuswide.

Times are changing

But change is coming. In the US, state-level legislative battles rage on the topic of age verification. Federal legislation, in the form of KOSA and COPPA 2.0, is set to lay out a new agenda for online safety nationwide. Meanwhile, the EU Digital Services Act is already being actively enforced, and the Online Safety Act in the UK is due to be from early next year. The coming years will see legislative kinks hammered out and regulators starting to flex their enforcement powers as they work to hold platforms accountable for perceived failings.

The numerous new global legislations are ambitious, attempting to encompass as much online safety 'ground' as possible. For online platforms, the most important provisions surround age assurance – preventing underage users from accessing inappropriate content, products and services – and content moderation – removing harmful and illegal content from platforms.

Platforms certainly haven't been idle on this front, but much more can and should be done before regulators start to 'force hands'. The key to this is moving the industry from a reactive – "fix it when it's broken" – to a pre-emptive footing. This means not waiting until content is flagged and reported by users, but going upstream and dealing with content and harm at its source.

The rapid adoption of generative AI has fuelled deep-fakes, explicit content, child sexual abuse material (CSAM) and other content issues. New tools being introduced which enable no-holds-barred image generation to have exponentially increased the already huge firehose of content that platforms must tackle every day.

While regulation takes time to bite, the regulatory environment is driving huge interest and investment in technologies that will support tomorrow's safer internet. Indeed, the global content moderation solutions market is forecasted to grow to _$17.45B by 2027_. At the heart of this rapidly evolving sector is AI and Machine Learning (ML), spearheading the development of online safety tools.

AI-powered online safety

Innovation, driven by necessity, creativity and the regulatory agenda, is bringing a new era of online safety technologies from a range of different players. At the heart of these technologies are increasingly advanced machine learning algorithms. With the right data, these algorithms are trained to discern harm from safety, manage access at scale and augment human capabilities. The combination of man and machine in online safety is critical, given the scale of the challenge which is now too big for humans to face alone.

With the worry over privacy and anonymity, age assurance innovation is largely focusing on behavioural analysis and user biometrics.

For example, highly-accurate, privacy-preserving methods currently being offered include age estimation using just an email address, facial age estimation, or even other promising biometric innovations including voice and even palm-based age estimation.

The recent '*Children's Code*' guidance from the UK's Information Commissioner's Office (ICO) is guiding innovation in this space, focused on methods with 'low friction' and 'high accessibility'. The intent is to make the internet safer whilst not encumbering valid users from accessing platforms.

Beyond age assurance, content moderation technology will also see huge and growing demand in the coming years. Thanks to AI, advanced algorithms can now sift through mountains of content to identify and flag that which poses threats or contains illegal materials. Beyond initial training, this human-machine dialogue is crucial to AI becoming ever more accurate at understanding what is and isn't harmful.

However, it's important to keep in mind that AI powered content moderation technology should go hand in hand with human moderation. And that humans should always make the final decision when it comes to tagged content, ensuring that the solution is as accurate and scalable as possible.

With AI critical to many innovations in online safety, high quality training data has become more important than ever. However, the trouble is that this data can be rare and often hard to access, particularly when it comes to Personal Identifiable Information (PII) of children. This challenge requires industry-level action and collaboration to make online safety solutions as accurate and effective as possible.

Looking to the future

As governments around the world focus on protecting our future generations, there is a collective optimism and a push for all stakeholders – tech giants, policymakers, safety tool providers and child safety advocates alike - to actively contribute towards moulding a secure online environment.

With a clear regulatory outlook, the stage is set for continuous innovation and investment into online safety technologies. Despite the challenges ahead, tech companies must continue to invest in online safety. It's not just about complying with laws; but embracing the innovation at hand to move from a reactive to a proactive online safety stance. The path forward is clear - embrace change and collectively strive for the safer internet that society deserves.

Ryan Shaw is the Founder and CEO of _Verifymy._

Meet Verifymyage, the frictionless solution for online age verification

MADDYNEWS UK

The newsletter you need for all the latest from the startup ecosystem

JE M'INSCRIS

Article écrit par Ryan Shaw