List of appendices

Appendix 1:
Call for evidence literature evaluation criteria

This evaluation criteria was developed by the eSafety Research and Evaluation team based on literature that detailed the key qualities of robust quantitative and qualitative research.[1] The criteria were tested using a random selection of sources cited in submissions and then adapted to ensure it was fit for purpose. The criteria development was led by one researcher and was subject to critical review and discussion from three other members of the research team during the development and testing process.

<table>
<tr><th colspan="3">Relevance of Studies</th></tr>
<tr><td colspan="3">Primarily relating to the location in which the research was conducted. Australia is classified as a WEIRD (Western, Educated, Industrialised, Rich and Democratic) country. Countries where relevant and comparable data comes from will also be industrialised, relatively economically stable and technologically advanced. They may not be Western but are likely to be educated, industrialised, rich and democratic.
Importantly, Australia's status as a settler colony is also relevant to the comparability of data. However, the experiences of Aboriginal and Torres Strait Islander people in Australia may have limited comparability.
Studies are more relevant the more recently they were published, and the more recently that data was collected.</td></tr>
<tr><td><b>High:</b> Studies based in countries with similar socio-economic and</td><td><b>Medium:</b> Studies based in countries with similar</td><td><b>Low:</b> Studies based in countries with low similarity to Australian</td></tr>
</table>

[1] Tong, A., Sainsbury, P. and Craig, J., 2007. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International journal for quality in health care*, *19*(6), pp.349-357; Stenfors, T., Kajamaa, A. and Bennett, D., 2020. How to... assess the quality of qualitative research. *The clinical teacher*, *17*(6), pp.596-599.

cultural contexts to Australia, or in Australia. Studies that take place in the last five years.	conditions to the Australian Context	conditions. Studies are more than 5 years old.
Relevance of Submissions		
How does most of the secondary evidence used in the submission rate in terms of relevance? How applicable is primary evidence (eg; based on the industry experiences of the submission author) to the contemporary Australian context?		
High: Most secondary evidence cited is high quality. Primary evidence is Australian based and recently attained.	**Medium:** Most secondary evidence cited is medium quality. Primary evidence is mostly Australian based and relatively recently attained.	**Low**: Most secondary evidence cited is low quality. Primary evidence is not Australian based and was not recently attained.

Quality of Quantitative Studies		
Quality of quantitative research will be evaluated based on the following factors: bias, reliability of methodology, use of validated measures, sample size & population, length of study, replication of results where relevant.		
High: No indication of bias (e.g., study substantially engages with high quality evidence that provides findings or arguments counter to those of the study, authors include positionality statement. Studies use robust methodologies, nationally representative sample populations and sizes (where relevant/appropriate). Studies are longitudinal and replicable where relevant. Use of validated scales to measure harm, effectiveness etc.)	**Medium**: Minimal indication of bias (e.g., study engages with some high quality evidence that provides findings or arguments counter to those of the study. Studies use robust methodologies. Sample sizes are small, and populations are not generalisable (where relevant). Validated measures and non-validated measures are used.)	**Low**: Indication of bias (e.g., study does not engage with high quality evidence that provides findings or arguments counter to those of the study. Study methodology is not robust, sample sizes are small, and populations are not generalisable, study does not use validated measures.)

<table>
<tr><th colspan="3">Quality of Qualitative Studies</th></tr>
<tr><td colspan="3">Quality of qualitative research will be evaluated based on the following factors:
• bias (ie; author positionality statement, reflexivity, engagement with high quality evidence that provides counter arguments/findings),
• reliability of methodology (ie; methodology has a sound theoretical framework, clear research question, justifiable sampling strategy and methods of data collection and analysis that are supported by literature/disciplinary convention),
• confirmability (ie; there is a clear relationship between data and findings), strength of analysis (ie; alignment with and contribution to existing literature).</td></tr>
<tr><td><b>High</b>: No indication of bias (e.g., study is highly reflexive, study engages with high quality evidence that provides counter arguments/findings. Methodology is robust and reliable. Findings are confirmable and analysis is strong.)</td><td><b>Medium</b>: Low indication of bias. Methodology is mostly robust and reliable (e.g., three or more elements of methodology are not highly reliable). Link between data and findings (confirmability) is moderate, analysis is valid but not strong.</td><td><b>Low</b>: Indication of bias (e.g., study does not engage with high quality evidence that provides findings or arguments counter to those of the study. Methodology is not very reliable or robust. Link between data and findings is weak and analysis is not strong.)</td></tr>
</table>

<table>
<tr><th colspan="3">Quality of Submissions</th></tr>
<tr><td colspan="3">What is the quality of the secondary evidence cited? Is evidence used accurately and appropriately? What is the quality of the primary evidence used? (eg; is it put into context of the author's lived/industry experience? Is it given appropriate weight in the submission?) Does the submission engage with evidence that provides contrary arguments to their own?</td></tr>
<tr><td><b>High</b>: Most secondary evidence cited is high quality & is used accurately. Primary evidence is well contextualised and is given appropriate weight relevant to its context and the secondary literature. Submission engages with evidence that provides counter arguments to their own.</td><td><b>Medium</b>: Most secondary evidence cited is medium quality and is mostly used accurately. Primary evidence is somewhat contextualised but most is not weighted appropriately. Submission engages with only some evidence that provides counter arguments to their own.</td><td><b>Low</b>: Most secondary evidence cited is low quality & is not used accurately. Primary evidence is not contextualised and is not weighted appropriately. Submission does not engage with evidence that provides counter argument to their own.</td></tr>
</table>

Appendix 2: Call for evidence request and thematic summary of submissions

Call for evidence requests by sector

Age verification, age assurance or third-party online identification providers

- Information on the state of the age verification (AV) and age assurance (AA) market in Australia
- Technical overviews of AV/AA systems and products – and information and details on: how they can be integrated with digital environments, platforms and services; user costs (service provider or consumer); and costs to businesses.
- How AA/AV system protect the privacy, safety and security of users – and how they address data protection, usability and accessibility considerations,
- Evidence of AV/AA system effectiveness and impact in preventing the exposure of children and young people under 18 to harmful online content (including pornography).
- Evidence of accuracy and error rates of AV and AA solutions across ages, gender and ethnicity
- Evidence and insights regarding business and consumer preferences and tools which use identification, age verification or age assurance technologies.
- Evidence of how AV/AA systems and products have been communicated and explained to the public – and the impact on subsequent uptake and use.

Digital environments, services and platforms, and platforms

- Measures already being used to mitigate the exposure of children and young people under 18 to oline pornography and other age-inappropríte material – and their impact and effectiveness.,
- Evidence and insights regarding the use of proprietary or third-party identity, age verification or age assurance tools – and their impact and effectiveness.
- Views on the proportionality of AV/AA measures on mitigating the exposure of children and young people under 18 to online pornography.
- Information and evidence of other effective technological measures or interventions tht can prevent access to online pornography by children and young people under 18.
- Evidence of how AV/AA systems and products have been communicated and explained to the public – and the impact on subsequent uptake and use.

Adult industry

- Measures that industry members already take to reduce or prevent the exposure of children and young people under 18 to online pornography – and their impact and effectiveness.
- Evidence and insights regarding the use of AV/AA or user identification systems and tools (including successes and areas for improvement).

Academia

- Academic publications and evidence (including grey literature) on the impact of exposure to pornography on children and young people under 18.
- Academic publications and evidence (including grey literature) on the effectiveness of age verification measures to reduce the exposure of children and young people under 18 to age-inappropriate online content

and products (for example, pornography, online gambling and the sale of alcohol and e-cigarettes).

- Insights into protective measures being used outside Australian and any examples of international good practice for mitigating the harms associated with the exposure of children and young people under 18 to online pornography.
- Evidence of the role, if any, that anonymity plays in shaping access to, engagement with and distribution of online pornography.

Not-for-profit

- Insights into how organisations are engaging with children and young people, parents, carers and teachers on the topics of: pornography, harmful online sexual behaviours, respectful relationships and consent. These insights can refer to resources, campaigns, training and educational programs, or independent evaluations of such initiatives.
- Research or evidence on the impact of exposure to pornography on children and young people.
- Lessons from other measures or interventions used for age verification across different public health issues.
- Evidence of other measures or interventions that could support or enhance any age verification system
- Evidence of the role, if any, that anonymity plays in shaping access to, engagement with and distribution of online pornography.

Civil society

- Research or evidence on the impact of exposure to pornography on children and young people.
- Lessons from other measures or interventions used for age verification across different public health issues.
- Evidence of other measures or interventions that could support or enhance any age verification system.

- Reactions to age verification technology being used to verify a minimum age before online pornography can be accessed.
- Perceptions of the benefits, limitations and concerns about age verification being used to limit access to online pornography

Thematic summary

This summary is a compilation of the emerging themes from the responses to the call for evidence. The views and opinions are those of the authors and do not reflect eSafety's position. They are an important contribution to informing the next phases of consultation and the development of the implementation roadmap.

Policy considerations

Governance

- There should be a single oversight body that investigates, audits, monitors and assesses compliance with the AV regime. This body should have adequate enforcement powers.
- Specific technologies should not be prescribed. Instead, principles of proportionality, community standards and consumer choice should guide which technologies are used. The regime should accommodate innovation and tool advancements.
- A recognised body should certify AV providers. Certification should test for overall effectiveness, privacy compliance and security (including for data storage facilities).
- There should be ways to identify and check compliance of new sites. Sites that do not comply should get no commercial advantage.
- Consideration should be given to the regulatory, administrative and financial burden to both industry and the consumer when determining what makes a proportionate, feasible and effective AV regime.
- Businesses, not consumers, should bear the cost of using AV technologies. Tools should be cost effective for all pornography sites – from individual businesses to large-scale platforms.

Alignment

- The regime should consider the outcomes of the review of the Australian classification system.
- A whole-of-government approach should be deployed to digital regulation to ensure online safety measures align with online privacy and security policy.
- Consideration should be given to aligning the age of access with the age of consent.

Scope

- Age verification should extend to all commercial pornography sites, not just sites which allow users to generate content. Any online service provider that poses a risk of exposing children to pornography should have measures to prevent children gaining access.
- Requirements should be proportionate (based on risk). Blanket requirements across all relevant platforms and services should be avoided. Industry should not have to scan for online pornographic material.
- Limit and focus AV tools on sites which provide direct access to pornographic content or bear the closest relationship to pornographic material. Requirements should not apply to private messaging or end-to-end encrypted (E2EE) systems.
- The AV regime should be holistic. It should consider mobile device filtering, ISP filtering and parental controls as well as AV tools. Consideration should be given to the regulator overseeing a list of relevant URLs, which are captured by all filtering services.

Privacy and security

- There are privacy concerns and cyber security risks relating to commercial pornography websites directly processing user data. These sites should use third-party verification tools.
- Age verification and assessment technologies should meet clear and transparent standards and technical requirements. They should also be certified, independently audited and demonstrate robust privacy and security settings.

- A data minimisation approach should be followed for AV/AA tool standards. Only age attributes should be shared between the AV/AA technologies and content hosts. Other data should not be shared.

Research and evidence on the impact and effects of pornography

- Studies point to accidental exposure from ages 11 to 13 – with a significant proportion of young people having viewed sexually explicit content by the age of 16.
- Access and exposure are not limited to pornography websites. This also occurs on gaming platforms, social media and search engines.
- Access and exposure to pornography for under 16s is generally seen as inappropriate. A more nuanced approach to pornography should be considered for young people who have reached the age of sexual consent (16- to 17-year-olds).
- Intentional access to pornography occurs for educational purposes, self-exploration, to understand 'expectations' when having sex, excitement and entertainment. Lack of access to comprehensive sex education is associated with greater intentional access to pornography.
- The influence of pornography on the sexual practices and beliefs of young people varies. It has been associated with:
    - A greater likelihood to pressure and coerce others to perform unwanted, derogatory and violent sexual acts.
    - Influencing perceptions of sexual expectations and negatively impacting awareness, attitudes and understanding of consent.
    - Increased frequency of watching pornography is associated with a greater likelihood of accessing extreme and violent pornography, and presenting greater levels of sexual aggression, sexual objectification and sexual coercion. There are factors other than pornography which may also contribute to these attitudes and behaviours.
    - Helping young people to learn the practicalities and mechanisms of sex and explore their sexual identities. This is particularly important

when this subject matter is not adequately discussed in the school curriculum or with parents or carers.

- The negative impacts of pornography are more pronounced in children under 14, marginalised and at-risk young people, and high frequency users of pornography.
- Negative impacts are mostly associated with accidental exposure to pornography and exposure to violent or extreme pornography. However, not all young people are negatively impacted by exposure to such content.
- Age verification should not block access to vital sexuality and sexual health information for young people, restrict adults' access to online pornography, or reduce safe online spaces for sex workers and the sale of adult products.
- Technologies should be designed so they are easy for children and parents to understand. This includes how tools works and how we use, store and protect data.

Educating young people on pornography and online harmful sexual content

- Educating young people on healthy sexual relationships, behaviours and sexuality can help counter the negative impacts of pornography. This includes risky or violent sexual behaviour, and reinforcing unhealthy or unrealistic expectations regarding gender, power, sex and relationships.
- Sex education in schools could be enhanced through:
    - Providing authorised leadership and fostering a culture that encourages age-appropriate information-sharing with young people.
    - Introducing a national, comprehensive sexuality and relationships education curriculum that is developmentally appropriate and runs across all year levels with a focus on:
        - improving young people's sexual literacy, which aligns respectful relationship education with digital literacy education

    - encouraging active participation, in recognition that young people see relationship and sexual health education as an important part of their psychosexual development
    - peer-led discussions among 16- to 18-year-olds and teacher or expert specialist-led discussions and workshops
    - the risks associated with viewing online sexually explicit media.
- Whole-school approaches to address student wellbeing and pornography exposure are important. These approaches can include policies, staff professional development, parent and community partnerships, guidelines for student education practices and evaluation, and parental support/advice on managing technology in the home.
- There should be community-based programs to support young people who are vulnerable to missing out on school-based relationship and sexual health education. These programs can be in residential care, flexible learning centres, community health centres and youth justice centres.
- Evidence-based public health content and programs can support young people, parents, educators and frontline professionals to talk openly about pornography and enhance the wellbeing and moral development of young people.
- Relevant industry stakeholders and experts should review any advice for navigating adult content online.
- Experts or relevant organisations could be funded to develop evidence-based public health resources and training.
- Parents and carers should have access to guidance on the tools that can better protect children online (including verification tools, filtering and parental controls). Parents and carers are critical to providing children with key literacy skills on pornography.
- Families are a significant source of information and support. However not all young people have equal access to adults they can turn to for advice.
- Young people often turn to peers, so it’s important to advance young people’s sexual wellbeing and pornography literacy.

- Targeted programs could be established for young people identified as problematic users of pornography.

Impact on the adult industry

- The economic impact and regulatory/compliance burden on both the adult industry and consumers should be considered within the roadmap.
- The cost of AV/AA tools may be by disproportionate and prohibitive for smaller producers and sites.
- The cost of implementing more stringent age verification processes may be anti-competitive for smaller producers and individual sex workers.
- AV/AA may disproportionately impact on LGBTQI+ and female producers.
- AV/AA may push sex workers onto unsafe platforms and systems.
    - AV/AA requirements should apply to every online pornography service available in Australia.
    - This provides a level playing field for all platforms and services.
    - Transparent policies, processes and guidance should exist to ensure compliance.
    - The regulator should have the capacity to take swift action against non-compliant platforms and services, so that compliant platforms and services are not disadvantaged.
- The role and use of social media by sex workers and industry-wide approaches to professional sex workers should be considered, as should the effect on sex worker advertising.

Current initiatives by platforms and services

- Dating apps and age-restricted content-sharing platforms and apps use stricter age-checking measures than most social media platforms. These include user registration with credit card details and in some instances user verification. App stores prohibit apps that contain or promote pornography.
- Filtering software and parental controls settings on mobile devices and computers can also help to prevent access to more mature content.
- Platforms have mechanisms for:

    - restricting the content younger users are exposed to (including advertising)
    - blurring sensitive content, providing warnings, or making content unavailable
    - blocking 18+ users from contacting under 18 users
    - preventing age-restricted content from being viewed by unregistered users
    - reporting and blocking inappropriate content
    - parental controls and filters which also allow parents to monitor children's activity
    - asking for age verification through credit card or identity verification to check if a user is underage
    - blocking users who breach terms of service by using IP addresses, device signatures or other data
    - deleting accounts if users are unable to prove they meet the minimum age.
- Some platforms use proactive technologies and machine learning to moderate content or identify underage users, including:
    - age estimation technology to determine whether a user is under 18 or 18+
    - artificial intelligence tools that help to understand someone's real age
    - persistent cookies that platforms place on devices to prevent children from attempting to circumvent age restrictions (e.g., multiple attempts entering a valid birth year)
    - proactive detection tools for identifying and removing sexually explicit images or videos.
- Some platforms have policies which:
    - restrict pornography, nudity and sexually explicit content and advertising

  - prohibit content which endangers the emotional and physical wellbeing of minors.
  - Search engines can:
  - apply filtered search tools which prevent search results for sexually explicit content or websites
  - block hyperlinks that drive traffic to commercial pornography sites, prohibit pornography ads or ads made against pornography websites
  - remove sexual and violent terms from autocomplete search functions.

Age assurance measures and alternative safety tech

- AV and AA providers offer digital apps, application programming interfaces (APIs) or send links which individuals can use to provide their age attribute.
  - Digital identity apps were proposed as an effective AV method as an individual's personal information remains stored on their phone. QR codes or links can be used to connect with the app and allow for an age attribute to be shared with the requesting site or platform.
  - Physical age tokens can be used to generate an online password for any age-restricted content. Mobile operator age checks were also suggested.
  - Facial analysis technology was presented as a suitable biometric option, particularly for individuals who do not have government identity documents. It allows for one-time facial scans which estimate a user's age – no data is stored.
  - Submissions noted that the technology is still nascent and few providers have achieved high accuracy.
  - Some research has raised concerns of ethnic and gender bias by some facial analysis technologies.
- Submissions raised some concerns regarding the use of:
  - Database checks, which may be less proportionate as they are typically used for 'Know-Your-Customer' (KYC) and anti-laundering or

  fraud regulation. Document verification can also be a more costly process.

  - Credit card checks, which may be easy for young people to circumvent by using their parent’s or a third-party’s details.
  - Official documentation (e.g., driver’s licence or passport), which is often used for regulated, age-restricted retail and services (such as alcohol sales and gambling) and may raise privacy concerns
  - Biometric data (such as facial recognition), which may raise privacy and security risks and concerns of surveillance.

- Submissions demonstrated mixed attitudes to one-off age checks, which present less friction for user experience, and to single-use checks, which require repeat user verification.

Appendix 3: Methodology information – Adults perception of age verification technology

Methodology

Overview

All aspects of this research were undertaken in compliance with the Privacy Act 1988 and Australian Privacy Principles, the Research Society Code of Professional Behaviour, and Privacy (Market and Social Research) Code 2014.

A short peer review consultation process was undertaken to explore wording, comprehension and understanding of the proposed survey items (including any gaps in the questioning against the research objectives). Five in-depth interviews of approximately 30 minutes in duration each were held on 15 and 16 March 2021 with staff at eSafety to discuss the wording of the draft questions. Interviews were not recorded. Participants in the peer review worked across the agency and were not part of the immediate project team, however, they will have come with a more informed understanding of the topic in comparison to the general community target audience. In the interests of time and budget, this tradeoff for feedback was considered acceptable.

Data collection

The i-Link non-probability online panel was used to access a sample of the Australian general community of adults. Quotas were put in place to reflect age, gender and location (State/Territory) data from the Australian Bureau of Statistics (ABS) 2016 census data. The final approved questionnaire was provided to i-Link for programming and distribution to their panel members. Data collection was conducted between 29 March and 13 April 2021. The survey was approximately 12 minutes in duration on average. A total of 1,200 surveys were completed.

A detailed profile of respondent characteristics is available below (Tables 1-3).

Weighting

A weight was calculated for each survey respondent. The data was weighted to ABS data for age, gender and location (State/Territory). Weighting made no significant differences to the results at the total level but has been applied in analysis for consistency across sub-group and total results reported.

Limitations

A non-probability online panel provider was used to access a sample of the general community for the online survey. Note, that this is a non-generalisable community sample and although online panel providers make efforts at recruiting a broad population, there is research that indicates online panel samples may under-represent some subgroups compared with others (AAPOR, 2010[2], Centre for Social Research and Methods, 2018[3]). Statistical testing should be taken with some caution and confidence intervals have not been established.

Survey Sample

Table 1 General individual characteristics (unweighted %) (n)

Characteristic	%	n
Age		
18-24 years	12%	146
25-34 years	18%	215
35-44 years	19%	222
45-54 years	18%	215
55-64 years	15%	182
65-74 years	10%	119
75 years and older	8%	101

[2] American Association for Public Opinion Research (AAPOR) (2010). Research synthesis AAPOR report on online panels. Public Opinion Quarterly, 74(4), 711–781

[3] D W Pennay, D Neiger, P Lavrakas, K Borg. (2018) The Online Panels Benchmarking Study: A Total Survey Error comparison of findings from Probability-based surveys and Non-probability online panel surveys in Australia, CSRM Methods Series, 2/2018, Centre for Social Research and Methods. https://csrm.cass.anu.edu.au/research/publications/online-panels-benchmarking-study-total-survey-error-comparison-findings

Gender at birth		
Female	51%	610
Male	49%	588
I use a different term (please specify)	0%	2
Gender currently recognised		
Man or male	49%	584
Woman or female	50%	601
Non-binary	1%	10
Prefer not to answer	0%	5
Do you identify as Aboriginal or Torres Strait Islander		
Yes	4%	44
No	95%	1142
Refused	1%	14

Table 2 General individual characteristics (unweighted %) (n)

Characteristic	%	n
Are you a parent		
Yes	56%	675
No	43%	517
Refused	1%	8
Highest level of education		
Primary school	0%	2
Years 7-9	2%	27
Year 10	8%	96
Year 11	3%	34
Year 12	16%	193

Trade/apprenticeship	6%	66
Other TAFE/technical certificate	17%	203
Diploma	12%	138
Bachelor degree	25%	300
Post-graduate degree	11%	133
Other	0%	3
Refused	0%	5

Table 3 Geography and location (unweighted %) (n)

Characteristic	%	n
State/Territory		
NSW	32%	389
VIC	25%	301
QLD	20%	244
SA	8%	90
WA	11%	126
TAS	2%	29
NT	0%	2
ACT	2%	19
Region		
Capital city	65%	779
Non-capital city (regional) area	35%	416
Refused	0%	5

Appendix 4:
Methodology information – Young people’s views on online pornography and age verification

Overview

The research was conducted to explore young people's current perspectives on online pornography, and their attitudes towards age assurance, by examining their lived experiences with pornographic content encountered online. This research contributes to the evidence base for the development of eSafety’s age verification roadmap and extends prior work conducted in the Australian context.

The objectives of the research were:

- to examine and provide insights into the lived experiences of young people in relation to their encounters with, and ideas about, online pornography
- to examine what support young people want in order to feel safe and prepared to navigate encounters with online pornography and to mitigate any potential harms that arise from these encounters
- to examine what young people think about age restriction and age assurance.

The research was conducted from 19 to 21 September 2022, with a survey of 1,004 young people aged 16–18 and six focus groups (n=32) with young people aged 16–18.

eSafety understands the impact of researchers’ intersecting experiences of power and marginalisation on our research and analysis. The team that authored this report is made up of cis-gender women of European and Asian heritage. Identities represented in the team include queer women and those with disability. Our team has expertise in quantitative and qualitative methodologies, online harms and safety, and the lived experiences of young people.

Ethical considerations

Various steps were taken to address ethical considerations during project development and recruitment. The project was submitted as part of the Human Research Ethics Committee (HREC) approval process. Ethics approval for the project was received from Bellberry Ethics Committee on 26 August 2022, ID 22CeSC117. eSafety collaborated with Professor Bronwyn Carlson and Maddi Day of Macquarie University's Department of Indigenous Studies to review the methodology and instruments for cultural safety and to ensure that questions were worded in a culturally sensitive manner.

Recruiting most participants directly, rather than via their parents or carers, was a key requirement for this project. Recruiting via parental or carer referral could prevent the participation of many young people, given the sensitive and personal nature of the research topic. Young people in out-of-home care, who are not close with their parents, carers or guardians, and those who don't feel comfortable discussing subjects such as online pornography with their parents, stood to be excluded from this study if contact via parents was the main form of participant recruitment.

In accordance with the National Health and Medical Research Council (NHMRC) Guidelines, only mature minors were accepted into the study as participants. To ensure they could make an informed decision about study participation, and to ascertain their suitability to take part as mature minors, participants were asked their age, followed by two pre-screening questions.

Steps were taken to reduce the risk of harm to participants, to ensure that their best interests were served, and that the research conducted provided for participants' safety, emotional and psychological security, and wellbeing.[4] Informed consent to participate in the research was obtained for study participation by providing information about the kinds of questions participants would be asked and by explaining the potential risks of participating~~.~~ In addition, a protocol was developed to ensure the safety and wellbeing of all participants in case of distress during study participation. More specifically, help-seeking/self-support information was made available to participants throughout the study.

[4] NHMRC, 2018. *National Statement on Ethical Conduct in Human Research (2007)*. Canberra: NHMRC.

The study

The study was comprised of two phases: an online survey, followed by online focus groups.

Online survey

The 15-minute survey consisted of 40 questions covering topics including:

- demographics
- whether they have encountered online pornography
- age when they first encountered online pornography
- locations where they encountered online pornography
- who they were with when they encountered online pornography
- how they encountered online pornography
- their responses to encountering online pornography
- their attitudes towards online pornography
- preferences on sources of support and information
- views on age-restricted access to online pornography
- views on who should be responsible for age-based restrictions on access to online pornography
- views on age assurance and verification technologies or processes.

Survey sample

A total of 1,004 young people participated in the survey. The sample included young people with disability (n=228), those who speak a language other than English at home (n=247), those who are LGB+ (n=219), trans and gender-diverse young people (n=31), and First Nations young people (n=31) (Table 4).

A non-probability-based online panel provider (Octopus Group) was used to recruit survey participants for this project. The survey was conducted Australia-wide, covering all states and territories, including regional and metropolitan locations. Survey participants included young people who are attending formal education as well as those who are not. It included those who are attending high school, university, or vocational training (TAFE), and those who are working or seeking employment.

Table 4 Survey respondents: Key demographics

	Number of young people (n)	% of sample
Aged 16 years	338	34
Aged 17 years	333	33
Aged 18 years	333	33
Men	356	35
Women	614	61
Trans and gender-diverse	31	3
LGB+ (lesbian, gay, bisexual and more)	219	22
First Nations	31	3
With disability	228	23
Speak a language other than English at home	247	25
Total sample size	**1,004**	

Online focus groups

The qualitative phase of the research comprised six one-hour online text-based focus groups of Australian young people aged 16–18. Questions asked in the focus groups aimed to complement the survey findings, adding depth and nuance and drawing out young people's opinions on and experiences of online pornography in their own words.

Topics covered included:

- views about what pornography represents
- views on what content should be restricted to various age groups
- what young people believe they need in order to feel empowered to navigate sexual content online
- perspectives on age assurance and verification technology.

Focus group sample

A total of 32 young people aged 16–18 participated in the six focus groups. The focus groups were made up of 12 sixteen-year-olds, 11 seventeen-year-olds and 10 eighteen-year-olds (Table 5). There were 15 women, 11 men, 4 non-binary young people, 1 trans man and 1 demiboy in the focus groups. Nineteen focus group participants identified as straight/heterosexual, one as lesbian/gay/homosexual, three as queer, five as bisexual, and one each as pansexual, asexual, demisexual panromantic and questioning (Table 6).

Based on participants' expressed preferences, one focus group consisted of LGBTIQ+ young people only, and another consisted only of heterosexual young men (Table 5). Four focus groups were not categorised according to gender identity or sexuality (Table 5).

Participants in three of the focus groups were recruited through a market research recruiter company (Q&A), while participants in the other three groups were recruited directly from their participation in the online survey. The focus groups were run by members of the eSafety Research Team through an online platform (VisionsLive). Participants' participation in the focus groups was pseudo-anonymous.

Table 5 Online focus groups: Characteristics

	Number of young people (n)	Composition	Recruitment
Focus group 1	5	Mixed	Not survey
Focus group 2	6	Mixed	Not survey
Focus group 3	5	Mixed	Not survey
Focus group 4	5	LGBTIQ+	Survey
Focus group 5	5	Heterosexual men	Survey
Focus group 6	6	Mixed	Survey
Total sample size	**32**		

Table 6 Online focus groups: Participant demographics

Participant sexuality	Number of young people (n)
Heterosexual/Straight	19
Homosexual/Gay/Lesbian	1
Queer	3
Bisexual	5
Asexual	1
Pansexual	1
Demisexual Panromantic	1
Questioning	1
Participant gender	
Woman/Female	15
Man/Male	11
Non-binary	4
Trans Man	1
Demiboy	1
Participant age	
16 years	12
17 years	11
18 years	10
Total sample size	**32**

Analysis

Octopus Group hosted the survey, collected and cleaned the survey data, and provided eSafety with raw data as well as descriptive analysis. eSafety checked and analysed the data further using Q Research software. Findings were checked and confirmed by a second eSafety researcher using SPSS Statistics software.

Focus group transcripts were thematically coded using the research software Condens. A coding framework was iteratively developed following Braun and Clarke’s (2019) method for thematic analysis.[5]

Limitations

Online panels are technically convenience samples. ‘Non-probability-based sampling’ means that not everyone has an equal chance of being selected to participate in the research. Results may be subject to a range of biases when compared with results from research using probability-based sampling. Although it is possible to control for demographic skews using quotas, controlling for psychographic skews arising from differential approaches to participation attraction is more problematic. However, it should be broadly noted that there is no perfect sampling approach for humans. A random digit dialling telephone approach is generally regarded as best practice in the development of true probability samples. However, this approach is complicated by a range of factors, including the prevalence and usage of mobile phones and landlines, and further by issues such as social desirability in responses – particularly relevant for this study, in which participants were asked a range of sensitive questions.

Qualitative findings reported from this research should be read with the understanding that they emerged from young people who had chosen to participate in the research. As a result, the findings may be slightly skewed towards those who are more comfortable talking about online pornography, as well as towards those who may have more experience of this topic.

Due to the nature of an online survey, we cannot be sure whether there is a reticence to report experiences of some encounters with online pornography by young people. For example, even though the survey was anonymous, some young people may feel more comfortable reporting accidental/unintentional encounters with pornography than deliberate encounters.

Specific survey findings for First Nations and trans and gender-diverse young people were not separated out in the main data collected. This was due to the small sample size for each group, which is an inherent challenge in statistical analysis. For example, even though we were able to recruit close to a nationally

[5] V. Braun & V. Clarke, 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health,* 11(4), pp. 589–597.

representative proportion of First Nations young people as part of our sample (3%), this was too small a number of participants in a 1,000-person sample to provide separate findings for this cohort. As a result, our findings are unable to reflect the experiences of these groups. We know that both cohorts are highly likely to experience online harms, and we acknowledge that reaching larger sample sizes, as well as doing research specifically with marginalised and at-risk groups, is a priority for eSafety Research. Specialist recruiters, as well as alternative recruitment strategies, will be considered for future studies.

This research does not contribute to the literature on the harms that may be associated with young people encountering online pornography. The focus of the study is not the impacts and effects of seeing online pornography. As such, our research cannot claim to represent a full, long-term picture of young people's experiences with online pornography.

Appendix 5: Consultation summaries

Age Verification Roadmap Consultations: Round 1 (May 2022)

Safety conducted targeted consultations with stakeholders for the development of the Age Verification (AV) roadmap between November 2021 and April 2022.

These targeted consultation sessions allowed for close examination of the evidence submitted in response to the call for evidence and to gain insights from industry and experts.

eSafety's consultation process was focused on exploring the evidence that will support the development of recommendations that are holistic, viable and proportionate.

The following are high-level, anonymised summaries of the consultation sessions held between November 2021 and February 2022, and appear in the order of each consultation date. Summaries of the remaining consultations will be published separately.

The summaries in this document represent a variety of views that were discussed by stakeholders during the consultation sessions and summarises them into key themes.

The views and opinions in these summaries are those of the stakeholders and do not reflect eSafety's position.

They are an important contribution to informing the next phases of consultation and the development of the AV roadmap.

More information about the roadmap is available on our website.

Adult industry: International stakeholder group

Consultation group overview

Stakeholders: International adult industry website providers.

Consultation date: 17 November 2021.

Overview: These stakeholders (First stakeholder and Second stakeholder) discussed age verification and assurance in practice and implementation, as well as social media platforms.

Age verification and assurance in practice

First stakeholder

- This stakeholder has found that there are many proposed age verification or assurance options, and many factors need to be considered when adding additional layers of protection to avoid creating greater harms.
- They noted that measures which create too much friction will deter customers or users from accessing compliant sites. They emphasised that this could create incentives for users to follow the path of least resistance and access alternative adult sites, which are non-compliant with any age assurance regulations. It may also be less secure for users, as well as less ethical in the production and distribution of adult content, and is more likely to contain harmful content.
- They explained users must be able to trust that their data is secure and will not be used improperly, and likewise, platforms must be able to trust the age verification or assurance providers are approved by the regulator. The stakeholder recommended age verification or assurance providers should be vetted by the regulator to ensure their data is secure and is not shared without consent. The stakeholder encouraged the development of international standards, regulation and certification.
- This stakeholder noted difficulty in engaging with the broader adult industry and tech industry on a variety of trust and safety issues. They believe that other players in the tech industry have a key role to play to ensure that adult sites' efforts to prevent children and young people from seeing pornography are successful, and do not create unintended consequences. For example, if there is a change in user traffic patterns away from sites, which require some form of age assurance and towards those without access restrictions, they highlighted that search engine algorithms could end up listing noncompliant, harmful sites in their top search results. Accordingly, the stakeholder

expressed that there must also be requirements for search engine providers in order to achieve desired outcomes.

- The stakeholder's current age assurance measures include:
    - Terms of Service which state users must be 18+.
    - The application of Restricted to Adults (RTA) meta tag, supplied by the
    - Association of Sites Advocating Child Protection (ASACP). This ensures registered sites are blocked or filtered by internet service providers (ISPs), search engines or when parental controls or filters are activated on devices or internet connections.
- The stakeholder felt that parents are responsible for the primary prevention of underage access to adult content through filters, parental controls and communication with their children, as they are able to directly mitigate access at home. However, the stakeholder acknowledged that many parents have a limited awareness of the available technology and how to implement it. They believe this should be an integral element for inclusion in the AV roadmap.

Second stakeholder

- This stakeholder expressed commitment to preventing children's access to pornography but noted that complying with regulations across jurisdictions globally is challenging as a business.
- The stakeholder reported the following measures currently in use:
    - The site contains a warning that the content is for adults only. The site also contains instructions to parents on how to block the site.
    - 'Restricted to Adults' (RTA) labelling and meta tagging, which stops the site showing up on search engines (if a 'safe search' or similar filter is applied).
    - The provider does not offer an app, as major app stores do not permit pornography, so the site is only accessible through browsers. Given that safety tools are available to block sites with RTA labelling and meta tagging, the provider suggested these measures should be sufficient to prevent children from viewing the content.
- The stakeholder observed that some social media apps are hosted by app stores and available for download by under-18 users, despite allowing explicit adult content under their Terms of Service. They also noted that RTA labelling does not apply to apps.
- The stakeholder highlighted a lack of awareness and skills for use of parental controls.
- They felt that tech companies often avoid collaborating with pornography companies and this stigma prevents the sharing of useful tools (for example, tools to detect and remove child sexual exploitation and abuse material).

Implementation

First stakeholder

- This stakeholder noted that age assurance should be cost-effective for all platform sizes and able to scale to large traffic volumes. They noted that the adult market is very diverse, and that a small independent producer would be less likely to have the budget or resources to implement technological solutions to reduce children and young people's access to pornography.
- The stakeholder suggested the introduction of a blacklist/whitelist model operated by the regulator, which ISPs reference to block or allow access to sites based on compliance with regulations. Other companies within the internet ecosystem could also act as a point of access enabling restrictions, such as operating system (OS) developers, browser operators, or device manufacturers. They also noted that other enforcement methods such as payment blocking can play a significant role in enforcing compliance with regulations.

Second stakeholder

- The stakeholder believes the cost of compliance is very high. They are concerned that while overarching European regulation merely requires some form of age verification or assurance, countries can make local laws which prescribe specific technology – resulting in many different compliance measures.
- They noted the frequency of age verification checks for users impacts both the cost for the company as well as the effectiveness of the measure. Asking for checks too frequently may drive users away, but less frequent checks (for example, one check per month) leaves opportunities for children to use an adult's device to access content.
- They stated that the current range of technological solutions are too expensive for the stakeholder's business model. They would prefer to see a variety of solutions available and felt this would be better both for business and for consumer choice. In this stakeholder's experience, document verification was the most expensive form of age verification.
- This stakeholder does not want to collect and process user information in house, as they feel it is too costly and presents a high risk of creating 'data honey pots'.
- The stakeholder is concerned that consumers would move toward using non-compliant sites to avoid sharing personal information.
- The stakeholder raised concerns about legislation applying only to pornographic websites and not to social media or other sites. They would prefer to see rules applied to all services where this type of content is available. In their experience, there is adult

content on many social media sites, but limited warnings, filtering, or due diligence in checking uploaded content.

Social media platforms

First stakeholder

- This stakeholder believes measures should be applied to a wide scope of platforms and services, as pornographic content can also be found on social media and other non-adult sites. Again, the stakeholder raised concerns that if measures only apply to adult sites, explicit content may migrate to unregulated sites.
- The stakeholder observed that social media services currently take different approaches to adult content. While some do not permit adult content and make efforts to detect and remove it, others allow this content under their Terms of Service and make no effort to moderate it.
- The stakeholder also observed that children and young people’s engagement on social media services presents a variety of risks beyond access to adult content, including exposure to hateful and violent content. They expressed the view that, despite these risks, many social media services do not verify their users’ age.

Adult industry: Australian stakeholder group

Consultation group overview

Stakeholders: Australian-based adult industry.

Consultation date: 18 November 2021 and 11 February 2022.

Overview: This stakeholder group discussed age verification and assurance and the adult industry in Australia, the risks associated with new technological measures, current 'age-checking' practices, language and definitions and provided recommendations.

Adult industry in Australia

Stakeholders highlighted the Australian-based adult industry is different to the international adult industry. Producers are often:

- women
- members of the LGBTIQ+ community
- operate independently.

These factors should be considered, to ensure measures are not solely designed to regulate larger, international websites and companies.

- Stakeholders indicated production of paid/subscription-based content was more common than free 'tube' sites[6] in the local industry. A paid model already requires more checks and barriers before users can access content (for example, age verification through credit cards).

Current 'age-checking' practices

- Stakeholders noted there are already using existing technological solutions to control access to adult content. Stakeholders shared examples of current industry practices. This included:

- Sites complying with guidance from the Association of Sites Advocating for Child Protection (ASACP). These sites have home or index page warnings that indicate the

[6] A 'tube' site refers to a site which is free to view and relies on user-uploaded content, rather than the operator uploading content.

site contains adult content and also use meta tags[7], which allow their sites to be blocked if parental filters are active on devices.

- Many sites have paywalls, requiring credit card details for page access and payment. Payment companies check for meta tags on sites.
- Sites also often have features which make it difficult for users to download and share videos outside of the protection of the paywall. Producers can blacklist users and ban IP addresses that share content to other sites. This can prevent the content from being seen outside the paywall/age gate.
- Stakeholders reiterated it is industry standard to use the 'Restricted to Adults' meta tags, which allows search engines and safety software to identify sites containing pornography.

Risks associated with new technological measures

- Stakeholders in this group felt there should be explicit measures to prevent age assurance or technology companies from storing users' data. Stakeholders were concerned about the influence of large, international adult industry members who may have a commercial benefit to collect and sell user data.
- Stakeholders said that the use of third-party age assurance services (that is, assessments conducted by a party other than the site or platform hosting the content) also raises risks for sex workers by giving third-party providers significant power of their businesses. For example, they noted major payment processers already impose, and enforce, certain policies that can be challenging for adult content providers to navigate and comply with, resulting in legal content and business activity being blocked. The adult industry has very few avenues to review third parties' decisions or feed into their policies and would welcome further opportunities for engagement.
- Stakeholders pointed out that many adult content producers and sex workers operate as micro-businesses or sole traders, and may have limited technical expertise and financial resources to apply new requirements. They felt that it is unfair and onerous to expect them to implement complex technological solutions.
- Stakeholders identified the following risks from their perspective in applying age verification measures:
  - The potential for scope creep and interference with legitimate access to adult content and sex education material. Stakeholders raised that measures which are designed to prevent children and young people from accessing pornographic

[7] Site owners can include information about their site in 'meta tags' in their website HTML code. These tags are not visible to users on the site, but filter software and search engines can read them and identify that the site is for adults only.

content could, unintentionally prevent access to educational material. They also expressed measures which restrict access to these sites could be employed to block types of legally permissible adult content from adults on other moral or religious grounds. ,

    - The over-collection of information – that is, stakeholders felt the implementation of child protection measures may lead to the collection and storage of private data (name, address, date of birth) when it is only necessary to record an age attribute (such as ‘over 18’).
    - Concerns with the potential use of facial recognition or analysis technology for the purpose of assessing age.
        - Stakeholders noted when platforms verify uploaders, often they do not provide users with information about the third-party company processing their information, nor how their data is assessed or stored.
        - They raised that artificial intelligence (AI) biometric technologies can present issues with ethnic and racial biases.
        - They highlighted that children and young people may not be aware of, or consent to, their data being collected and used for the purpose of other AI-based age estimate technology, such as key-stroke pattern analysis.
        - They emphasised that there would need to be substantial trust and transparency requirements for any third-party assessment tool.
    - Data surveillance of users – they raised that both consumers of content, as well as performers and sex workers may be at risk of being tracked, stalked or abused if their personal information is made publicly available.
    - The logistical and financial impost of compliance, which they felt would likely have a disproportionate impact on smaller local producers.
- Stakeholders suggested that:
    - placing age-barriers around moderate pornographic content produced by compliant sites could deter users and instead funnel them to fringe or extreme sites, or sites which host content that is not ethically produced.
    - a reliance on content moderation to reduce children and young people’s access may have unintended consequences. For example, they stated that AI moderation on platforms is often imprecise and over-moderation tends to disproportionately affect marginalised people and communities, such as sex workers.
    - measures to address adult content online often come at the expense of sex workers’ safety and their ability to work.

Challenges of a global ecosystem

- Stakeholders highlighted the local industry works within the Australian legal framework; however the policies of large social media services are based on laws and settings in other jurisdictions, which often negatively impact the local adult industry. They recommended that services should adhere to the legal settings of each jurisdiction they operate within.
- Stakeholders shared specific challenges they have encountered in their dealings with social media services. They expressed concern that these issues may worsen should age verification requirements be introduced:
  - Difficulties with social media companies – stakeholders find it difficult to engage with social media companies.
  - Being de-platformed – adult content producers and sex workers have been deplatformed from social media or other digital platforms, despite working within the Terms of Service for platforms.
  - No warning – stakeholders reported receiving no advance warnings from services to changes in Terms of Service or moderation practices. These changes can have significant impact on how they conduct their business practices. Notice of changes would enhance their ability to comply.
  - A gap in service reporting processes – services also offer limited means for reviewing decisions. Stakeholders felt that they are subject to arbitrary or vexatious complaints but have minimal options for challenging those complaints.
- Some stakeholders would like to see a liability introduced for poor moderation practices, as well as best practice guidance on content moderation with input from the adult industry.

Language and definitions

- Stakeholders reflected on the importance of terminology and how it will frame the issue and AV roadmap report:
  - They felt that using the term ‘exposure’ misses the behavioral triggers involved in seeking access to adult content online. Stakeholders also felt that it has the potential to stigmatise pornography as something inherently negative, in the same way that society talks about being ‘exposed’ to disease.
  - They preferred ‘accidental’ viewing or exposure as a term, which reduces the stigma around adult content producers and does not position pornography producers as preying on children.

    - They sought greater clarity around the type of content which would be captured by any regime, particularly on social media platforms. Stakeholders recommended that the report must define pornography and what content is intended to be covered under the AV roadmap.

Recommendations

- Stakeholders felt that improving sex education was a preferable way to address the issues associated with young people accessing pornography.
    - Stakeholders advised that young people have the right to privately explore their own sexuality, and measures that are education-based would be preferable. They expressed that current porn education contains anti-sex work messaging, which may stigmatise sex workers.
    - Stakeholders suggested that eSafety could play a role in producing educational materials, such as resources for navigating adult content, relationships, and targeted content for people with diverse sexual and gender identities.
    - Stakeholders noted that the harms of pornography are contested and suggested greater clarity was required on the definition of pornography, the harms the AV roadmap seeks to address, the relevant age groups, and how success would be measured.
    - Stakeholders suggested curriculum development should be informed by engagement with children and young people. They highlighted that sex workers should also have a role as stakeholders and collaborators in education development, not perpetrators.
- Stakeholders felt that parental controls and family filters may provide a more targeted response. There were concerns about blocking sites at an internet service provider (ISP) level, noting this could censor whole domains. If ISP-level measures were to be considered, stakeholders felt they should be opt-in measures for consumers rather than automatic.
- Concerns were canvassed on regulatory responsibilities falling on ISPs and private companies (such as banks), which may lead to limited recourse for appeal for smaller industry members. Stakeholders felt that accountability measures for ancillary service providers should be made explicit in the AV roadmap.
- Stakeholders expressed that the current regulatory framework was problematic, particularly the reliance on the National Classification Scheme, which was seen as outdated. They suggested postponing implementation of the AV roadmap until after the classification review is completed.

- Stakeholders suggested:
  - More could be done to strengthen existing technologies, such as parental controls and search filters. They suggested government should provide financial support for low-income families to purchase filtering software, as well as provide more educational resources for parents to support the installation and use of such technology. Stakeholders felt that increasing parents' digital literacy would provide them with greater agency and choice over how they raise their children. Stakeholders also expressed that these technologies are less intrusive on people's online activities.
  - The AV roadmap should consider a tiered, proportionate approach to best practice and enable industry to choose the most suitable age assurance methods (based on their size and other risk factors). Search engines already have the means to identify adult sites. They felt this could be leveraged more effectively.

Academics: Australian stakeholder group

Consultation group overview

Stakeholders: Australian academics.

Consultation date: 26 November 2021.

Overview: This stakeholder group discussed age verification and assurance in a wider context, young people's access to pornography and education.

Wider context

- Stakeholders expressed that the AV roadmap development process should consider and provide recommendations for the wider socio-political context. In particular, they drew attention to regulatory regimes affecting the pornography industry, the criminalisation and licencing of sex work, as well as attacks upon LGBTIQ+ communities. Examples of relevant factors raised by stakeholders included:
    - The pending review of the classification system – stakeholders highlighted that the current classification scheme prohibits consensual fetish content and therefore has the potential to censor queer and differently-abled sexual practices.
    - The discriminatory policies of financial institutions in relation to pornography and sex work – stakeholders expressed that these policies can prevent the sale of much independent sexual content and thereby place workers in economic precarity.
    - The introduction of religious discrimination legislation – stakeholders spoke about the way that these laws may been seen to sanction discrimination against LGBTIQ+ people.
    - Efforts to limit gender diversity in education – stakeholders advised this impacts the availability of comprehensive, culturally relevant sex, consent and relationship education.
- Stakeholders highlighted the inconsistency between the age to access to pornography (18) and the age of consent to sex (which varies across states and territories but is consistently under 18). They felt that the age of consent could be used to guide the age of access to online pornography.
- Stakeholders observed platforms and services have a role to play and can set the tone and norms for content by improving upon individual government approaches. They felt this could be achieved by taking guidance from relevant international human rights

instruments which encompass sexual rights obligations pursuant to the rights to health, expression and non-discrimination. They stated more could be done to demonstrate and uphold contemporary community standards, including queer, feminist and sex-positive approaches to sex and sexual content.

- Stakeholders highlighted age verification or assurance measures (including age estimation), which use bio-analysis, can collect significant amounts of data from users, including young people. They held significant concerns in relation to young people being able to consent to providing this information or withdrawing access to it.
- In terms of current practices to limit unintentional access, stakeholders noted that 'age gates' (where a site or service asks a user to enter a birthdate or confirm their age) can be effective barriers by creating a point where a decision must be made to access the content. Even though a user may provide a false answer, stakeholders noted the user must make a conscious decision to do so – which can prevent young people from stumbling upon content unintentionally.
- Stakeholders pointed to the need for a regulatory framework that encourages the decentralisation of internet infrastructure, ends financial discrimination by banks and payment processors, and prevents discriminatory platform Terms of Service. They felt this could prevent pornography tube monopolies from forming, and consequently enable independent pornography providers to flourish.

Access to pornography by young people

- Discussion on the impact of pornography on young people and how young people access pornography focused on the following themes:
    - Stakeholders emphasised that non-consensual exposure to pornography is a discrete issue, which is already addressed under the law, including through a variety of criminal offences. They noted that there could be better enforcement of existing laws, for example, in relation to pop-up advertisements.
    - Stakeholders pointed out the problem in talking about 'exposure' to pornography as if it is inherently harmful. They noted that a distinction is often made between children's 'accidental' and 'intentional' access to pornography. They felt it is more meaningful to discuss access to pornography pre-puberty versus post-puberty. In their experience, prepuberty access tends to occur as a result of stumbling upon the material. Their research demonstrates that where this experience results in harm, that harm often stems from the subsequent negative reactions of trusted adults (such as anger or shame) as opposed to the material itself.
    - Their evidence shows that it is common for young people to seek out sexual material post-puberty, and that this behaviour is not harmful. They highlighted

    - research showing that young people want information on how to have pleasurable sex and there is a gap in mainstream sex education for this material.
    - Stakeholders pointed to a recent interdisciplinary literature review, which found no evidence of harm as a result of post-pubescent viewing of pornography. They noted that some research categorises sexual adventurousness as a type of harm associated with pornography, but emphasised this is based on a narrow and potentially discriminatory definition of healthy sexual activity. They pointed to the need to avoid heteronormative, ableist and kinkphobic assumptions about behaviour, and to more carefully examine the label of 'unrealistic' sexual activity.
    - Stakeholders expressed that whole categories of sexual activity should not be deemed inherently harmful merely because they carry a risk of physical harm. They identified a lack of education and support available to inform and guide safe behaviours as a key factor contributing to the risk of physical harm.

Education

- The importance of improving and expanding sex education was also discussed by stakeholders:
    - They noted that pornographic content can be educational, including content which models and depicts consent and safe practices. In particular, they emphasised that pornography can be validating and affirming for LGBTIQ+ people. They expressed the current restrictions within the classification system, and the criminalisation of production in some jurisdictions, limits the availability of content which explores these themes.
    - They felt resources being spent investigating mandatory age verification could be diverted to address these issues instead.
    - Stakeholders stated young people can be savvy, critical consumers when accessing pornographic material if they are properly equipped to do so. They felt the focus should be on the systems and environments supporting young people.
- Stakeholders highlighted that education measures could:
    - address both digital literacy and sex, relationship and consent education
    - explain the social, cultural, historical value of pornography, and include information about its regulatory and labour frameworks. Stakeholders noted that discussions about working conditions can help humanise performers and challenge harmful norms, such as expectation of access to women's bodies, as well as encourage consumers to support independent producers and sex worker rights

- include resources which equip adults to have supportive and helpful discussions with the young people in their lives without shame or stigma
- be co-designed by young people and youth-centred for the adults in their lives. Stakeholders noted recent experience with this particular model and that it had been effective.

Children and young people's wellbeing: International stakeholder group

Consultation group overview

Stakeholders: International academia, advocacy and youth wellbeing groups.

Consultation date: 24 November 2021.

Overview: This stakeholder group discussed age verification, the impact of access to pornography by young people, technological solutions and responsibility, and education.

Pornography and young people

- Stakeholders advised that research in this field is highly contested, and it is unlikely a consensus will be reached on pornography's impact on young people.
- Stakeholders observed that differentiating between 'accidental' and 'intentional' access to pornography is not necessarily helpful. They felt the use of the term 'intentional' may invite blame and stigma, while 'accidental' may not reflect realities of teen culture in relation to sexuality and pornography use.
- Stakeholders noted it can be difficult to measure the adverse outcomes of exposure to pornography as there are ethical challenges involved in researching young people's relevant experiences.
- Stakeholders recommended that the AV roadmap could consider different approaches to age assurance and access to material based on age, highlighting differences in what is appropriate for a 15 to 16 year old compared to what is appropriate for a 10 to 11 year old.
- Stakeholders reflected on the various scenarios through which young people may be exposed to explicit content outside of visiting pornography websites (for example, through online chat services, older students or siblings, or website pop-ups).
- Accordingly, they called for a multi-layered approach to the AV roadmap, looking at all likely means of exposure.

Technological solutions and responsibility

- Some stakeholders expressed the view that stringent age verification controls should be imposed alongside educational measures. These stakeholders felt that system-level controls are more likely to be effective at preventing children and young people's access to pornography.

- Stakeholders supported the promotion of parental controls and filtering, but believed that reliance on these measures alone would not be sufficient. They emphasised that these measures currently exist but have not been effective in mitigating children and young people's exposure to adult content to date.
- Stakeholders highlighted technological solutions should reflect the nuanced nature of content and how content circulates on the array of apps and services that children and young people use.
- Stakeholders discussed the importance of creating friction in the user experience for children and young people seeking to access adult content. They emphasised that creating boundaries – even those that can be easily overcome, for example, by entering a false birthdate – can help children understand the context of what they are accessing and that it is not meant for them.
- Stakeholders raised developments in the United Kingdom calling for minimum standards for age assurance technology, including for the technology to be privacy preserving. They felt that similar considerations should inform the development of the AV roadmap within the Australian context.
- Stakeholders suggested internet service providers (ISPs) could play a role in implementing measures to block access to content or assessing the age of users.
- Stakeholders highlighted parents who are less familiar with these issues, and less comfortable using technology, can be more restrictive of their children's internet use due to fears of online harms, including exposure to inappropriate content. They noted this can result in children and young people missing out on beneficial educational content and social interactions, and emphasised the importance of a balanced approach.

Education

- Stakeholders suggested that digital literacy and pornography education should start from approximately 11 to 13 years old.
    - They advised health, sexuality and relationships content should be compulsory in school curricula.
    - They highlighted the benefits of education delivered by third parties in the school context.
    - They expressed resources should be free and scalable, and suggested eSafety could play a role in their creation and/or distribution.
- Stakeholders stated that children and young people need help contextualising what they see online and building resilience to confronting content. They saw these skills as a key

element of violence prevention, enabling young people to critique the representations of men, women, power, and pleasure in pornography.

- Stakeholders highlighted that education measures should be co-designed with young people and/or consider the views of young people.
- Stakeholders felt that attentive and positive discussions are most effective in communicating with young people about sex and pornography. They identified honest and open engagement as critical for educators and parents to assist young people transitioning to adulthood.
- Stakeholders also suggested that educational programs should provide more information about health risks (for example, physical injuries) and legal risks (such as, the criminal offences that may apply to the distribution of explicit images) of sexual or intimate activities.
- Stakeholders suggested that government should offer training and resources to parents, carers, educators, frontline health workers, social workers and trusted adults in children and young people's lives.

Children and young people's wellbeing: Australian stakeholder group

Consultation group overview

Stakeholders: Australian advocacy and children and youth wellbeing groups.

Consultation date: 29 November 2021.

Overview: This stakeholder group discussed age verification and assurance, pornography education in school-based settings, as well as safe digital platforms for children.

Pornography and young people

- Stakeholders shared their experiences working as educators and advocates who speak about pornography with children, young people and schools. They have:
  - received requests for content from primary schools and reported a growing demand for sessions in boys' schools;
  - observed an increase in reports from teachers and students of peer-to-peer sharing of pornography in school environments (for example, showing porn videos on the school bus);
  - responded to frequent requests from girls asking for help in negotiating boundaries and seeking advice on sexual assaults they have experienced at school (for example, groping or up-skirting);
  - perceived that a culture of pornography is normalised among young people. For example, students watching content on laptops at school; and
  - identified that many people do not have an awareness of the extent of harmful pornographic content that is available online.

Digital platforms

- Stakeholders emphasised that online spaces should be safe for children and young people to explore and express themselves. They felt there was scope to improve regulation of large technology companies, so they provide a safer experience for young users.
- Stakeholders recommended a range of measures, including better access to parental filters and more robust content moderation on social media services, particularly in

relation to recommendation algorithms, which they felt may serve up inappropriate content to young users.

Educating in a school-based setting

- A holistic approach – stakeholders expressed that a whole of school approach to pornography education is beneficial but should be structured by age-appropriate bands of content. In their experience, schools are eager for external support to build awareness and critical thinking around pornography and respectful relationships. In addition to third-party presenters, stakeholders stated that schools need support to have the right procedures, policies and staff resources in place to prevent and address incidents in a timely way.
- Third-party presentations – stakeholders observed that third-party presenters may be able to share their own experiences with pornography more openly (compared with teachers) and this can encourage a more open dialogue with students.
- Student-led learning – stakeholders highlighted the benefits of student-led learning strategies. For co-ed and sister/brother schools, they felt this gives students opportunities to work together and relate to each other outside of a party context.
- Resource support for teachers – anecdotal reports from teachers highlight that they feel inadequately prepared and resourced to discuss pornography with students, and are in need of support to personally process trauma, while also guiding students.
- Upskill support workers – stakeholders noted that marginalised young people are more likely to have bad experiences online. They felt that residential care, youth and allied health workers should be upskilled on online safety issues, the signs that a young person may be having negative online experiences, and how to respond.

Safety technology: Australian stakeholder group

Consultation group overview

Stakeholders: Australian age verification technology providers, safety technology providers, and internet service providers.

Consultation date: 30 November 2021.

Overview: This stakeholder group discussed age assurance and verification and safety technology, technical intervention and risks, education and resources, and privacy.

Age verification and assurance in practice

- Stakeholders were supportive of a holistic, multi-layered approach to the AV roadmap. They agreed that all segments of the online ecosystem have a role to play in enhancing children and young people's online safety and restricting their access to online pornography.
- Stakeholders raised there are many features and elements of social media and gaming apps that can have health implications for children and young people beyond pornography, and that government should also consider acting on these matters.
- Stakeholders also noted there are no 'silver bullet' protective measures. They emphasised that all measures can be circumvented in some manner, and that it will be difficult to get participation from all sites and services online that could contain pornography.
- Stakeholders advised there are effective filtering programs available, but consumer demand and uptake of these programs appears to be relatively low. They suggested this could be due to a lack of interest, awareness or understanding of how to activate such programs, or possibly the associated cost. Stakeholders also raised that some parental controls are easy for young people to remove once they are over 13.

Technical intervention recommendations and risks

- Stakeholders proposed the AV roadmap should look up and down the stack, including websites, device manufacturers, operating system developers and internet service providers (ISPs), and consider different levels of intervention at each layer.
- Stakeholders supported device level filtering. They advised device level filters can reduce privacy concerns and the number of actors required to implement compliance measures.

- Stakeholders said device developers need to be effectively engaged in the roadmap process to ensure parental controls and age restrictions are being applied effectively.
- Stakeholders raised there are gaps that could be addressed with existing technology – such as implementing filters on devices provided to children in school settings and imposing restrictions on public Wi-Fi.
- Stakeholders highlighted the impact of developments by device and operating system manufacturers. They warned that unilateral decisions from major companies regarding privacy and encryption settings could undermine the effectiveness of some filtering systems and age assurance interventions.

Education and resources

- Stakeholders advised there are existing resources provided by ISPs and telecommunications companies to support parents in creating safer online environments for their children. For example, they pointed to filtering programs and resources on how to use them and how to report illegal content. Stakeholders believe these resources could be better promoted to the public.
- Stakeholders suggested an online hub could be established as a 'one-stop-shop' for parents to access information and online safety tools or programs.

Privacy

- Stakeholders noted that trust is critical, and that any technological measures should be privacy preserving and data minimising. Stakeholders also preferred third parties to hold age assurance data, rather than adult websites.
- Stakeholders suggested that the focus should be on age assurance, not age verification, measures. They noted that only the 'over-18 attribute' is important for purposes of restricting children and young people's access to online pornography, not the user's identity or specific age.
- Stakeholders considered biometrics-based solutions (such as facial recognition or analysis) to be highly problematic for individual privacy.
- Stakeholders advised obtaining consent from a child, to process age determination data, is very complex and would differ based on jurisdiction. They advised this would be challenging for international sites and services to implement.

Safety technology: International stakeholder group

Consultation group overview

Stakeholders: International age verification technology providers, safety technology providers, and internet service providers.

Consultation date: 30 November 2021.

Overview: This stakeholder group discussed age verification and safety technology, including standards and frameworks, existing and future challenges, trust and privacy, as well as current and potential solutions.

Standards and frameworks

- A 'technology neutral' standard – stakeholders advised that work is underway on an international standard for age assurance technology. The standard would be 'technology neutral' and could be used to assess existing and emerging technologies.
- Interoperability – stakeholders stated that interoperability is crucial across national ecosystems and frameworks. They believe that mutual recognition of national standards should be a focus of regulators.
- Adopted by users – stakeholders noted that trust is critical for the adoption and success of age assurance and verification technologies by users. They suggested that a mechanism or resources be established by government, for consumers to confirm that service providers are legitimate, can be trusted and meet government's minimum standards.
- European Union project – stakeholders shared information on the European Union's EU Consent project, noting the project aims to develop an interoperable network of age assurance providers where the provider will be certified to a specific standard.

Existing and future challenges

- Individual identity information – some stakeholders felt it was not appropriate to use individual identity information to access pornography sites as this may not be the most privacy-preserving or data-minimising approach.
- Public awareness – stakeholders acknowledged the challenge of public awareness and educating people on what age assurance and verification tools are, how to use them, and which tools are safe and reliable.

- Data disclosure – stakeholders suggested increased data disclosure may lead individuals to attempt to evade compliance measures. They suggested that this could lead pornography users (and potentially children) to access non-compliant platforms, which may host more extreme and harmful content.
- Permissions and content filtering – stakeholders flagged challenges in some services allowing users who are 13 and over to access their platforms, while also permitting content that may be inappropriate for those young users under their Terms of Service. They recommended greater exploration of content moderation and filtering on such platforms.
- Risks to children and young people – stakeholders noted the importance of embedding a children's rights lens and considering the risks to children and young people from product features at a granular level. Stakeholders suggested that all content that is agerated should be age-gated in some way, to create experiences online that are age appropriate and safe.
- Costs – stakeholders reflected that in other spaces, such as online gambling which follows Know Your Customer (KYC) regulatory requirements, business models were able to adjust for the additional costs of verifying users. Stakeholders noted that this may not be the same for all pornography sites – some of which are small scale businesses. Despite this, stakeholders observed that the cost of age verification and age assurance had decreased significantly over time, and that this trend would likely continue as the age assurance tech industry grows.
- Small businesses – stakeholders suggested that small businesses operating pornography websites under a particular revenue threshold could be given access to age assurance solutions for free or at a lower subsidised cost.
- Digital rights – stakeholders suggested that a perceived imbalance between user privacy, safety and security enhances aversion to adopting and accepting child protection measures. They submitted that the digital rights of both children and adults need to be actively promoted and explained in relation to any technological solutions recommended under the AV roadmap.

Current and potential solutions

- Stakeholders highlighted that there are age assurance technologies which operate entirely on individual devices. They explained that this allows age estimation/assurance to occur on a device, or a device's internet browser, preventing an individual's underlying data from moving across different servers. This assists apps or services accessed on that device in gauging the user's age attribute.

- Stakeholders explained that identity verification is not necessary for age verification because age attributes can be decoupled from identity information. Stakeholders suggested that this can be achieved through vectors of trust such as 'zero knowledge proofs'.
- Stakeholders suggested there may be an opportunity for re-usable digital IDs, specifically those that allow people to selectively share age attributes only.
- Stakeholders noted that if interoperable measures are introduced, then consumers who have already done age checks for other reasons, such as purchasing alcohol, could re-use that age check to access adult content.

Trust and privacy

- Stakeholders advised that enabling choice for consumers is critical for building trust in age assurance measures.
    - Stakeholders suggested that a prerequisite for establishing user trust is to ensure there are 'no surprises', meaning there should be transparent disclosure and explanation about why the age assurance or verification measures are in place, how they work and how data is processed and/or stored.
    - Stakeholders noted that the average person may be more likely to engage with information that is available on the platforms and services they frequently access, rather than seeking out information on a government website. They proposed that information on regulatory requirements, platform safety measures and age assurance or verification certification details should be shared by platforms and services.
- Stakeholders discussed that many age assurance or verification solutions that rely on government ID are appropriate for use cases such as gambling or purchasing alcohol. However, stakeholders acknowledged that accessing pornography is a very different user experience (for example, tube sites often do not require accounts or payments).
- Stakeholders noted that companies interested in using age assurance technologies still hold concerns about providing data to third parties.

Age Verification Roadmap Consultations: Round 2 (October 2022)

This document contains insights and feedback that eSafety received from industry, stakeholder groups and experts on options to address the risks and harms associated with children's access to online pornography.

In June 2021, the Government responded to the House of Representatives Standing Committee on Social Policy and Legal Affairs' report on age verification for online pornography, tasking eSafety with the development of a comprehensive roadmap exploring 'if and how a mandatory age verification mechanism or similar could practically be achieved in Australia'.

In August 2021, eSafety issued a call for evidence and received 33 submissions. A thematic analysis of these submissions is available on our website.

Between November 2021 and July 2022, eSafety conducted targeted consultations with a wide range of stakeholders across different sectors. These sessions allowed for closer examination of the information submitted in response to the call for evidence. They also provided a forum for experts to share broader insights. The evidence received through consultation will support the development of recommendations that are holistic, viable and proportionate.

This document contains summaries of consultation sessions that were, for the most part, conducted between March and July 2022. Summaries of the previous consultation sessions, largely held between November 2021 and February 2022, are available on our website.

The information in this document represents a variety of views that were discussed by stakeholders during the consultation sessions and these have been consolidated into key themes. The views and opinions in these summaries are those of the stakeholders and do not reflect eSafety's position. They will help to inform the development of the roadmap.

More information about the roadmap is available on our website.

Children and young people's wellbeing: Australian stakeholder group

Consultation group overview

Stakeholders: Australian academia, youth wellbeing groups and anti-violence advocacy groups.

Consultation dates: 24 January, 4 February and 22 June 2022.

Overview: These stakeholders discussed age verification, children and young people's access to pornography, and educational approaches, measures and opportunities.

Children and young people's access to pornography

Comments from the first stakeholder

- This stakeholder was of the view that watching pornography is not inherently harmful for young people.
    - The stakeholder advised that young people have natural interests and curiosities that can be explored through sex-positive pornography in a healthy and developmentally appropriate way.
    - They suggested that government may not be best placed to determine what constitutes sex-positive pornography.
- However, this stakeholder did note that young people who had not received any education or support about online pornography may be more likely to be harmed when viewing it.
- In addition, the stakeholder felt that 'accidental' interactions with pornography should be differentiated from 'intentional' interactions. They offered research demonstrating that the median age of young people accidentally viewing online pornography is markedly younger than the median age of those who intentionally view the content.

Education measures and opportunities

Comments from the first stakeholder

This first stakeholder added that:

- most young people can comprehend that pornography is not 'real', but they have difficulty identifying which aspects of it are unrealistic (this observation was based on the stakeholder's own research)
- educational measures are key to providing countervailing information and can help to inform young people about healthy sex and relationships
- they support the implementation of some measures to restrict young people's access to online pornography, however, do not believe that age verification is the most appropriate measure for balancing freedom of access to information and protecting young children from harmful content
- there is research indicating that young people oppose content blocking tools but support education programs addressing healthy pornography consumption.

Educational approaches

Comments from the second stakeholder

The second stakeholder made the following points:

- Education about pornography and sexuality should be sustainable and integrated into schools.
- Teachers should be well trained, supported and appropriately resourced to provide this education. They acknowledged that every teacher's level of experience will vary - some will be confident, while others may feel unsupported, lack confidence, be unsure about how to have appropriate conversations, and know very little about contemporary pornography.
- External presentations, if they occur, should be complementary to lessons embedded within the curriculum, not a substitute for lessons currently addressing these issues.
- This stakeholder also pointed to advice from the United Nations Educational, Scientific and Cultural Organisation (UNESCO) which supports an integrated approach to education, rather than relying on external presentations. UNESCO says that content should be embedded in the school curriculum, as age-appropriate content can be delivered every year and contextualised within the broader curriculum. The stakeholder also expressed the view that:
    - one-off external presentations – which rely on an individual presenter – can be ineffective pedagogically (as they are often in the form of a lecture, rather than participatory and learner-centred). There could also be challenges associated with quality assurance, highlighting the need to ensure that external presenters use a sound conceptual framework and avoid a shame or fear-based approach.

    - young people may be more comfortable participating in an external presentation and willing to volunteer information and have conversations about pornography with a person they do not know, rather than with their teachers. However, it was noted how difficult it can be for a third-party to provide ongoing support and follow up with students. It can also be hard for teachers and students to have sensitive conversations. Accordingly, this stakeholder highlighted the importance of supporting teachers to develop confidence in talking about these issues.
- The stakeholder noted that educational content regarding pornography aligns with the health curriculum, however, there is also value in finding ways to integrate discussions and learning outcomes into media and digital literacy lessons as well.
- The stakeholder said that it was important to consider the welfare of children when introducing these types of topics.
    - They suggested that schools may feel comfortable initially starting education about pornography in Years 9 to 10, as that is where the connections to the curriculum are the strongest. However, as parents and school communities are brought along on the journey, they noted there are opportunities to introduce content addressing pornography in earlier years, using less explicit language. They felt this would allow the lessons to build on each other and enable age-appropriate learning at each stage.
    - They emphasised the importance of educational measures being age-appropriate and child-focused.
    - The stakeholder reported that, in their experience, many schools have expressed an interest in obtaining information and support to address the topic of speaking with students about pornography, healthy sexual behaviours and incidents involving pornography among students or on school grounds.

Young people's access to and use of pornography

Comments from the third stakeholder

The third stakeholder made several points and they:

- emphasised the gendered nature of pornography. They noted this can include rigid and hierarchical gender stereotypes and problematic notions of power and consent. They suggested this was an important lens for consideration within the roadmap
- noted that their work was not focused on stopping young people from watching pornography, but rather encouraging them to engage with it critically
- felt that a holistic and gendered approach was needed to address the impact of pornography on young people – particularly considering the different ways in which girls

and boys are affected. This would entail working with young people as well as parents, schools, the media and government

- advised that their research shows young people hold conflicting and nuanced views on pornography. They can find it concerning, but also informative; find it enjoyable, but also be critical of it.

Messaging and engaging young people

Comments from the third stakeholder

- The third stakeholder creates resources for young people and made the following comments:
  - From their research, young people and children commonly get information about sex and pornography from their peers and pornography websites.
  - They have found a peer-to-peer and youth-led learning approach is effective, noting that young people have capacity and resilience to navigate complicated issues themselves, with the right information and support.
  - It is important that resources are online, as this is where young people are searching for information.
  - There is a need to tailor resources differently for different genders, as they seek information from different sources.

Privacy and digital rights: Australian stakeholder group

Consultation group overview

Stakeholders: Australian digital rights advocates.

Consultation dates: 4 March and 12 July 2022.

Overview: These stakeholder groups discussed age verification and data, privacy and technological measures, scope and education, standards and compliance measures and provided recommendations for the design of an age verification regime.

Data, privacy and technological measures

Comments from the first stakeholder group

- These stakeholders raised concerns about the collection, use and storage of personal information, noting that once data is collected, it has the potential to be misused. The group also:
    - suggested that the risks of harmful consequences, data breaches and privacy intrusions should be considered. Similarly, the point at which the risk of harmful, unintended consequences for privacy outweighs the age verification regime's objectives to prevent harm to children, should be made clear from the outset
    - said it could be difficult to 'turn off' aspects of a mandatory age verification regime, should substantial data and privacy breaches occur
    - did not support the collection of biometric information for purposes of estimating or verifying the age of users
    - expressed apprehension about the uptake of measures which may normalise surveillance, as well as introducing technological tools that could be linked with other identity systems.
- The stakeholders felt that content moderation at scale, particularly systems with little to no human intervention, are of significant concern. Stakeholders flagged the challenges and risks of automated content moderation at scale, emphasising the risk of overblocking and stigmatising marginalised groups.
- Stakeholders stated that age assurance technology providers should be transparent about their systems and processes. They felt that the public should have access to

evidence which substantiates the effectiveness of providers' technologies and demonstrates adherence to information privacy and security standards.

- Stakeholders advised that interventions which consolidate power in large platforms and inhibit competition should be avoided.

Comments from the second stakeholder group

- One stakeholder suggested that it is nearly impossible to identify every stakeholder that may be affected by an age verification regime, advising that some parties may not realise they are impacted. They said that even though it might be difficult, it would be important to help those parties understand any risks which may affect or apply to them.
    - They also felt it was important to understand that even if the risk to an individual may not be significant in aggregate, the risk to a number of people may be significant. There was also a need to consider the risk of projecting unknown and unidentified risks onto individuals.
    - A point was made that the roadmap should acknowledge the context and implications of other movements to identify people on the internet, and how age verification may intersect with those discussions.
- Another stakeholder suggested eSafety should consider options that aren't 'perfect', although it would be necessary to determine an acceptable error rate. The stakeholder noted the need for fast, empathetic and caring appeal or solution processes for groups more likely to be impacted by potential errors.
- The stakeholder also said systems, such as the verified credentials proposed in New South Wales, could be a potential measure for consideration – particularly where the system shares verified credential or attribute information only (for example, if a user is 18 years or older), and not other unnecessary data (such as a birthdate).
- Stakeholders supported a data-minimisation approach and emphasised the right to participate online free from surveillance, arguing that safety and security can be promoted by collecting, using, storing and sharing less data.
    - Stakeholders identified specific privacy risks raised by age verification measures, including that the risk of re-identification is possible even from seemingly anonymous data.

Scope and education

Comments from the first stakeholder group

These stakeholders:

- emphasised the need to create a clearly defined scope for the roadmap to reduce the risk of capturing and restricting more content than necessary
- felt that the age of consent and the evolving capacity of young people as they grow up should be considered when determining access to sexual and pornographic online content. Stakeholders expressed that not all pornography is harmful and that access to this content can form an important component of sexual education, particularly for diverse or marginalised communities for whom generic sex education may not be relevant or inclusive
- supported the role of respectful relationships education and holistic sex education and its inclusion in the roadmap's recommendations
- said that education is a critical component to building digital literacy and minimising any negative impacts stemming from underage access and the use of pornography
- were concerned about any regime which might rely solely on technological solutions to address what they consider to be fundamentally a social problem. Stakeholders preferred to prioritise educating and empowering children and young people and their families.

Standards and compliance

Comments from the second stakeholder group

- Stakeholders held mixed opinions and discussed the role of the International Organization for Standardization (ISO) and other international standards, and the potential for independent or government certification.
- Stakeholders noted the value of alignment with ISO and international standards. They said, however, that a conflict of interest could potentially be a risk to standards and certification processes.
- The group discussed the need for continuous auditing processes for the duration of age verification implementation to identify where protections fail, and where measures cause mass risk to people's privacy or other risks. They suggested that:
    - any regime should be non-prescriptive about standards and any testing and auditing authority should be kept separate from legislation
    - the relevant regulator should not inspect every age verification service but rather should be the backstop to third-party auditing
    - the cost of independent auditing should be absorbed by the online industry as a routine cost of doing business.

- Stakeholders raised general concerns about the adequacy of current legislation to protect individuals’ privacy.

Summary and recommendations: Privacy and digital rights

Comments from the first stakeholder group

- Stakeholders in this group indicated that:
    - the legislation should include a review period and/or should have sunsetting provisions
    - there should be clear consequences for data breaches or other misuse or mishandling of data
    - accessible and affordable legal avenues should exist for individual and collective rights of action against privacy breaches and other forms of possible harm
    - the roadmap should empower parents and carers to make choices suited to their individual circumstances.

Comments from the second stakeholder group

- A stakeholder in this group noted that a multifaceted approach is needed, suggesting that a package of holistic measures would be most likely to achieve a measured and effective response.
- Another stakeholder stated that a regime should not be ‘set and forget’, especially in relation to privacy and security requirements. They expressed that it will need to be developed iteratively, as many risks and consequences may not be apparent until the implementation starts. o Stakeholders thought it was important to continue to consult with stakeholders through the implementation phase, particularly on solutions to issues that may arise over time.
- A stakeholder suggested designing friction into the process of changing or expanding age verification measures to manage associated risks – for example, requiring the Australian Parliament to embed measures in any primary legislation, not just under subordinate legislation.

Education: Australian stakeholder group

Consultation group overview

Stakeholders: Australian education authorities.

Consultation date: 10 March 2022.

Overview: This stakeholder group discussed age verification and children's access to and use of pornography, support for schools and teachers, broader educational measures and current technological approaches to reduce students access to pornography in schools.

Children's access to and use of pornography

- Stakeholders shared examples of incidents that had occurred in Australian schools involving students' use of pornography. The incidents demonstrated that:
    - Pornographic content is being viewed and shared at primary and secondary levels, as early as Year 1. A pattern had been observed of students viewing pornography from increasingly younger ages over the medium term, with the use of pornography being normalised by the time students reach secondary school.
    - Students are sending sexually explicit images and hyperlinks to sexually explicit content through messaging apps, as well as using personal devices or virtual private networks (VPNs) to bypass content restrictions on school Wi-Fi networks.

Support for schools and teachers

- Stakeholders reported many schools are seeking advice on how to respond to students accessing pornography during school hours.
    - Stakeholders felt that both primary and secondary school teachers should be supported to have conversations with students about pornographic or sexualised content online, noting that teachers may lack awareness around the nature of this content.
- People in this group suggested that continuing professional development (both courses or resources) would be helpful for supporting teachers to have difficult conversations with students and parents, and to better understand the issues relating to online pornography.
- Some saw ongoing professional development as more effective than including new curriculum content in pre-service teacher training as it could be updated more frequently to reflect developments in technology. Stakeholders were complimentary of

eSafety's existing resources to support parents and carers to have conversations with children, and suggested that eSafety could be involved in developing these resources for educators.

Broader educational measures

The stakeholders in this group:

- discussed developments to curriculum and noted there may be opportunities to include information about online pornography in education about digital literacy, respectful relationships and consent and sex
- acknowledged that although students may find it easier to speak with non-school staff (such as external presenters) on the subject matter, educational measures should be embedded in curriculum and schooling. They made the point that this facilitates longer term and sustainable impact as there would be age-appropriate learning outcomes embedded consistently into each schooling year. Stakeholders noted that teachers and school leaders can also provide consistency in content implementation and ongoing support
- said that schools across Australia are at different stages in transitioning from a biology focused approach to sex education to a more contemporary approach. They felt that health teachers are currently the most experienced and best placed to deliver educational content on this subject matter.

Current technological approaches to reducing students' access to pornography in schools

- Stakeholders said technological approaches to preventing students' access to pornography varied across jurisdictions. Measures included:
  - device level filters
  - network level filters
  - pro-active scanning for language of concern and individual incident alerts on the school Wi-Fi network.
- People in this group identified challenges, such as students using a mobile hotspot or personal devices (not subject to school controls) to avoid the content filters and restrictions.
- It was noted that content filters and restrictions could either be applied to the whole school population or by year group to align with age-appropriateness.

- Stakeholders also flagged the role of technology usage agreements, codes of conduct and policies for personal mobile phone use at school as measures which also address access to pornography.

Understanding harmful behaviours and providing support to young people: Australian stakeholder group

Consultation group overview

Stakeholders: Law enforcement intelligence, researchers and clinical therapists for young people who have been court sanctioned for sexual offences.

Consultation dates: 8 November 2021 and 17 March 2022.

Overview: This stakeholder group discussed age verification and young people's access and use of pornography, how young people access information, and gaps and opportunities in education.

Young people's access to and use of pornography

Comments from the first stakeholder

- The stakeholder's evidence suggested a potential correlation between increased access to online pornography and increased sexual behaviours among 13 to 15 year-olds. The stakeholder's data also suggested increasing numbers of teenage girls presenting for medical treatment for injuries sustained through sexual activity that may be influenced by online pornography.
- The stakeholder acknowledged the influence of pornography on child development is still relatively uncharted territory and there is likely significant variety in the amount and type of content watched by young people.

Comments from the second stakeholder

- This stakeholder noted that the rate of young people accessing pornography is quite high, however they said that it is difficult to determine any common effect on young people committing sexual offences.
- They expressed that identifying a causal link between pornography consumption and harmful sexual behaviours is sensitive and not straightforward as there is substantial complexity in the way harmful sexual behaviours present among young people.
- The stakeholder explained that they discuss pornography consumption as a potential factor in a young person's behaviour on a case-by-case basis with their clients. It has been identified as a factor in some, but not all, cases.
- They observed that it can be difficult to disentangle any potential impacts of a young person's pornography consumption from other factors that may have contributed to their

harmful behaviours, such as adverse childhood experiences including maltreatment, dysfunctional families, sexual victimisation or a lack of protective factors. They also emphasised that factors such as experiencing domestic violence may attract young people to watching or re-enacting the coercive scripts found in some pornography.

- The stakeholder expressed concerns that restrictions on mainstream content may unintentionally encourage young people to access fringe content, which may accelerate interests in extreme, violent or abhorrent content by young people who are already vulnerable.

How young people access information

Comments from the first stakeholder

- The participant noted that law enforcement can provide support by sharing relevant data and intelligence to inform understanding about associated risks and harms but may not be best placed to deliver prevention messaging in relation to children's access to adult pornography.

Comments from the second stakeholder

- This stakeholder noted that, unlike other frontline workers, trained clinicians are skilled in navigating conversations about sex and pornography with young people. They made the point that clinicians also have specific resources available for clients who have preoccupations with, or compulsions towards, unhealthy sexual behaviours.
- They advised that young people should have a safe place to find information and ask questions. In their experience, young men have many questions about sex, their own sexuality and whether they are 'normal'.
- The group noted that young people will seek out information online (such as through YouTube), however it is hard to establish the quality of the information available to them. They stated that:
  - it is difficult for young people to find quality, scientifically backed resources unless they are specifically directed to them
  - it is important to enable young people to ask questions and seek information online and anonymously to reduce associated stigma.

Gaps and opportunities in education

Comments from the second stakeholder

- The stakeholder believed that children and young people who are not engaged in school are likely to miss out, even though school education about sex and consent has improved. Similarly, they highlighted that vulnerable young people, such as those in outof-home care or those living with disability, may not have a clear path or responsible trusted adult to discuss their questions and concerns regarding healthy sexual behaviours and pornography.
- They acknowledged the importance of timely education, suggesting that late primary school (ages 8 to 10) is a suitable time to start having conversations with children about topics like consent, in a way that is appropriate to their age.
- The group also said that many parents and carers wait for children to start conversations about topics like online pornography, which they felt can often be too late as it suggests that children may have already accessed pornography.

Business and consumers: Australian stakeholder group

Consultation group overview

Stakeholders: Australian financial service providers.

Consultation dates: 28 and 31 March 2022.

Overview: These stakeholders discussed age verification in other contexts, business and consumer needs and expectations, and complementary measures for age verification and assurance technologies.

Age verification in other contexts

- Both stakeholders provided overviews of how proof of age can be verified in commercial contexts that do not relate to online pornography, for example, online banking or alcohol sales.
- They explained how points of purchase (merchants) are connected to digital verification services (verification services) via a third-party exchange (exchange).

Comments from the first stakeholder

- Whereas some commercial contexts require identity to be verified, the stakeholder noted that a provider of adult content who wants to prevent access by underage users only needs limited information, confirming that a prospective user is a real person and 18 years or older.
- The stakeholder highlighted that there are:
    - methods of limiting the amount of information shared through an exchange between parties o methods for only sharing individual attributes determined by individual users
    - options for individuals to have information stored and reused on a verification service or used in a one-off information exchange and then erased
    - various international standards which can be used to guide the development and maintenance of privacy, security and safety standards of digital verification services.

Comments from the second stakeholder

- The stakeholder said that digital identity solutions should minimise data sharing and use zero-knowledge proofs. This means the merchant can only see confirmation that a user is a real person and is over 18, but does not see the user's name, date of birth or other identifying information.

Business and consumer needs and expectations

Comments from the first stakeholder

- The stakeholder explained that exchanges can be jurisdiction and technology agnostic, which could support its use by both international and domestic businesses, as well as adult sites operating across borders.
- The stakeholder noted that currently identity verification is a high friction, high frustration and high cost experience for most businesses and consumers. They expect however that verifying identities digitally will become the easier path, as digital IDs become more mainstream and more options are available to businesses and consumers.
- They advised that businesses want to create more seamless customer experiences and lower costs while also ensuring they can confidently comply with privacy, safety and security requirements.
- Similarly, they noted consumer perspectives on identity verification and data-sharing concerns. They stated consumers are often more comfortable providing information to trusted providers for transactional purposes (for example, purchasing online from a familiar brand).

Comments from the second stakeholder

- The stakeholder stated that payment services have limited visibility as to the nature of purchases, however they can intervene in instances of illegal conduct (such as the sale of counterfeit goods).
- The stakeholder advised that they have internal mechanisms and rules about illegal content, however it was more complex to apply to situations where the product is legal, but age restricted. The stakeholder believed that responsibility for enforcement of age restricted goods (such as gambling or pornography) should generally lie with the 'retailer' or bank, rather than payment services.
- The stakeholder highlighted that legitimate and legally compliant businesses can face impediments to accessing banking services which can often increase risks for already vulnerable people.

Complementary measures for age verification and assurance technologies

Comments from the first stakeholder

- The stakeholder emphasised the importance of public awareness and education around verification and assurance technologies, standards and requirements for individual privacy, safety and information security.
    - They said that services have a role to play in highlighting the relevant features of their verification tools, however government could also play a role in improving public awareness of the design and appropriate use of these technologies – particularly as they become more common in commercial settings such as banking, gambling and age-restricted goods.

Digital platforms and services: Australian and international stakeholder group

Consultation group overview

Stakeholders: Digital platform and service providers.

Consultation date: April to June 2022.

Overview: eSafety held multiple individual and roundtable discussions with digital platform and service providers. Stakeholders included social media platform providers, device manufacturers, search engine providers, app store providers, industry groups and other digital service providers.

The following summarises the comments made during individual consultation meetings and a roundtable session.

Current measures

- Stakeholders operating digital platforms each used different measures to protect children and young people from age-inappropriate or harmful content. The measures across industry varied as the platforms and services had different rules or terms about pornographic content.
- Platforms and service providers that **do not allow** pornographic content under their terms of service discussed how they employ measures such as:
  - Machine learning and human-based content moderation, and filter systems to remove pornographic content.[8]
- Platforms and service providers that **do allow** pornographic content discussed how they employ measures such as:
  - Machine learning and human-based content moderation, and filter systems to identify content which is not permissible under their terms of service (such as violent or non-consensual content).
  - External and internally operated age verification measures for users who upload content.

[8] Machine learning is a type of Artificial Intelligence (AI) which uses data and algorithms to make decisions without human input while also 'learning' to gradually improve its accuracy.

  - Age gating either the platform, or certain areas of the platform; allowing users to report suspected underage users; and verifying user age through identity documents in appeals processes.
  - Measures which limit the reach of pornographic content, or rules that prohibit pornographic content from being visible in areas of the platform accessible to users who are not signed in, or users with an under-18 profile.
  - Measures enabling users to identify content as sensitive or appropriate for those who were 18 years or older, which require additional consent or input from users to view such material (for example, blurring an image until a user clicks to confirm they wish to view the post).
- Some platforms, regardless of rules about adult content, employed other additional safety measures, such as:
  - Behavioural analytics or machine learning tools to identify underage users; imposing more privacy-preserving default settings; and limiting content and contact for younger users.
  - Default settings for younger users who have declared their age, including limiting public and private communication with other users, and limiting the reach of user-generated content.
  - Providing information to parents about the tools available to curate their child's experience on the platform.

Proposed measures as part of the roadmap

- Stakeholders acknowledged that platforms and service providers have a responsibility to keep all users safe, especially young users. They agreed that this includes measures which prevent the distribution of age-restricted or harmful content to children and young people.
- The stakeholders held a variety of views on how this objective was best supported. Across the sessions, stakeholders discussed various measures that could support the roadmap's objectives. These included:
  - The need for multifaceted and holistic measures to address children and young people's access to content, including less onerous age assurance for lower risk settings and measures to support education about pornography in addition to technical solutions.
  - The implementation of age assurance measures further up the digital stack (for example, at the device level) as a way of addressing efficacy, improving child

protection and minimising potential privacy risks. They felt this would require less collection and duplication of data.

- Encouraging use of device and app store parental controls, which already exist but may have low familiarity and uptake. They suggested the roadmap could include recommending increased government or eSafety resources in this area. The stakeholders also noted an opportunity for industry and government to think collectively about how to get the right information to parents at the right time.

Types of measures

Stakeholders broadly supported the notion of preventing children from accessing pornographic or other age-inappropriate content but noted that it could be approached in many ways. Some stakeholders felt that age assurance and age verification could assist in preventing access to age-restricted or harmful content, but that a multi-layered and holistic approach was required. Many stakeholders held differing views about the efficacy and benefits of the different safety measures that are available.

Moderation

- One stakeholder noted that due to the scale and volume of content on their platform, they use machine learning and human moderation. The stakeholder noted that while machine learning models are essential to respond to the scale of content, human moderation is important to apply context.
- The stakeholder noted that it is easier for machine learning based moderation techniques to pick up content which contains nudity as it is less reliant on textual/audio context.
- Another stakeholder reflected on their own age verification practices. They identified sensitivities, also noting the importance of human intervention in age verification and in content moderation, to reflect the nuance of online communication. The stakeholder advised that it is important to have human intervention in verification processes, especially for certain groups. For example, they pointed out that there are members of the trans community, where their lived identity does not necessarily match their identity documents.

Parental controls

- Stakeholders discussed the variety of parental controls and measures currently available, including those where devices for children can be limited to allow only certain apps, or to limit web browsing functionality.
- One stakeholder said parents should be empowered to make their own choices in operationalising tools and settings that curate experiences for their child, based on their own values and their child’s maturity levels.

- Another stakeholder noted that many of the current safeguards rely on parents actively choosing to implement parental safety controls.
- Several stakeholders said reliance on parental controls and device filters should be considered in the context of vulnerable and disadvantaged young people.

Age assurance and age verification

- One stakeholder suggested that age assurance is a good compromise for some circumstances. The stakeholder felt that while verification is an extra burden, it is effective as a safety measure.
- Another stakeholder advised that many in the industry view age assurance technology as new and unreliable and are uncomfortable with its use until the technological solutions are more robust and there is clearer guidance on the privacy implications. They suggested its use, potential and applicability as a solution in the roadmap should be further studied.
- A comment was also made about digital tokens as a form of age assurance or age verification, noting that they are a relatively lower burden compared to other age assurance measures. The stakeholder noted that tokens are based on simple technology but would be best supported by wide uptake and non-proprietary options.

Impacts of age assurance and age verification

- One stakeholder stated that the cost of age assurance or verification measures is not prohibitive, but is a valid cost of doing business, akin to insurance. The stakeholder noted it will impact on industry revenue, but the ability to make revenue should not be unfettered at the expense of safety.
- The stakeholder, drawing upon their experience in jurisdictions where age verification has been mandated, noted that the increased friction did result in fewer users.

Age assurance and age verification models

- Several stakeholders suggested that responsibility for verifying the age of users should be raised as high as possible in the digital stack (that is, at the operating system or device level). Stakeholders noted it is harder to maintain consistent information about age when relying on individual platforms and services to verify information. Similarly, they felt it would remove duplication of effort for both end users and companies.
    - A stakeholder believed this would also remove the need for parents and carers or users to configure individual settings for every website, platform or service and instead allow them to opt in or out only once.
    - Another stakeholder offered an example of devices verifying the age of its users through information already held for contractual or security purposes. They noted

  devices could then share the age attribute (such as '18+') with apps or platforms downloaded to the device. In that case, the platform would not collect or store any additional identifying information about the person.

  - A stakeholder noted that individual platforms can subsequently design age-appropriate experiences relevant to an individual user's age and any jurisdictional requirements. For example, some content or functions may only be accessible based on users being 18 years and older (such as sexually explicit content). Alternatively, bespoke experiences could also be shaped for different age groups (for example, ages 10 to 11, 12 to 13, 14 to 15).
  - A stakeholder acknowledged that this approach could consolidate market dominance by manufacturers and emphasised that interventions should balance safety, privacy and competition.

- Another stakeholder did not believe device security measures (such as existing PINs or biometrics used to unlock devices) were appropriate tools to use for the purposes of the roadmap.

Trust in age assurance and age verification

- Many stakeholders raised public trust in age assurance and age verification measures as a concern.
- One stakeholder felt age verification and age assurance may have an important role to play but noted that many people have concerns with the digital industry's use of data and with the potential unreliability of age estimation. The stakeholder noted that public trust is key.
- Another highlighted that implementing any form of verification requires users to trust the platform and to feel comfortable sharing personal information.
- Another stakeholder noted that transparency is key. They proposed user awareness of processes, and platforms not asking for more information than needed, or keeping that information longer than necessary, were fundamental elements to good age verification practice.

Age verification regime design

- One stakeholder suggested a legislative obligation was necessary in order to create an online environment where age verification is the norm. They felt without such an obligation creating consistency across platforms, users would have broken or inconsistent experiences. They also acknowledged the business cost in developing age verification or assurance systems and the challenge of compliance across different jurisdictions.

- Another stakeholder acknowledged that all parties in the tech industry are facing the challenge of moderating content. They advised that while their approach of having all content scanned by technology and subject to human moderation was not easy, it was worthwhile. The stakeholder considered that the onus should be on platforms to do more – and they did not need to wait for legislation to force them to do more.
- Several stakeholders advised that the roadmap and any subsequent implementation should account for the different business models and types of platforms. The stakeholders also considered that the nature and age of the user base was a relevant consideration in determining proportionate measures.
- One stakeholder in this group was concerned that age assurance or age verification technologies were being presented as quick and definitive solutions to a complex problem and stressed the need for a holistic response that does not rely on a single solution.

Reaching parents and carers

- Stakeholders discussed how they communicate safety information to users and parents and carers of younger users.
- A stakeholder reported that it can be challenging to engage with parents and carers about safety tools. The platform provider explained that they partner with nongovernment organisations (NGOs) and specialist organisations for their engagement and outreach to parents and carers.
- Another explained that they are exploring partnerships with influential creators on their platforms, to encourage peer-to-peer learning about safety tools and strategies.
- One stakeholder said that industry practices around parental controls and educating parents about them are inconsistent. The stakeholder discussed parent-led learning as an opportunity to share safety information. This would mean taking safety information to places where parents communicate (such as parenting groups on social media) or to parent organisations to distribute, rather than requiring the parent to reach out to the platform.
- A stakeholder flagged the challenge of reaching parents with safety information and suggested such efforts should be part of the roadmap. They also expressed that disadvantaged and marginalised children and young people are more at risk of online harms and should be a particular focus in the roadmap. For example, the children who may not be receiving sex education in school and who may not have parents that are willing to discuss sex education matters at home.
- The point of sale for devices was also seen as a potential safety check point where relevant safety settings could be explained and applied. However, stakeholders noted

that many young people receive second hand devices from adult family members, meaning point of sale will not always be the ideal time to engage parents and carers, and children.

Accessibility and fairness

- Stakeholders discussed issues of capacity, equality and online participation as important considerations when assessing types of interventions. They acknowledged that user age verification is very complex and can have significant drawbacks for accessibility and fairness.
- Participants generally felt that verification (such as using official identity documents) was problematic in that it is not fool proof, as documents can be shared and access to documentation is not universal – which potentially affects individual rights to digital access and participation; and the collection of data raises security and privacy concerns.
- One stakeholder highlighted a tension between protecting children's safety and wellbeing and protecting the privacy of all Australians online. They noted that a broader societal conversation should occur to identify the public's views on how to balance each objective.

Efficacy and risk

- Stakeholders discussed the risks associated with ineffective age verification measures. Stakeholders believed it was inevitable for some young people to misrepresent their age or attempt to navigate around restrictions. They emphasised the importance of having multiple intervention layers to make evasion more challenging.
- One stakeholder highlighted that, as mainstream platforms and services do more to protect users, the interventions applied can push users to seek content from alternative platforms, browsers and websites which may be less safe, secure and compliant with applicable laws.
- Another noted that age restrictions can incentivise young people to misrepresent their age. To counter this, the stakeholder discussed ways to create unique and age-appropriate experiences for young people, such as creating teen-only spaces for young people to safely connect with peers over mutual interests.
- Stakeholders were concerned about a lack of harmonisation across international approaches, noting that an Australia-only approach could be ineffective, as mechanisms such as virtual private networks (VPNs) allow users to bypass jurisdictional measures.
- Stakeholders felt the international harmonisation of measures was important for efficacy and cost.

- A stakeholder also noted that treating all users as under 18 (unless proven otherwise) carries risks. For example, this can prevent the use of systems which limit the ability of unknown adults to contact minors.

Future technology and digital identity

- Some stakeholders discussed opportunities for age assurance and age verification measures that could be implemented through decentralised systems and settings where users would have control of their data and could choose what personal attributes were shared.
- One stakeholder said that open standards to facilitate a digital identity or attribute exchange was a viable option, but felt it was inappropriate for a single corporate entity to design or own related technology.
- Another noted that many identity checks involve needless disclosure of information. For example, providing a driver licence to purchase alcohol requires you to provide your age (necessary) as well as your full name and address (unnecessary). The stakeholder described a model where trusted authoritative sources (both government and community-based) provide attestation for an attribute. This attestation could be issued to a cryptographic key or token on your phone and the user could share and withdraw the credential as needed.
- The stakeholder advised such a model needed to be open-source to allow users portability, choice and privacy, noting that if users cannot change the provider of the token, they could be more susceptible to censorship.

Appendix 6: Consultation participants

Organisations that participated in multi-sector consultation process
18North
5Rights
Australian Curriculum, Assessment and Reporting Authority (ACARA)
ACT Education Directorate
Association of Heads of Independent Schools of Australia (AHISA)
Alannah and Madeline Foundation
Apple
Age Verification Providers Association (AVPA)
AgeChecked
Bumble
Collective Shout
Communications Alliance
Crisp Thinking
Digital Industry Group Inc. (DIGI)
Digital Rights Watch
Discord
eftpos
Electronic Frontiers Australia
Equifax
Eros Association
euCONSENT
Family Zone
Google
Griffith Youth Forensic Service
Interactive Games and Entertainment Association (IGEA)

IIS Partners
It's Time We Talked
Jumio
Mastercard
Meta
Microsoft
MindGeek
OnlyFans
Optus
Our Watch
Privately
Queensland Police Service
Queensland Department of Education
Reddit
The Reward Foundation
Roblox
Safecast
Scarlet Alliance
Snap Inc
Spectrum Ai
Teach Us Consent
TikTok
TrustElevate
Twitter
VerifiiD
VerifyMyAge
Victoria Department of Education
xHamster
YOTI

Academics associated with the following institutions: • Aston University • Burnet Institute • London School of Economics and Political Science • Middlesex University • Queensland University of Technology • Technological University Dublin • University of New South Wales • University of Sydney • Western Sydney University
Federal Government departments and agencies consulted
Attorney-General’s Department
Australian Cyber Security Centre
Australian Human Rights Commission
Australian Competition and Consumer Commission
Department of Home Affairs
Department of Infrastructure, Transport, Regional Development, Communication and the Arts
Department of Education
Department of Social Services
Digital Transformation Agency
Office of the Australian Information Commissioner
Services Australia
eSafety groups consulted
National Online Safety Education Council
Online Safety Youth Advisory Council
Trusted eSafety Providers
Other
Members of the International Working Group on Age Verification
Australian and New Zealand Children’s Commissioners and Guardians (ANZCCG)

Appendix 7: Cross-sector workshop summary

Workshop overview

eSafety held an independently facilitated cross-sector workshop with a wide range of stakeholders on 19 July 2022 to inform the development of an age verification roadmap.

This workshop followed on from consultation sessions held between November 2021 and July 2022, where several stakeholders told eSafety that it would be a useful and constructive experience to bring together all of the different sectors affected by this work. They said this would give stakeholders the chance to share their perspectives with eSafety and with each other.

eSafety invited a selection of stakeholders from academia, the adult industry[9], children's wellbeing advocacy groups, digital rights and privacy experts, education authorities, safety technology providers and digital services and platforms.

The cross-sector workshop provided participants with an opportunity for multi-disciplinary discussion, as well as an opportunity for eSafety to further understand the issues raised during previous consultation sessions.

This document provides a high-level, anonymised summary of the cross-sector workshop. Summaries of the prior individual meetings and sector-specific roundtables are available on our website.

The views and opinions in this summary are those of the stakeholders and do not necessarily reflect eSafety's ultimate position. However, they are an important contribution to informing the development of the age verification roadmap.

More information about the roadmap is available on our website.

[9] Adult industry covers commercial enterprises (individuals, businesses or peak bodies) involved in the sale or purchase of sex-related entertainment services.

Workshop objectives

These were the objectives of the workshop:

- Bring stakeholders together to ensure previously expressed views have been properly captured and to help us all build an understanding of the full range of competing factors that must be balanced.
- Define principles that can help shape the direction of the report and its recommendations, balancing areas of consensus and disagreement.
- Identify and consider proportionate and feasible measures for reducing the risks and harms of underage access to online pornography.
- Deliberate the roles and responsibilities of government, industry, family and others.

eSafety thanks participants for their time and thoughtful contributions to the workshop.

Discussion of draft principles

eSafety's consultations involved wide-ranging discussions and broad, diverse views.

Based on these discussions, eSafety drafted a series of six high-level principles to distil the points of consensus, to serve as guideposts in drafting the roadmap. eSafety asked workshop participants to provide feedback on the draft principles.

The following table includes the draft principles as presented at the workshop, and a snapshot of emerging themes from both the workshop and prior consultations.

Principles the age verification roadmap should follow

Draft principle 1 - Take a proportionate approach based on risk and harm

Understanding the nature of the risks and harms – and the areas where there is greater or lesser evidence and agreement – will enable measures which are reasonable and targeted.

What we've heard from other consultation meetings	This is what stakeholders have told us previously: • There is general agreement that younger children may experience harm from viewing online pornography, especially if they haven't had any previous education or support on the matter. Some also experience harm from negative reactions of trusted adults. • There is lesser agreement, but a growing evidence base, about potential harm to older children, especially where pornography use is frequent. • Not all uses of pornography are harmful, and notions of 'harm' should be examined for bias. • The risk of harm depends on the context, and there is no single intervention that is applicable to or appropriate for all contexts. All interventions carry their own risks, consequences and trade-offs which must be identified and considered.
Workshop feedback	This is the feedback from participants at the workshop: • The research informing the roadmap should be drawn from a diverse range of interdisciplinary sources, including the field of pornography studies. • Not all access to pornography is harmful, and reducing access is not necessarily the best way to reduce harm in all cases. • The roadmap should draw a distinction between younger children and older teens, who will experience different levels of risk and harm. • The roadmap should acknowledge that there are cultural differences in understanding harm and pornography.

Draft principle 2 - Respect and promote human rights

Making the online world a safer space is ultimately about fulfilling the human rights of those who inhabit it.

What we've heard from other consultation meetings	This is what stakeholders have told us previously: • The best interests of the child should be the paramount consideration, informed by children's rights to protection, health, participation, expression, information, education, privacy and non-discrimination. • Children have agency, resilience and evolving capacities and should be equipped with skills to navigate the online environment and interpret the content they encounter. • eSafety should consult with children of different ages as the roadmap is being developed (and rolled out) to define what is in their best interest. • The roadmap should acknowledge children's agency, and that children are not merely passive consumers of content. • The rights of all other affected persons must also be taken into account, including the rights of adults who lawfully make, share and/or consume online pornography.

	• It is particularly important to consider the rights of people who may be most at risk online, including children and adults: ○ with disability ○ who are First Nations people ○ from culturally and linguistically diverse backgrounds ○ who are members of the LGBTIQ+ community ○ who identify as girls or women.
Workshop feedback	This is the feedback from participants at the workshop: • The roadmap should reference and reflect UN General Comment 25 on children's rights in relation to the digital environment.[10] It should also explore the intersections between the rights of children and the rights of adults in a productive way, acknowledging there are some areas of convergence and some areas of tension. • Teens' agency and their own rights to participation and expression should be considered. • Particular consideration should be given to LGBTIQ+ communities, including how LGBTIQ+ young people engage with pornography as a form of sex education and expression. The roadmap should not have the effect of excluding LGBTIQ+ and other young people who may not have support from their parents. • The rights of sex workers (including pornography performers and producers) should be acknowledged. • Some age verification tools (including those which rely on ID) are not accessible to all, which potentially affects individual rights to digital access and participation.
Draft principle 3 - Propose a holistic response, recognising that everyone has a role to play There is no 'silver bullet' technology and a whole-of-community approach is required.	
What we've heard from other consultation meetings	This is what stakeholders have told us previously: • There is a role for government, the online industry, the adult industry, safety tech providers, educators and other frontline workers, parents and carers, other trusted adults – and children and young people themselves – to reduce the risks and harms associated with online pornography. • The nature of their role and responsibility may shift depending on the context, for example, parents and carers may have a more prominent role for younger children. • Technology alone – and particularly age assurance or verification alone – will not resolve these issues. A level of failure or error is inevitable, so it is crucial to provide a fast, empathetic and caring pathway to redress. • Tailored, inclusive and age-appropriate education about sex, relationships, consent and online pornography is key for both children and the adults in their lives.
Workshop feedback	This is the feedback from participants at the workshop:

[10] UN, *Convention on the Rights of the Child: General Comment 25 on children's rights in relation to the digital environment*, available at: https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021childrens-rights-relation.

	• A holistic response requires consideration of different stakeholder roles, but also consideration of the different ways and reasons why children and young people access online pornography. • Different types of technical interventions may be appropriate for different situations – for example, inadvertent access by young children compared to deliberate access by older teenagers. • The general public should be engaged in developing the roadmap, so they can feel that policy changes are done with them, not to them. • Educators have an important role to play, but will need support with guidance and training. More research may be needed about the effectiveness of different types of educational interventions. • The role and interests of parents and carers in protecting and supporting their children should be more clearly reflected. • Parents and carers should be engaged in a meaningful way rather than merely being asked to tick a box to give consent for their child to engage online. • Empowering parents to protect and support their children should be a principle underpinning the roadmap, consistent with a child rights approach. Currently, many parents feel excluded from, or powerless in relation to, decisions that affect their children and families. • There should be a principle that explicitly recognises the need to protect access to sex education and ensure resources and support are relevant for LGBTIQ+ young people. If sex education is not accessible and inclusive, young people will continue to turn to pornography to learn about sex. • A whole of community approach should include sex workers who produce content about consent and sex education for the community. • It may be too broad to say that 'everyone' has a role to play. There may be risks with certain groups playing a role, such as concerns about the role of police in enforcement or conflicts of interest by safety technology providers or digital platforms. Government may not be best placed to determine what constitutes sex-positive pornography.
Draft principle 4 - Ensure any technical measures are data minimising and privacy preserving Safety measures will not work unless they are private, secure and trustworthy.	
What we've heard from other consultation meetings	This is what stakeholders have told us previously: • Interventions to improve safety must also respect privacy and security, and promote trust, which is critical for the adoption and success of age assurance or verification technologies. • International standards for age verification are being developed to protect privacy and minimise data collection. These standards should be upheld domestically, with some stakeholders expressing that they should be incorporated into Australian law. Standards should be iterated and improved over time, and technologies should be continuously monitored and audited for compliance and the potential risk of reidentifying anonymised data. • Children (and adults) need to be aware of the data and privacy implications of using assurance technologies, and should be equipped to make informed choices and have access to remedies for breaches. • There are pros and cons to:

	○ placing technical interventions up and down the digital stack[11] ○ introducing in-house versus third-party solutions ○ implementing age estimation (which is often data minimising but may rely on biometrics which are considered sensitive) compared to age verification (which often relies on government identification and carries its own sensitivities and issues of equity).
Workshop feedback	This is the feedback from participants at the workshop: • Trust and privacy are critical to young people's help seeking. • The roadmap should consider both the actual risk of data breaches – particularly where data is centrally held by a government or corporation – as well as public perceptions in relation to those risks, which may or may not be aligned. • Age assurance may prompt users, including children, towards less secure parts of the internet or more extreme pornography sites to avoid real or perceived surveillance. • The uptake of certain measures may normalise surveillance and encroach on personal privacy. Safety and security can be promoted by collecting, using, storing and sharing less data. • There should be mandatory standards relating to data minimisation to promote trust and give people confidence that their data is not at risk if they use age assurance technology. • Data privacy and online anonymity for sex workers are crucial to their safety and access to income. • The roadmap should consider device- and operating system-level age verification as users tend to already trust these providers with their information. • Age assurance technology providers should be public about their systems and processes, and should demonstrate adherence to information privacy and security standards. There should be consequences for data breaches or other misuse of mishandling of data.
Draft principle 5 - Consider the broader domestic and international regulatory context Potential responses cannot be considered in isolation.	
What we've heard from other consultation meetings	This is what stakeholders have told us previously: • Both the domestic and international regulatory landscapes are highly dynamic, and influenced by a range of social, cultural and political factors. • It is important to take a consistent and coordinated approach to age assurance requirements which are either in place or being considered across state and federal governments in Australia. • It is also important to collaborate globally and promote harmonisation and interoperability across national standards, while also recognising local contexts. • Consideration should be given to the potential for regulatory burden, particularly for sole traders and small businesses.
Workshop feedback	This is the feedback from participants at the workshop: • Australia should develop a regime that matches its national policy priorities, values and ideals, however, international regulatory cooperation and dialogue is important to promote coherence globally.

[11] Digital stack refers to layers of technology – for example, accessing a website involves a device and its operating system, an internet service provider, a browser, hosting providers, as well as the site itself.

	• The impact and conflict of obligations under international laws should be considered. For example, some age verification methods may put individual pornography producers and performers in violation of the European General Data Protection Regulation. • The roadmap should differentiate between dedicated adult industry sites compared to sites that provide more broad-based content and experiences (such as social media platforms). • The roadmap could look to examples such as the UK Age appropriate Design Code (Children's Code)[12] which mandates privacy protections and child-friendly terms and conditions. • Adoption of international standards for age verification allows for both a convergence of approach and flexibility of application. • There was a difference of opinion about the utility of this principle; one participant felt it made sense given the global nature of the internet while another felt that it did not provide meaningful guidance and may not be helpful to the roadmap.
Draft principle 6 - Consider what is feasible now and into the future Measures should not be 'set and forget'.	
What we've heard from other consultation meetings	This is what stakeholders have told us previously: • There is a range of technical options currently available and in use to reduce children's access to online pornography and/or to make services age appropriate for children more broadly. These are associated with different levels of assurance and accuracy, as well as concern regarding issues such as bias and equity. • The development, uptake and improvement of age assurance and other safety tech measures is rapidly evolving, with a number of companies announcing new measures since consultation began. • As interest in the move towards a decentralised Web 3.0 grows, more initiatives around self-sovereign identity are being proposed.[13] • The roadmap should be future focused and consider this changing online landscape as well as the need for standards which don't prescribe specific technologies (technology neutral standards). • Democratic checks and balances must be in place to limit the potential expansion of verification systems and the passing off risks without accountability.
Workshop feedback	This is the feedback from participants at the workshop: • Policy should not rely solely on the potential of future technology; it should also consider the feasibility of current technology. • Some of the technology in the age assurance space is potentially 'hyped-up'. An ecosystem-wide approach is more likely to be effective. • The roadmap should explain both how measures will be implemented and how the impacts of measures will be evaluated to inform future reform. • There should be support for the adult industry in Australia to adapt to changes, as it is predominately made up of small businesses.

[12] UK Information Commissioner, *Age appropriate design* code, available at: https://ico.org.uk/for-organisations/guide-to-data-protection/ico-codes-of-practice/age-appropriate-design-code/

[13] Self-sovereign identity refers to models of digital identity in which individuals have sole ownership and control of their digital identity credentials, separate from third-party organisations.

Discussion of case studies

Having discussed the principles at a conceptual level during the workshop, the next step was to discuss how they might apply in practice.

eSafety assigned participants into three multi-sector groups and asked each group to consider a different case study. The three case studies reflected the evidence received to date and involved children of various ages, who were accessing online pornography through different means and for different purposes.

The objective of the exercise was to consider the practical implementation and use of different safety measures, including age verification or assurance options, as well as broader technical and non-technical interventions. eSafety was interested to hear views on how these measures could be applied to address risks and harms associated with online pornography in an effective and proportionate manner, and in different contexts, as the age and capacity of children develop.

Participants were asked to leverage their specific expertise and consider:

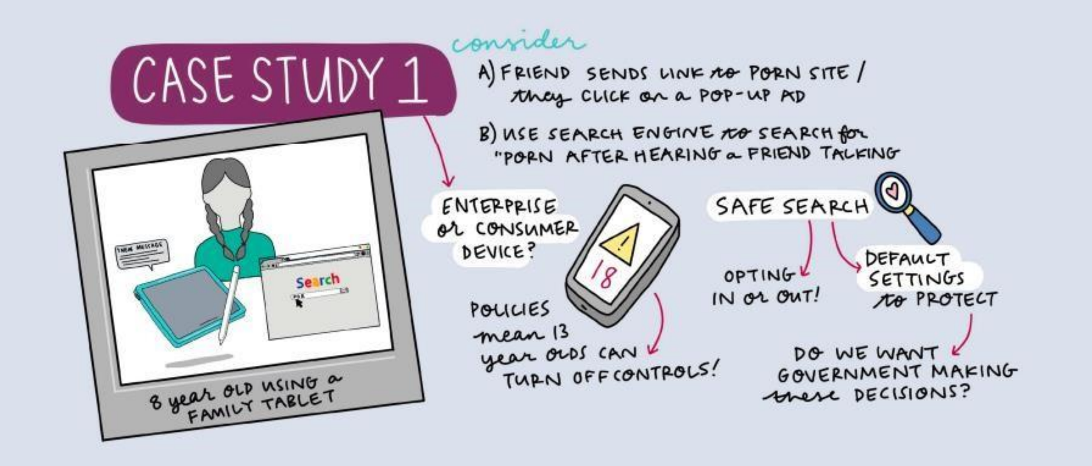
- types of measures, including tech-based measures and non-tech measures (such as education initiatives) that could be implemented
- what role families, governments and industry have in enabling the measures • what the unintended consequences or risks associated with the measures may be
- how feasible the measures are in the context of the case studies.

Each case study contained various scenarios to prompt further discussion. Participants discussed a variety of potential measures, including device-level age verification, educational initiatives and other safety tools, such as content moderation and user tools.

Workshop facilitators provided live graphic scribing of the discussion of the case studies. Extracts of these images have been included with each summary and draw from the conversation with all participants after the small group discussion.

The inclusion of potential measures for discussion in this activity does not reflect an endorsement by eSafety.

Case study 1: 8-year-old who uses a shared family tablet


CASE STUDY 1
consider
A) FRIEND SENDS LINK to PORN SITE / they CLICK on a POP-UP AD
B) USE SEARCH ENGINE to SEARCH for "PORN AFTER HEARING a FRIEND TALKING
ENTERPRISE or CONSUMER DEVICE?
18
POLICIES mean 13 year OLDS CAN TURN OFF CONTROLS!
SAFE SEARCH
OPTING IN or OUT!
DEFAULT SETTINGS to PROTECT
DO WE WANT GOVERNMENT MAKING these DECISIONS?
NEW MESSAGE
Search
8 year OLD USING a FAMILY TABLET

Participants received variations on the scenarios as discussion prompts:

- Friend messages the child a link to a porn site.
- The child accesses a pop-up ad for a porn site.
- The child uses a search engine to search for 'porn' after hearing friends talking about it.

The intention of this case study was to enable discussion of:

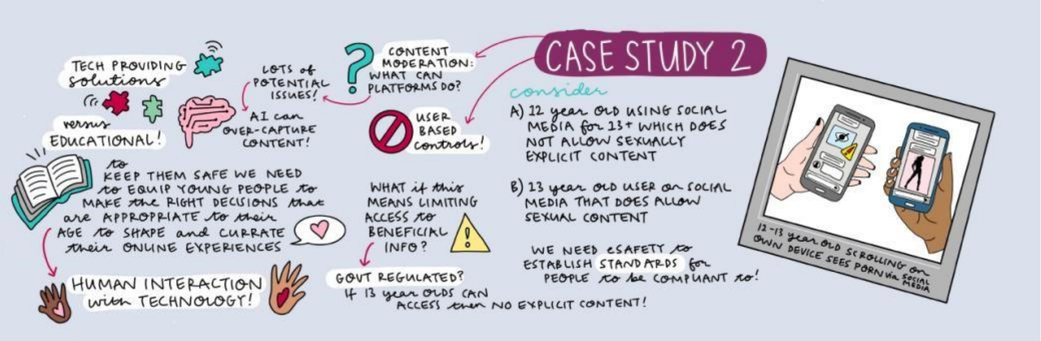
- whether different interventions may be more appropriate for accidental exposure compared to a curious search or click
- whether family devices provide opportunities or challenges for managing children's access to material online.

Participants discussed device-level measures as well as site-level measures. Participants held different views on whether introducing device-level age assurance would be simple or complex to implement.

Participants said that parents, carers and families need to be supported to understand how to use safety tools, noting however that some children lack access to family support. Participants suggested that even if safety settings are on by default, there remains a need for families to have discussions about their expectations for technology use and make decisions about the types of content that may not be suitable for their family.

Participants discussed the process to determine which sites are appropriate for children, and raised considerations about which stakeholder(s) have that responsibility.

Case study 2: 12 or 13-year-old who uses their own mobile device to scroll through their social media newsfeed and sees pornography


TECH PROVIDING solutions
versus EDUCATIONAL!
LOTS of POTENTIAL ISSUES!
AI can OVER-CAPTURE CONTENT!
CONTENT MODERATION: WHAT CAN PLATFORMS DO?
USER BASED controls!
CASE STUDY 2
consider
A) 12 year OLD USING SOCIAL MEDIA for 13+ WHICH DOES NOT ALLOW SEXUALLY EXPLICIT CONTENT
B) 13 year OLD USER on SOCIAL MEDIA THAT DOES ALLOW SEXUAL CONTENT
to KEEP THEM SAFE WE NEED to EQUIP YOUNG PEOPLE to MAKE the RIGHT DECISIONS that are APPROPRIATE to their AGE to SHAPE and CURRATE their ONLINE EXPERIENCES
WHAT if this MEANS LIMITING ACCESS to BENEFICIAL INFO?
GOVT REGULATED? If 13 year OLDS CAN ACCESS then NO EXPLICIT CONTENT!
WE NEED eSAFETY to ESTABLISH STANDARDS for PEOPLE to be COMPLIANT to!
HUMAN INTERACTION with TECHNOLOGY!
12-13 year OLD SCROLLING on OWN DEVICE SEES PORN via SOCIAL MEDIA

Participants received variations on the scenarios as discussion prompts:

- 12-year-old using social media platform for 13+ users which does not allow sexually explicit content.
- 13-year-old using social media platform for 13+ users which does allow sexually explicit content.

The intention of this case study was to enable discussion of:

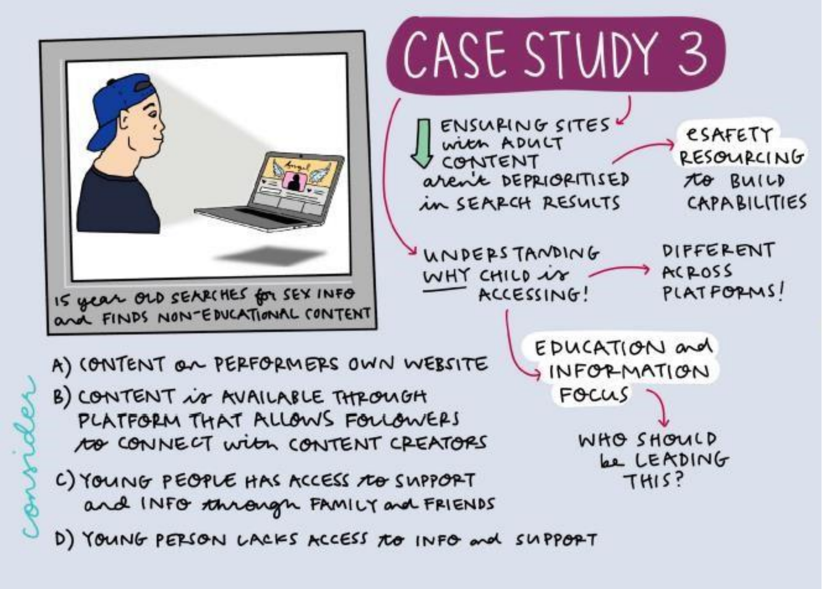
- the age assurance options for enforcing minimum age requirements to participate on various sites
- the extent to which different options may be suitable on sites where sexually explicit material is permitted compared to where it is not allowed
- the role of community rules and content moderation tools of a service
- the evolving role of parents as children grow up and begin to use their own devices and online accounts.

Participants considered the role of parental controls, device-level age verification, user-based controls, content detection and removal tools as well as educational measures in relation to this scenario.

Participants suggested that while content detection, removal tools and parental controls are available currently, and can be effective in blocking access, they also carry risks. Such tools may be used in harmful ways, and they could also block important sexual health resources.

Participants discussed that government could have a role to support a consistent approach and ensure a ‘level playing field’ for all platforms and adult sites when implementing measures.

Case study 3: 15-year-old who uses their own laptop to search for information about having sex and finds explicit content produced by an Australian adult content producer where the primary purpose is not educational


CASE STUDY 3
15 year OLD SEARCHES for SEX INFO and FINDS NON-EDUCATIONAL CONTENT
Angel
ENSURING SITES with ADULT CONTENT aren't DEPRIORITISED in SEARCH RESULTS
eSAFETY RESOURCING to BUILD CAPABILITIES
UNDERSTANDING WHY CHILD is ACCESSING!
DIFFERENT ACROSS PLATFORMS!
EDUCATION and INFORMATION FOCUS
WHO SHOULD be LEADING THIS?
consider
A) CONTENT on PERFORMERS OWN WEBSITE
B) CONTENT is AVAILABLE THROUGH PLATFORM THAT ALLOWS FOLLOWERS to CONNECT with CONTENT CREATORS
C) YOUNG PEOPLE HAS ACCESS to SUPPORT and INFO through FAMILY and FRIENDS
D) YOUNG PERSON LACKS ACCESS to INFO and SUPPORT

Participants received variations on the scenarios as discussion prompts:

- The content is available on the performer’s own website.
- The content is available through a large platform, which lets followers connect with creators of user-generated, sexually explicit content.
- The young person has access to information and support through family, school, community or services.
- The young person lacks access to information and support.

The intention of this case study was to enable discussion of:

- the extent to which different measures may be appropriate for large platforms compared to sole traders or small businesses
- how a young person’s circumstances and access to support may impact measures.

Participants in this group discussed options to provide young people with educational content about sex, relationships and online pornography. Participants considered different referral pathways, and education sources, including schools and mentors.

Participants explored the challenges in returning age-appropriate search results for children and young people while also avoiding over-capturing data about them. They also said that adult sites should not be deprioritised in searches by adult users.

Participants reviewed the role of not-for-profits, schools and government in providing children and young people with information, while noting that producing the information would require support and funding.

eSafety will use these case study discussions to guide consideration and assessment of the available measures within the roadmap.

Appendix 8: Independent assessment of age assurance and safety technologies (Enex TestLab report)

Age Verification Technology Evaluation

Prepared for

Australian Government | eSafetyCommissioner

DATE: 04/01/2023

REPORT VERSION: FINAL

Enex Pty Ltd
ABN 77 096 299 099

+61 3 9436 7454
www.testlab.com.au

Office Address: 10/355-365 South Gippsland Highway, Dandenong South VIC 3175
Mailing Address: PO Box 45, Yarrambat VIC 3091

This page has been left intentionally blank

Contents

Executive Summary

A wide variety of technologies are available to contribute towards the task of protecting children from viewing unsuitable online content while ensuring adults can access this material if they choose. Broadly speaking, there are two approaches:

- **Age assurance.** Restricting content at an online service to ensure only people of an appropriate age can access an online service, or parts of that service, using techniques including age verification and biometric age estimation.
- **Online safety.** Filtering internet content, either at a service provider, home network or individual device level using techniques including filtering and safe search, or minimising children's exposure to inappropriate content or behaviour on a website or online service using techniques including age gates, tagging, and moderation.

The fundamental question with these approaches is how can a website or online service provider *know* if anyone using their service is a child or an adult? The Office of the eSafety Commissioner has engaged Enex TestLab to evaluate the variety of current and emerging methods to verify or estimate the age of online users.

This paper is our evaluation of age assurance and online safety technologies available in the market as of November 2022.

We examine existing and emerging standards for age verification available internationally. We outline the various technology available for age assessment and related safety methods. We then evaluate a range of commercially available age assessment and online safety technologies using criteria including:

- Feasibility
- Security
- Technical integrity
- Need to provide sensitive data
- Potential for bias
- Barriers to inclusion
- Cost
- Transparency of decision making
- Flexibility of implementation
- Ability to mitigate unintended consequences.

Finally, we identify positive development, challenges and gaps in the current age assessment technology landscape.

We can draw three clear conclusions from this assessment:

- The age verification industry and associated technologies are relatively new and still evolving
- No single solution is bulletproof in identifying age or protecting children from harmful online content
- Robust standards – nationally or globally – will be necessary to make it easy for service providers to implement age verification technologies consistently and effectively.

Glossary

Term	Definition
Age assurance	The process of verifying or establishing the age of a person, typically for the purpose of age-restricted activities (using techniques such as age verification and biometric age estimation).
Age estimation	Age estimation is the process of determining or predicting the age of a person, typically based on physical or behavioural characteristics or other information.
Age estimation token	A digital token that represents a predicted age of a person and can be used for age-restricted activities.
Age gate	A website feature that requires users to self-verify their age by answering a question, typically 'Are you over 18?', or providing their date of birth before they can access the site's content.
Age token	Services that use age estimation techniques and or hard identifiers (such driver's licences) and generate an electronic token people can use to verify their age online. People can store age tokens in their digital wallets.
Age verification	A technical process that confirms the age of a person using their attributes or other confirmed sources of information. Examples include tokens or licences, third party verification, and government e-ID systems.[1]
Artificial intelligence (AI)	Machine-based artificial intelligence that can be trained to perform tasks requiring human-like intelligence.
AVP	Age verification provider
Biometric age assessment	Services that analyse a photograph of a person's face or other aspects of their person to infer their age.
Cookie	An HTTP cookie is a small piece of data sent from a server to a user's browser. The browser stores the cookie and can later send it back to that same server on request. It is a way to retain or remember stateful information in the stateless HTTP protocol.
Digital identity	A digital identity helps Australians verify their identity in a safe and secure way, to access government and other services online. It removes the need for individuals and businesses to visit a shopfront with their identity documents, saving time and money.[2]
Digital wallet	A virtual storage system that can store age estimation tokens, which represent a predicted age and can be used for age-restricted activities. Digital wallets can be accessed through devices such as smartphones or computers and can also store other types of digital payment methods, such as credit or debit cards.

[1] See eSafety Commissioner, Glossary of terms
[2] See Digital Transformation Agency, Digital Identity

Term	Definition
euCONSENT	euCONSENT is a European Commission–funded project to develop a European Union–wide age verification system.
FFF	Family Friendly Filter
Hard identifiers	Services that verify government-issued identity documents such as drivers' licences, typically using a photo of the document taken on a smartphone or computer.
Internet content filter (ICF)	Products installed on personal devices or network routers or implemented by service providers to prevent access to certain types of content.
ISO	International Organization for Standardization
Moderation	Manually or automatically checking and managing the content of conversations, to ensure that users participate according to the site rules. Some social media services and online chat rooms and forums use moderators to block both individual comments and users who do not participate appropriately. They aim to keep conversations on topic and free from offensive or derogatory comments.[3]
OeSC	Office of the eSafety Commissioner
Online safety	Online safety refers to the measures taken to protect people's security and well-being when using the internet and other online technologies (using techniques such as age gates, safe search, filtering, tagging and moderation).
PAS standard	Publicly Available Specification standard
PASS	Proof of Age Standards Scheme
PII	Personally identifiable information
PUF	Prohibited URL Filter list
Restricted to adults (RTA) metatag	The restricted to adults (RTA) metatag is a HTML tag that can be added to the head section of a web page to indicate that the page's content is intended for adults only and should not be displayed to children.
Safe search	A search engine filter that blocks explicit content
Tagging	A method that relies on content posters or hosts to voluntarily tag or label the age appropriateness of their content.
Trusted Digital Identity Framework (TDIF)	An Australian accreditation framework for digital identity services.
Uniform resource locator (URL)	The address of a resource on the internet

[3] See eSafety Commissioner, Glossary of terms

Phase 1 a): Scan of available age assurance and relevant safety technology in the market

Enex TestLab reviewed commercially available age assurance (age verification and age estimation) and safety technologies (age gates, safe search, filtering, tagging and moderation) in the market as of November 2022. We identified six technology types under two broad categories of age assurance and online safety technology. The age assurance technologies were:

- **Age token** – services that use age estimation techniques and or a known age verified hard identifier (such as a driver's licence) and convert it to an electronic token for use online. People can store age tokens in their digital wallets within their device. Age tokens may also be stored as cookies in a user's internet browser to be reused on multiple websites. Age tokens stored in browsers have challenges as each browser or private browser session will need to reauthenticate the users age to store a new age token.
- **Biometric age assessment** – services that analyse a photograph of a person's face and or voice to infer their age.
- **Digital identity** – services that provide a digital identity to verify a person's name, age and other identifying characteristics, sometimes accompanied by a physical identity document.
- **Hard identifier** – services that verify government-issued identity documents such as drivers' licences, typically using a photo of the document taken on a smartphone or computer.

The online safety technologies were:

- **Artificial intelligence moderation** – services that use machine learning technologies to identify and block inappropriate content and behaviour on websites and online services, often supplemented by human moderators.
- **Internet content filter** – products installed on personal devices, network routers or implemented by service providers to prevent access to certain types of content.

Category	Vendors
Age assurance	
Age token	Privo SuperAwesome
Biometric age assessment	AGEify Ageware FinGo Innovative Technology Privately VerifyMyAge Yoti
Digital identity	Citizen Card MasterCard ID TDIF[4]

[4] TDIF was listed as a government framework accreditation for digital identity services within Australia

Category	Vendors
Hard identifier	Age Check Certification Services Limited AgeChecked AU10TIX BlueCheck Experian Integrity
Online safety	
Artificial intelligence moderation	Spectrum Labs AI
Internet content filter	Asus router Bing safe search Duck Duck Go safe search Google safe search McAfee Safe Family McAfee Secure Home Platform Norton Family

Phase 1 b): Scan of existing standards, including relevant ISO standards.

Existing standards

Enex TestLab examined existing and proposed standards for age verification systems as of November 2022.

PAS 1296:2018 online age checking

The *Publicly Available Specification (PAS) standard 1296*, published in 2018, is a code of practice developed by the British Standards Institute and the Age Verification Group of the Digital Policy Alliance.[5] PAS 1296 is not a registered ISO standard. Its goal is to help businesses comply with the legal requirements of conducting age checks for their online services.

Compliance with this standard is a requirement for membership in Age Verification Providers Association (AVPA) which currently has 27 members, two audit members and one associate member.

PAS 1296 describes an online framework that website operators can use to check the age of those buying from or accessing age-restricted sites. It does not endorse any specific tools for implementation.

In December 2019, the Digital Policy Alliance proposed an upgrade to PAS 1296:2018 that would make it auditable, similar to ISO standards.[6] The upgraded PAS defines:

- The key terms and definitions applicable to the age verification process.
- The roles and responsibilities of key actors including the requirement to establish their own age verification policies.
- The requirements for establishing levels of confidence, either Standard, Enhanced or Strict and associated testing methodology and statistical analysis to validate the age verification system.
- A common specification for how sources, outputs, levels of confidence and a trust framework are established.
- Requirements for privacy protection, data security and information system management.
- A common language to enable systems interoperability, for example tokenised age verification systems.
- Certification, monitoring and testing systems including frequency and detail of required ongoing monitoring and testing activities.

The upgraded PAS:

- Is technologically agnostic in terms of the age verification systems and their methodologies.
- Does not establish thresholds or required levels of confidence in the age verification process.
- Does not deal with financial or commercial models related to the age verification process.

[5] British Standards Institution, PAS 1296:2018 Online age checking. Provision and use of online age check services. Code of Practice

[6] Digital Policy Alliance, PAS 1296:2018 Upgrade to a Specification (& then to an International Standard)

PASS scheme

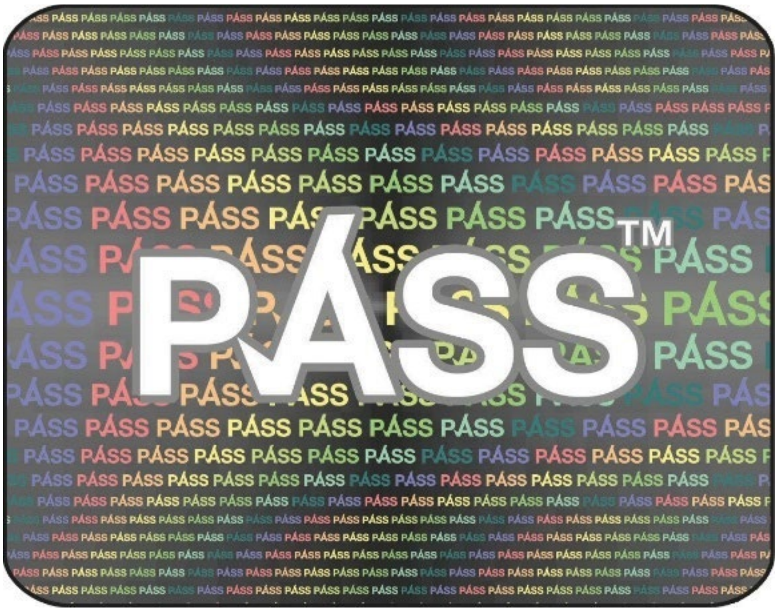
The United Kingdom's Proof of Age Standards Scheme (PASS) and was founded in 2001.[7] Its current implementation is based on the PAS 1296:2018 code of practice.[8]

The PASS scheme describes itself as follows:

> "The PASS Scheme is operated by PASSCO, a community interest company providing accreditation to suppliers of proof-of-age cards in the United Kingdom. The accredited providers are assessed against strict standards by qualified auditors to ensure that they operate to the highest standards. Sellers of age-restricted products and venues can be confident in accepting cards or digital proof of age with a PASS hologram or dynamic graphic, safe in the knowledge that the scheme is supported by the police, Trading Standards and a wide range of trade bodies." Currently the PASS card is only used at the point of sale or point of entry, and not for access to online sites.

PASSCO is a community interest company registered in England, UK.

The scheme has an identity hologram registered as a trademark in England:



At time of writing, seven card providers support the scheme and use the logo. Of the seven, five operate nationally:

- CitizenCard
- My ID Card
- OneID4U
- TOTUM
- Young Scot

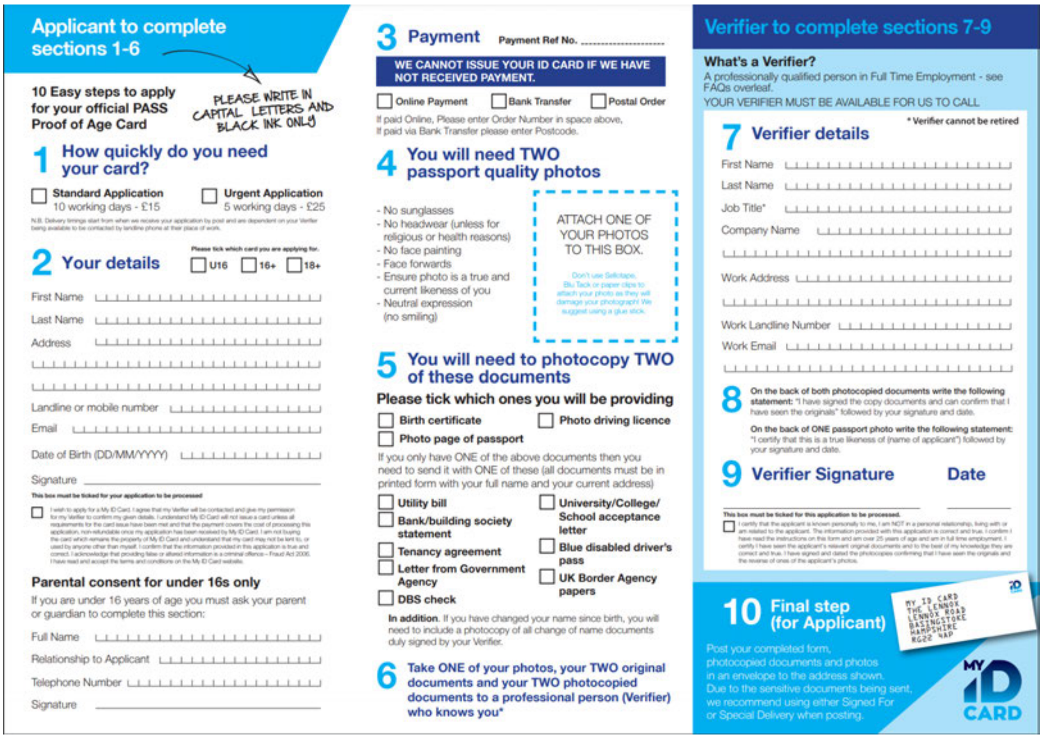
The other two providers operate regionally in Britain.

The cards support three age ranges: 18+, 16+ and under 16.

[7] Proof of Age Standards Scheme, What is PASS?
[8] See 'Pass scheme accreditation standards' at Proof of Age Standards Scheme, Your free downloads

Cards must include a photo, date of birth and the PASS hologram logo. The 16+ cards include the date that the user turns 18 and the under 16 cards include the date the user turns 16.

The application process is relatively simple. We have included a sample application form from My ID Card below. A My ID Card costs £15 for a standard application or £25 for an urgent application.



Applicant to complete sections 1-6

10 Easy steps to apply for your official PASS Proof of Age Card

PLEASE WRITE IN CAPITAL LETTERS AND BLACK INK ONLY

1 How quickly do you need your card?

□ **Standard Application** 10 working days - £15 □ **Urgent Application** 5 working days - £25

N.B. Delivery timings start from when we receive your application by post and are dependent on your Verifier being available to be contacted by landline phone at their place of work.

2 Your details

Please tick which card you are applying for. □ U16 □ 16+ □ 18+

First Name ____

Last Name ____

Address ____

____

____

Landline or mobile number ____

Email ____

Date of Birth (DD/MM/YYYY) ____

Signature ____

This box must be ticked for your application to be processed

□ I wish to apply for a My ID Card. I agree that my Verifier will be contacted and give my permission for my Verifier to confirm my given details. I understand My ID Card will not issue a card unless all requirements for the card issue have been met and that the payment covers the cost of processing this application, non-refundable once my application has been received by My ID Card. I am not buying the card which remains the property of My ID Card and understand that my card may not be lent to, or used by anyone other than myself. I confirm that the information provided in this application is true and correct. I acknowledge that providing false or altered information is a criminal offence – Fraud Act 2006. I have read and accept the terms and conditions on the My ID Card website.

Parental consent for under 16s only

If you are under 16 years of age you must ask your parent or guardian to complete this section:

Full Name ____

Relationship to Applicant ____

Telephone Number ____

Signature ____

3 Payment Payment Ref No. ....

WE CANNOT ISSUE YOUR ID CARD IF WE HAVE NOT RECEIVED PAYMENT.

□ Online Payment □ Bank Transfer □ Postal Order

If paid Online, Please enter Order Number in space above,
If paid via Bank Transfer please enter Postcode.

4 You will need TWO passport quality photos

- No sunglasses
- No headwear (unless for religious or health reasons)
- No face painting
- Face forwards
- Ensure photo is a true and current likeness of you
- Neutral expression (no smiling)

ATTACH ONE OF YOUR PHOTOS TO THIS BOX.

Don't use Sellotape, Blu Tack or paper clips to attach your photo as they will damage your photograph! We suggest using a glue stick.

5 You will need to photocopy TWO of these documents

Please tick which ones you will be providing

□ **Birth certificate** □ **Photo driving licence**
□ **Photo page of passport**

If you only have ONE of the above documents then you need to send it with ONE of these (all documents must be in printed form with your full name and your current address)

□ **Utility bill**
□ **Bank/building society statement**
□ **Tenancy agreement**
□ **Letter from Government Agency**
□ **DBS check**
□ **University/College/School acceptance letter**
□ **Blue disabled driver's pass**
□ **UK Border Agency papers**

In addition. If you have changed your name since birth, you will need to include a photocopy of all change of name documents duly signed by your Verifier.

6 Take ONE of your photos, your TWO original documents and your TWO photocopied documents to a professional person (Verifier) who knows you*

Verifier to complete sections 7-9

What's a Verifier?
A professionally qualified person in Full Time Employment - see FAQs overleaf.
YOUR VERIFIER MUST BE AVAILABLE FOR US TO CALL

7 Verifier details * Verifier cannot be retired

First Name ____

Last Name ____

Job Title* ____

Company Name ____

____

Work Address ____

____

Work Landline Number ____

Work Email ____

____

8 On the back of both photocopied documents write the following statement: "I have signed the copy documents and can confirm that I have seen the originals" followed by your signature and date.

On the back of ONE passport photo write the following statement: "I certify that this is a true likeness of (name of applicant") followed by your signature and date.

9 Verifier Signature ____ **Date** ____

This box must be ticked for this application to be processed.

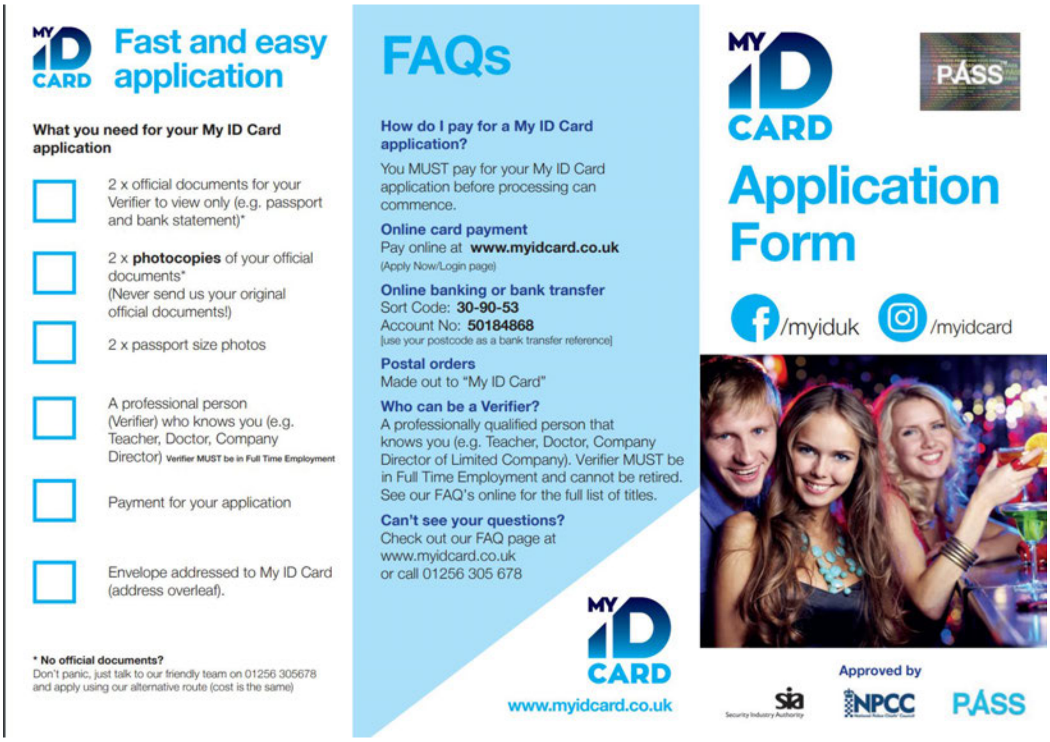
□ I certify that the applicant is known personally to me, I am NOT in a personal relationship, living with or am related to the applicant. The information provided with this application is correct and true. I confirm I have read the instructions on this form and am over 25 years of age and am in full time employment. I certify I have seen the applicant's relevant original documents and to the best of my knowledge they are correct and true. I have signed and dated the photocopies confirming that I have seen the originals and the reverse of ones of the applicant's photos.

10 Final step (for Applicant)

Post your completed form, photocopied documents and photos in an envelope to the address shown. Due to the sensitive documents being sent, we recommend using either Signed For or Special Delivery when posting.

MY ID CARD
THE LENNOX
LENNOX ROAD
BASINGSTOKE
HAMPSHIRE
RG22 4AP

MY ID CARD



MY ID CARD **Fast and easy application**

What you need for your My ID Card application

□ 2 x official documents for your Verifier to view only (e.g. passport and bank statement)*

□ 2 x **photocopies** of your official documents* (Never send us your original official documents!)

□ 2 x passport size photos

□ A professional person (Verifier) who knows you (e.g. Teacher, Doctor, Company Director) **Verifier MUST be in Full Time Employment**

□ Payment for your application

□ Envelope addressed to My ID Card (address overleaf).

*** No official documents?**
Don't panic, just talk to our friendly team on 01256 305678 and apply using our alternative route (cost is the same)

FAQs

How do I pay for a My ID Card application?
You MUST pay for your My ID Card application before processing can commence.

Online card payment
Pay online at **www.myidcard.co.uk**
(Apply Now/Login page)

Online banking or bank transfer
Sort Code: **30-90-53**
Account No: **50184868**
[use your postcode as a bank transfer reference]

Postal orders
Made out to "My ID Card"

Who can be a Verifier?
A professionally qualified person that knows you (e.g. Teacher, Doctor, Company Director of Limited Company). Verifier MUST be in Full Time Employment and cannot be retired. See our FAQ's online for the full list of titles.

Can't see your questions?
Check out our FAQ page at www.myidcard.co.uk
or call 01256 305 678

MY ID CARD
www.myidcard.co.uk

MY ID CARD PASS

Application Form

/myiduk /myidcard

Approved by
sia Security Industry Authority NPCC PASS

IEEE standard

The Institute of Electrical and Electronics Engineers (IEEE) standard IEEE Std 2089 – 2021, *IEEE Standard for an Age Appropriate Digital Services Framework Based on the 5Rights Principles for Children* was approved on 9 November 2021.[9]

The standard states its purpose as:

> "[Providing] a set of processes for digital services when end users are children, and, by doing so, aids in the tailoring of the services that are provided so that they are age appropriate. This is essential to creating a digital environment that supports, by design and delivery, children safety, privacy, autonomy, agency and health, specifically providing a set of guidelines and best practices and thereby offering a level of validation for service design decisions."

The standard is suited to online service or product providers where the services or products are age dependent, such as social media sites or the online sale of alcohol. The framework considers risk and children's age-appropriate needs. However, it does not specify any technology for determining the child's age.

The 5Rights Principles, on which the standard is based, are:

1. **The right to remove.** That is the right to easily remove what you yourself have put up. It doesn't challenge freedom of speech, but the first rule of conscious use is being able to control what your history will look like online, in the space you curate.
2. **The right to know.** That is the right to know who and what and why and for what purposes, your data is being exchanged. And a meaningful choice about whether to engage in that exchange.
3. **The right to safety and support.** What is illegal must be pursued by the law. But much of what upsets young people online is not illegal and support is sparse, fragmented and largely invisible to those children and young people when they need it most.
4. **The right to informed and conscious use.** It is simply undemocratic that young people are looped into technology that is deliberately designed to keep them attached, based on the same principles as casino slot machines. 'Addicting' is what product designers call it. 'Addicting' is what product designers work towards.
5. **The right to financial literacy.** Financial literacy from car insurance to budget management means understanding the purposes of the technology that you are using. Growing up as a creator and contributor as well as an informed consumer. And having a clear grasp of the likely social outcomes of that use.

ISO standards

The International Organization for Standardization (ISO) published a working draft titled *Information technology – Age assurance systems – Framework* in November 2021.[10] When completed, this international standard will define:

1. The key terms, definitions and abbreviations applicable to the age assurance process.
2. The roles, responsibilities and procedures of key actors in the age assurance process, including the requirement to establish age assurance policies.

[9] 5Rights Foundation, IEEE Standard for an Age Appropriate Digital Services Framework Based on the 5Rights Principles for Children

[10] euCONSENT, ISO Working Draft Age Assurance Systems Standard

3. The requirements for establishing the levels of confidence (zero, basic, standard, enhanced or strict) associated with output of the age assurance system.
4. A common specification for how sources, outputs, levels of confidence and a basis for a trust framework are established and communicated to other actors in the age assurance process, sharing, swapping or communicating the verified attributes or credentials.

In line with other existing standards, this standard will be technologically agnostic in terms of age verification methods and will not recommend age assurance thresholds.

The standard will not address, except at a high level, requirements for data security and privacy or establish detailed requirements for interoperability. Additionally detailed test methodologies for assurance components will not specified other than adopting evaluation methodology for IT security as set out in ISO 15408.[11]

Emerging standards

Through Enex TestLab's engagement with the market, we have learned that a proposal for a European age verification standard is currently being developed.

[11] International Organization for Standardization, ISO/IEC 15408-1:2022 Information security, cybersecurity and privacy protection — Evaluation criteria for IT security

Phase 2 a) Part 1: Outline and explain age assurance *(age verification, age estimation)*

There are three broad approaches to controlling children's access to unsafe online content:

- **Age verification** – using hard identifiers such as identity documents or third-party organisations to confirm a person's age.
- **Age estimation** – using biometric or other digital techniques to estimate a person's age.
- **Content filtering** – using age gates, safe search, filters, tagging or moderation to prevent a person's access to unsafe content.

The methods outlined below are relevant to children aged 18 years and under; methods that apply to adults – for example voluntary exclusion from online gambling or age verification for the purchase of alcohol – are not within the scope of this study.

The methods outlined in this study are not foolproof. In most instances, several methods may be combined to increase the accuracy of age verification, or estimation, and reduce the likelihood of circumvention. Some of these methods could be used for subsequent re-verification after establishing the user's age using more rigorous and reliable forms of verification.

This article provides an example of how multiple detection methods may improve the accuracy of age verification using digital identity, as opposed to physical documents.[12] It outlines a four-step process for age estimation in social media:

1. Biographic analysis of the user using natural language processing
2. Visual face recognition of age-related features
3. Social graph recognition – the ages of a user's friends and followers
4. Manual curation by an analyst in the more difficult cases.

We note that Australia, and other jurisdictions, may have some way to go on this front.

Age verification methods

It is important to stress that age verification is not, and should not be, *identity* verification.

Age verification assigns an age to an individual that online services can use to permit or block access to those services but does not uniquely identify that individual. Identity verification uniquely identifies an individual. Identity verification, if lost, stolen or intercepted, can be used to assume an individual's identity, including for criminal purposes.

The use and storage of user identities is a critical consideration. Age verification systems that require the user to upload photos or scans of identity documents could create a vector for these sensitive documents to be stolen or inadvertently leaked, creating a significant privacy issue.

Most of the solutions we evaluated carried out age checking in the cloud, requiring users to upload images or documents for processing. Ideally, the verification services would discard these ID documents immediately after processing and create a unique ID token that would verify the user's age but could not be used to *identify* the user.

[12] Luca Marchesotti, How to verify consumers' age on social media, LinkedIn, 18 August 2017

A few of the age verification systems we evaluated bypassed this issue by processing the data on the user's local device and storing a unique ID token on that device. This method, providing it effectively verifies the user's age, is in our opinion the best solution to prevent identity theft.

Age verification systems have three broad options for verifying a person's age:

- Scanning and validating **government-issued identity documents**
- Relying on third-party organisations such as **mobile phone providers** or **banks** to have verified the age of their customers
- Asking an adult, whose identity and age have been verified, to **vouch for** the person's age.

Self-declaration of age is not an acceptable verification or validation of age – a user could enter any age or date of birth they liked.

Government-issued identity documents

The typical method for using a government-issued ID (hard identifier) is for the user to photograph or scan and upload the ID so its details can be extracted using optical character recognition. The user would then upload or take a likeness photo to be compared with the photo on the ID document.

These government-issued identity documents may be useful in age verification:

- **Passports.** A very safe and accurate age verification method because of the strict proof-of-identity requirements of the Australian Passport Office and overseas passport issuers. Passports may present an issue for very young children if their likeness changes rapidly as they grow, and their passport photos don't match closely enough (we will discuss and other potential accessibility issues later in this report).
- **Driver's licenses and learner's permits.** These are of limited use in the Australian context because learner's permits are only granted to 16-year-olds (15 years 9 months in the ACT) and driver's licences to 17-year-olds (16 years 6 months in the NT and 18 years in Victoria). New South Wales and South Australia currently offer digital licences, and Queensland will soon offer them. For the purposes of age verification, digital licences operate very similarly to physical ones: the individual needs to upload a photo of the digital ID and may be required to provide a likeness photo to match the photo shown on the ID.
- **Boat (marine) licences.** These are available in restricted form at age 12 and full licences are available at age 16.
- **Student cards.** University student cards include photos and require solid identify verification. A primary or secondary school student card could be issued by schools or state education departments, and students must provide birth certificates or passports to enrol in schools. Given the wide use of mobile phones by students at quite early ages, it could also take digital form. This could cover ages 4 to 18.
- **National identity cards.** Australia does not issue these, but other countries do and employ them for age verification. For example Angola's national identity card includes date of birth, photo and a fingerprint.

Mobile phone accounts

In 2020, the Australian Bureau of Statistics revealed 33% of children aged 6 to 13 owned a mobile phone.[13] This figure is likely to be higher in 2022. However, children cannot take out their own

[13] eSafety Commissioner, Your child's first smartphone – are they old enough?, June 2019

mobile phone contracts and require a parent to sign up for them.[14] The mobile phone provider currently only needs to verify the parent's age and identity, so it's unlikely that the child's age would be recorded to act as an age verifier, or indeed, if that was even accurately recorded for verification.

Banks

In Australia children aged 14 and over can open bank accounts in their own name without parental assistance. For example, at the Commonwealth Bank a 14-year-old can open a Youthsaver account with a copy of their birth certificate or passport.

Until recently there was no simple way to use a bank account as a convenient form of age verification. However, some Australian banks have recently started providing identity verification as a service to retailers.[15]

Credit cards are not relevant as a form of ID for minors – you must be 18 years old to apply for your own credit card. Children under 18 can have a credit card in their parents' account name, but the bank only needs to verify that parent's age.

Vouching

This is where a credible member of the community, such as a child's teacher, doctor, parent or guardian, confirms a child's age. This can be a useful method for children who do not have readily available documentation such as a passport or birth certificate. The Open Identity Exchange (OIX) has proposed a "Digital Vouch with Photo" which, if successfully adopted, would provide a digital photo verification "token".

Self-declaration

Relying on users to declare their age voluntarily should not be considered a reliable or accurate method of age verification. UK communications regulator Ofcom, for example, suggests that in the UK up to a third of children aged between 8 and 17 have adult social media profiles.[16]

Age estimation methods

The methods discussed above require documentary evidence as proof of a person's age, either directly provided or relying on an authoritative third party such as a bank to have previously verified it. Biometric and algorithmic methods can estimate a person's age in other ways; however, they lack the hard proof of age verification. Here we discuss a range of available and proposed age estimation methods.

Biometrics

Biometrics is a method of using physical aspects of a person, as measured by a computer, to estimate their identity, age or other factors. An advantage of biometrics is it relies on something a person *is*, rather than something they *have* (which can be lost) or something they *know* (which can be forgotten). For the purposes of age estimation, biometric physical aspects of a person include:

- **Facial age assessment.** Several facial age assessment products are available, each at various stages of maturity. Vendors of such products include Yoti and Privately. These products use machine learning (ML) models to estimate a person's age based on their facial proportions and characteristics. One advantage of this method is that it doesn't need

[14] FindLaw Australia, Can children sign mobile phone contracts?

[15] James Eyers, Banks ready to launch new digital identity checking service, *Australian Financial Review*, 31 August 2022

[16] Ofcom, A third of children have false social media age of 18+, 11 October 2022

to store users' identities; instead, it can store a hash token of their facial characteristics and inferred age as a unique identifier. Even if this identifier is stolen, it can't be used for identity theft. Of all the biometric methods, this has the most potential in terms of expediency and accuracy.

- **Voice print age assessment.** This machine learning technology typically asks a user to talk about a particular topic, answer a question or read out a paragraph from the device's screen. It can assess many characteristics including the pitch, cadence and delivery of the voice, the sentence structure of the reply, the sophistication of the wording and the level of understanding of the concept under discussion. This method is not currently as mature as facial age assessment, but its accuracy is improving. Some vendors, Privately for example, use it in conjunction with facial age assessment.
- **Written age assessment.** This technology estimates a person's age by measuring how they express themselves when writing. It assesses many of the same characteristics as voice print age assessment, with the addition of how they type and keystroke patterns. This method is also maturing in accuracy.

It's important to note that language comprehension and fluency can vary for many reasons, so voice and written age assessments may be biased against people who have disabilities or who don't speak the local language fluently. Additionally, an accident or medical condition may also impact the accuracy of the facial characteristics of an otherwise healthy subject. It might be advantageous to combine several biometric methods to improve the accuracy and reliability of age estimation.

Social media profiling

Social media networks have significant problem with underage users. An Ofcom study in the UK found that 32% of users aged between 8 and 17 claimed to be adults in their online profiles.[17]

Social media networks have internal mechanisms for gauging the true age of their users, however these approaches are not transparent, and their accuracy is unknown. A number of academic papers have proposed approaches for estimating social media users' ages by analysing their postings.

The 2018 paper A Model for Age and Gender Profiling of Social Media Accounts Based on Post Contents examined a range of algorithms to identify the gender and age group of social media accounts by analysing the contents of their posts.[18] The researchers considered factors including 'socio-linguistics, grammar, characters and words'.

In 2022, the paper *ReportAGE: Automatically extracting the exact age of Twitter users based on self-reports in tweets* evaluated a method that automatically identified the exact age of users based on self-reports in their tweets.[19] The natural language processing algorithm these researchers developed analysed more than 1.2 billion tweets posted by more than 245,000 users and predicted ages for 54% of them.

[17] Ofcom, A third of children have false social media age of 18+, 11 October 2022
[18] Cheng, JK, Fernandez, A, Quindoza, RGM, Tan, S, Cheng, C (2018) A Model for Age and Gender Profiling of Social Media Accounts Based on Post Contents, *Neural Information Processing,* ICONIP 2018
[19] Klein AZ, Magge A, Gonzalez-Hernandez G (2022) ReportAGE: Automatically extracting the exact age of Twitter users based on self-reports in tweets, *PLoS ONE* 17(1)

Algorithmic profiling

Other examples of algorithmic profiling include methods for analysing mobile phone data and usage patterns.

This paper from 2019, *Predicting customer's gender and age depending on mobile phone data*, proposes a method that predicts mobile device users' gender and age based on their behaviour, services usage and contract information.[20] The model achieved 85.6% accuracy in predicting users' gender and 65.5% accuracy in predicting their age. While this technology was primarily developed for marketing purposes, it has the potential to be applied for eSafety's needs.

A study by Zhejiang University in China and the University of South Carolina in the United States claimed to distinguish between young children and adults based on their mobile phone screen swipes and tapping.[21] However, this study only claimed it could detect very young children (aged 3 and under) and very little further information is available.

[20] Al-Zuabi, IM, Jafar, A & Aljoumaa, K (2019) Predicting customer's gender and age depending on mobile phone data, *J Big Data* 6, 18

[21] Zhejiang University, This age-detecting algorithm can make your smartphone child-proof, 12 February 2018

Phase 2 a) Part 2: Outline and explain relevant safety technology

A variety of relevant safety technology methods are available and currently employed to control or limit children’s access to harmful online content.

These include:

- age gates
- safe search
- Internet content filtering (ICF)
- tagging
- moderation.

Age gates

Age gates are a commonly employed age assurance method to assess whether consumers are of an appropriate age to view or buy products. The most common method is to ask a user to enter their birthday or self-verify that they are of the appropriate age to enter the site. Age gates are a low-cost and low-effort method for content providers to implement. In most cases, age gates rely on an honour system and are easy to circumvent.

Safe search

Safe search is a search engine filter that blocks explicit content such as pornography. A number of major search engines – including Bing, Duck Duck Go and Google – offer safe search options. A child can circumvent safe search by disabling it or using another search engine. Many ICF products allow a parent to enforce safe search to be enabled, Google also enforce their safe search on Google family linked accounts.

Internet content filtering

Content filtering blocks access to internet content based on a child’s age and related criteria. This is a method enabling parents to control their child’s web surfing. The filter blocks access to content that is deemed age-inappropriate based on the filter settings and allows access to age-acceptable material.

Many internet content filter (ICF) products or services can be configured by selecting a child’s age range or age gate limit. It is then left up to the filter vendor to set and apply the criteria it considers appropriate for each age range.

The majority of content filters allow finer adjustments of filtered content in categories such as adult content, pornography and racism and hate speech. While this level of granular control may be beneficial for proactive parents or carers, those wanting a quick fix to controlling their children’s internet access may overlook these controls.

Content filters may also over-block, preventing access to sites that the child should be able to view. This may include content related to sexual health and support services, such as support for LGBTQIA+ children. Over-blocking can lead to frustration, leading parents to reduce the effective level of the product or even turn it off entirely.

Another issue with content filters is circumvention. Technically savvy children may be able to turn off or bypass the filtering service to access content they should be prevented from viewing.

Internet content filtering can be implemented in three ways.

- **By the internet service provider.** A number of Australian service providers broadly apply content filtering using third-party lists, to exclusively prevent access to the listed sites and images, such as the UK Internet Watch Foundation's URL list,[22] the Office of the eSafety Commissioner's Prohibited URL Filter list[23] and the Interpol worst of list.[24] However, content filtering at the service provider level traditionally has not proven a popular choice with the public or internet providers in Australia. Internet providers have claimed that implementing and delivering granular parental filtering controls to their customers is cost prohibitive and may affect the performance of service to all customers.[25] This claim was last independently tested in Australia in 2009, when performance impact was found to be minimal but the technologies implemented by service providers were easy to circumvent. Some Australian internet service providers may provide upstream filtering of the Interpol worst of list for their subscribers.
- **By software installed on the user device.** A commonly implemented method is filtering software installed directly on the child's or family's device and integrated into the browser. This is, at time of writing, the most flexible and the most accurate form of ICF implementation. For Australia, a list of third-party accredited filters can be found under the Family Friendly Filter Scheme which assesses ICF product effectiveness and vendor claims against the Office of the eSafety Commissioner's PUF list.[26]
- **By a home hardware device such as a smart router.** A number of home routers have parental control settings that provide control over:
  - The time children can access the internet, such as a set number of hours per day or a series of permitted and blocked times over each day of the week.
  - Blocking content using commercially available filters but implemented at the router level rather than in software on each device. These offer similar configuration options to software ICFs but with the complication of requiring a parent to navigate the router's login and menu system.
  - Manually configured block or allow lists. These are not widely used because it may be time consuming for parents to enter sites into the router's list and they require a relatively high level of technical skill to perform it successfully.

While the smart router may provide adequate control at home, a child is no longer protected once they step outside and connect their device to a mobile network or unfiltered internet service.

Tagging

Tagging relies on content providers or hosts to voluntarily label their content according to age appropriateness. Unlike films, video games and publications, most online content is not classified before its release. It's up to the poster or host of the content to label their content as suitable or not for various audiences. However, this is no more rigorous than relying on a social media poster to use a hashtag such as #NSFW (not safe for work) to indicate that content is not suitable for display or sharing in a working environment. Also, children can search for those tags as a convenient way to locate and access content that may be harmful to them, some services will block these results from displaying to accounts known to be children.

[22] Internet Watch Foundation, URL List

[23] eSafety Commissioner, What is illegal and restricted online content?

[24] Interpol, Blocking and categorizing content

[25] Internet Society, Internet Way of Networking Use Case: Content Filtering, 17 December 2020

[26] Communications Alliance, Family Friendly Filters

Another option is for content producers and hosts to include a restricted to adults (RTA) metatag (a tag that is read by web browsers but not visible to users), enabling filters that can read page-level metatags and tag attributes to filter out RTA content.

Some services apply community rules that require tagging and implement moderation tools that attempt to identify content that has not been appropriately tagged, enabling those service providers to take enforcement action against users who fail to tag their content.

Moderation

Moderation is where the owner or administrator of an online group, forum or platform oversees and decides on the activities occurring in that domain. This may involve actively policing and reviewing material – using human moderators, algorithmic filters or both – or providing a mechanism for users to complain about material. Moderation, and particularly automated moderation, may result in under-blocking or over-blocking of unsuitable content.

The larger social media companies have experienced considerable difficulties actively monitoring and policing content on their platforms, and sometimes harmful content is only taken down after complaints or negative publicity.

Any form of moderation relies on the moderator’s judgement and the policies of that platform to determine what is and isn’t acceptable. Further, any independent review or accountability of moderation practices depend on the transparency of the service provider.

Phase 2 a) i): Case study: euCONSENT

euCONSENT is a European Commission funded project to develop an EU-wide computer network for completing online age verification and securing parental consent when younger children wish to share personal data.

To date, euCONSENT has run two pilot projects to evaluate the effectiveness of age verification technologies.

Early pilot (September 2021)

The first pilot involved 63 adult participants from across Europe.[27] The participants were asked to provide their age, complete three missions (shown below) and answer general questions about usability and applicability of the test system. The missions were:

- Use an age verification provider (AVP) to have their age checked for the first time, in order to buy alcohol from an online store.
- Use a previous age check to get access to the online store, without having to prove their age again.
- Get access to the online store by using another AVP than the one directly supported by the store.

Trial results were primarily concerned with experimental design. It found that improvements to the user interface were required to prevent user confusion. Technical data is not available in the EU early pilot report.

Large-scale pilot (February to March 2022)

The second pilot involved over 1,700 people including children, parents and other adults.[28] Test subjects were selected from five countries in Europe: Belgium, Cyprus, Germany, Greece and the UK.

Participants did not all have the same missions in this trial. Five dummy websites were created, each translated into four languages as applicable to the countries involved. Sites included online stores for alcohol or knives, a dating site and various social media.

The trial used three AVPs: AgeChecked, AGEify and Yoti. Where required, parental consent was obtained using JusProg or Upcom. This requires that each platform needs to share information on the user.

Various age checking methods were used, including: age estimation from face, or verification of an ID document or credit card. Each mission given to a participant used a different AVP.

Missions included:

1. Access a dummy alcohol website where they had to verify their age using one of the methods provided. Facial recognition was the preferred method because participants were children without ready access to identity documents. Once the age was verified, the child was redirected to a ‘protected’ page informing them that they were too young to access that site.
2. Access a dummy social media website where they were recognised as being under the digital age of consent, based on their age as determined in the previous mission (using an

[27] euCONSENT, Pilot Execution Report – early pilot, 30 September 2021
[28] euCONSENT, Pilot Execution Report – first large scale euCONSENT pilot, 15 April 2022

age token stored as a cookie). Now they had to ask for parental consent to be able to move forward in the survey.

3. Access a dummy chat website where they similarly had to ask for parental consent to complete the mission.

Results

The pilot execution report illustrated a key issue with this kind of technology – humans. On mission 1, children were given the option to choose which age verification method they used to access an age-inappropriate website.

Verification method	% of participants
Face estimation	59%
Document scan	22%
Itsme (app)*	15%
Credit card verification	3%

*An AVP service delivered via an application on a user's smart device (Apple App Store and Google Play Store).

About 71% of participants scored this step as 8 to 10 out of 10 for easiness to complete. When asked if users needed parental help to complete the mission, 33% said yes.

As the report points out, this indicates that a significant proportion of parents were happy to give their child access to an inappropriate website using their own ID documents. (It is unknown if this is due to relaxed parental attitudes or a misunderstanding of the mission objectives.)

Children would likely not have succeeded with verification because they were underage. Parents most likely assisted by providing their own faces or identity documents. While the report didn't specify the reasons children needed parental help, the fact that document scans and credit cards were used in some cases suggests that parents helped circumvent the technology.

Ultimately, 35% of children gained access to an inappropriate website.

In addition, there may be issues relating to availability of checking-services, which makes it important to have options available to the user. Itsme was not available in all countries involved in the pilot. Yoti is only available with Android or iOS, not with Microsoft Windows. Not everyone has a driver's licence or credit card.

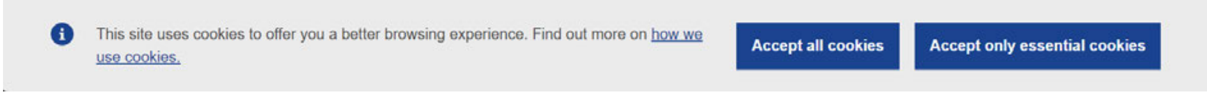
The report says nothing about the accuracy of age estimation techniques.

Interoperability between the systems responsible for age checks and parental permission was rated at almost 100% – AVP and parental permission services could reliably communicate with the webpages to which children were seeking access.

HTTP cookies and privacy

The euConsent documentation says little on the topic of cookies. However, under the EU ePrivacy Directive 2002, a website can't use cookies to process a visitor's data without first detailing the types of cookies used on the website and gaining explicit consent from the user for their use. Users must be informed of the category and purpose of cookies, what kind of data is processed, and the length of time the cookies persist. There must also be a clearly stated method for users to relinquish consent later on.

These are most commonly implemented in the form of cookie consent banners, such as the example below.



This site uses cookies to offer you a better browsing experience. Find out more on how we use cookies.

Privacy laws in Brazil, California and South Africa have similar requirements.

Personal data regulated under the EU directive includes names, addresses, email addresses, identification card numbers, location data, an internet protocol (IP) addresses, cookie IDs, mobile device advertising identifiers and other data that uniquely identifies a person.[29] Additional rules apply to safeguard special categories of personal data "… revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation."[30]

From an Australian context, while Australia's RAS Declaration 2022 requires websites to take reasonable steps to ensure the age of users when accessing adult material,[31] this is likely to have little impact on the majority of websites which are hosted in other jurisdictions. It is relevant for sales of alcohol and tobacco within Australia – or even on major international sales platforms – but its impact on access to internationally hosted pornography sites is unclear.

Additional items

euCONSENT considers not just technical items, but also the privacy and rights of users and the need to protect both children and website owners.

In most cases, apps and websites rely on self-declaration for age-verification and even parental consent.[32] This would generally not meet the expectations of government in situations where there is a legal obligation to protect children from online material by verifying their age.

Privacy policies

Privacy policies can be very general, but ideally should give children (and parents) options to selectively consent to privacy options. This not only gives more granular control to the user but may better inform the user what they are consenting to. How many children – or even parents – read the privacy policy they are consenting to?

Optimally to ensure privacy, websites should not be aware of the identity of users and services providing age verification should never be aware of the websites being accessed by the user,[33] unless in jurisdictions where legislation requires that data be retained.

[29] European Commission, What is personal data?

[30] European Union, Processing of special categories of personal data

[31] Australian Government, Online Safety (Restricted Access Systems) Declaration 2022, 13 January 2022

[32] euCONSENT, Methods for Obtaining Parental Consent and Maintaining Children Rights, September 2021

[33] Commission Nationale de l'Informatique et des Libertés, Online age verification: balancing privacy and the protection of minors, 22 September 2022

Children's rights versus protections

Parental consent is about giving parents options to control privacy and web-content protections for their children. It should not be mistaken for a parental right to invade a child's privacy online. Permission is being given for a child to share limited information online – and teenagers in particular may not want this sharing to include parents.

Verification methods must be inclusive

A child must not be excluded on the basis that the parent does not have access to a credit card or other identity document. Also, as noted above, not all verification technologies are available in all countries and all platforms.

Lessons to inform the eSafety roadmap

Enex TestLab believe eSafety can draw a number of lessons from these studies to inform its roadmap for age verification. These lessons relate to user rights, legal protections for users and providers, usability and accuracy.

- The large number of participants who circumvented age restrictions in the euCONSENT trials indicate the need for very clearly stated objectives in any Australian trials. For example, trials could include a broader range of sites (such as alcohol sales or dating sites) to avoid stigmatising an individual category such as pornography. If testing system accuracy, it should be very clear to parents that giving the child access to an adult site is inappropriate. (Though it might be appropriate if testing circumvention techniques.)
- If a child does get assistance from a parent, or a parent fails to secure passwords or identity documents that could be used by a child to misinform a website provider, who is liable for any exposure of the child to inappropriate content?
- It is essential to have two or more options for proof of age. This gives users a choice to provide details they consider less private. Also, it allows for people who don't have access to a particular method of verification – for example:
    - No children and even some parents may not have a driver's licence (ID verification).
    - A device may not have a webcam.
    - The age verification technology may not be compatible with the platform on the user's device.
- While English is Australia's official language, should age verification methods be made available in other languages commonly used in Australia?
- Privacy consent should ideally give users options:
    - To have granular control over their privacy
    - To understand the nature of the data being shared
    - To understand and consent to the use of HTTP cookies, if they're used to track age of user.
- AVP services should not pass user identity to website providers and website providers should not pass to AVPs what content users have accessed, as in the euCONSENT trial. AVP services could generate a token without personal identifiers and attach it to a particular user or machine, and websites could then be given access to the token which only provided age data.
- Given that facial recognition cannot provide precise age identification, its usage should be provisional. It may be part of a two-factor (or more) identification process where the estimated age is close to an age cut-off. Some faces are particularly hard to assess for age – even for humans. (Consider, for example, a child with a premature aging condition.) This would allow for very significant errors in either direction (either over- or underestimating a person's age), making alternative verification methods necessary.

- In any Australian trial, determining the level of assurance that is appropriate for the use case at hand, for example, access to online pornography.
- In an Australian trial, we would want to test a more representative sample, e.g., how biometric accuracy performed for users who identified as men, women or nonbinary, and for those for whom English was not a first language. Any method that relied on identity documentation would have to consider those whose live identity did not match the name or gender on their ID documents.
- euCONSENT is intended to be accessible beyond Europe. Should Australia consider sharing resources with euCONSENT for the sake of efficiency and economy, rather than reinventing the wheel? Such sharing could include mutual recognition, harmonisation on international standards and trialling technologies that had already been accredited as complying with requirements in other jurisdictions.

Identification of other relevant case studies

During the research phase of this study (November 2022), Enex TestLab did not discover any further relevant case studies.

Phase 2 b) Evidence based assessment of age assurance and relevant safety technologies

Enex TestLab assessed a range of age assurance technologies for the following criteria based on eSafety’s consultation feedback with stakeholders: feasibility, sensitivity of data, security and integrity, barriers to inclusion, potential for bias and accessibility. Wherever possible, we applied objective assessments and measures; in some instances, we relied on subjective assessment and observation.

Enex TestLab cannot comment on the installation, configuration and administration of these services. They are typically accessed through application programming interfaces (APIs) or software development kits (SDKs) and need to be integrated with existing services by software developers. In many cases, we could only use the demos that vendors had provided us. These aspects of the services could be further explored and quantified with in-depth testing.

All aspects of the services Enex TestLab could evaluate is included in the test results section for each service tested.

Phase 2 c) Evaluation Results & Analysis

Yoti

All testing was carried out over the Web at the Yoti World website using Yoti's estimate your age using your face demo.[34] This is presumably very similar to the application for ad hoc age verification in normal use.

The screenshots on the following page demonstrate the privacy disclaimer and the testing process for the demo.

The demo displayed fine granularity of age ranges compared to some of its competitors. During our testing, Yoti returned the following unique age ranges for test subjects under 13 years old: 0–4, 1–5, 3–7, 4–8, 5–9, 6–10, 7–11, 8–12 and 9–13. The AI does not limit itself to critical age points such as 13, 16, 18 and 21 – it can estimate a very large distribution of age ranges. This capability is probably more than most potential clients need.

The demo is easy to use; aligning your face in the template is not as slow and onerous as some.

Accuracy was quite good, far better than some competitors. Yoti had a slight tendency to overestimate the age of younger children. When we tested it with an older adult in their sixties, they were under aged by as much as seven years. However, in this age category, it is not of great consequence.

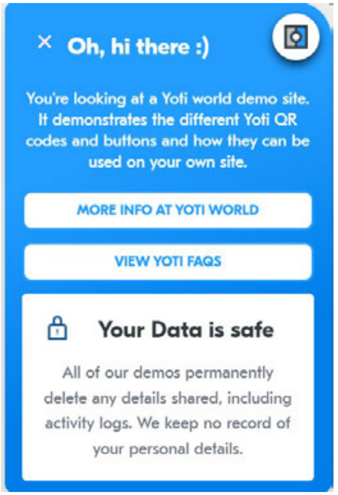
Yoti consistently provides age estimates with a width of four years. We assume the actual estimate is the midpoint of this range. This has the potential for a child who was under a particular age point to be estimated as being older than that point. This issue could be eliminated by using Yoti in conjunction with an age verification or age token system that required more reliable forms of ID.

Yoti in summary

- 64% of tests classified subjects in the correct age bracket
- 53% of results averaged over multiple tests classified the correct age bracket
- On average, the midpoint of the estimated age range differed from the true age by 21%.
- The true age was, on average only 6% outside the estimated age bracket.
- The assumed age estimation (mid-point of returned range) differed on average by 1.6±0.5 years from the actual age with a standard deviation of 2.8 years.
- If we consider only the magnitude of the error, without considering if the estimate is over or under the actual age, the average error is 2.6±0.5 years
- In some cases, the estimated age of a subject varied markedly over five tests; with one seven-year-old was assigned ages in the range 6 to 14.

[34] YOTI World, Age estimation

Yoti screenshots

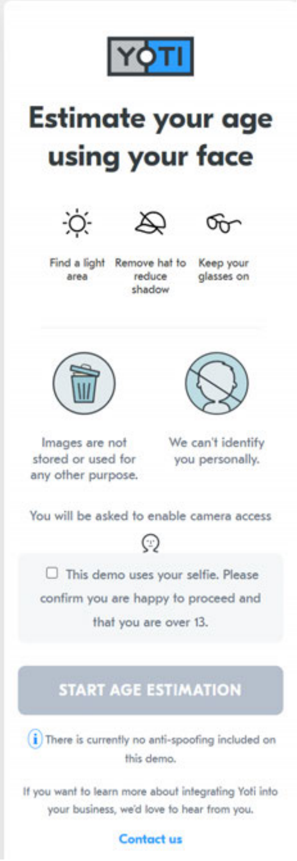


× Oh, hi there :)

You're looking at a Yoti world demo site. It demonstrates the different Yoti QR codes and buttons and how they can be used on your own site.

Your Data is safe

All of our demos permanently delete any details shared, including activity logs. We keep no record of your personal details.



Estimate your age using your face

Images are not stored or used for any other purpose.

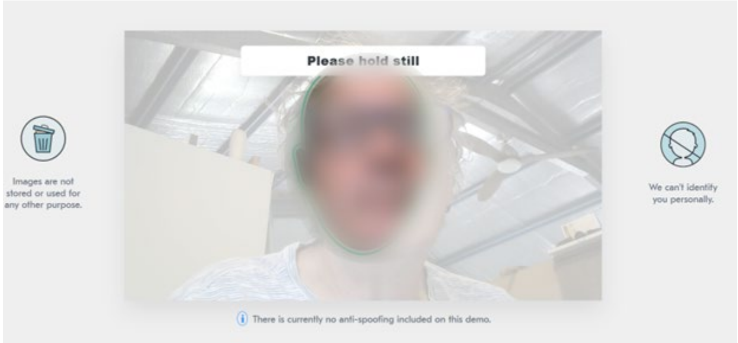
We can't identify you personally.

You will be asked to enable camera access

□ This demo uses your selfie. Please confirm you are happy to proceed and that you are over 13.

ⓘ There is currently no anti-spoofing included on this demo.

If you want to learn more about integrating Yoti into your business, we'd love to hear from you.

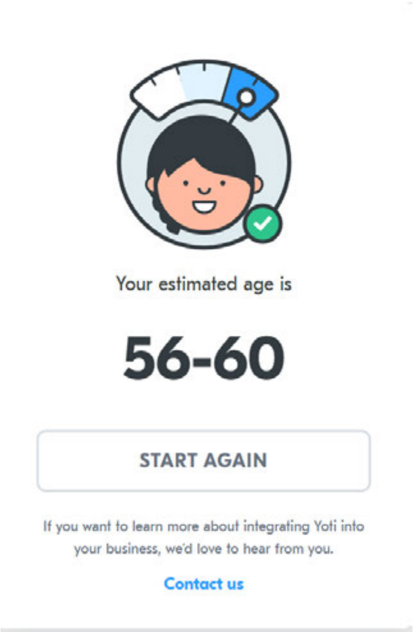


Please hold still

Images are not stored or used for any other purpose.

We can't identify you personally.

ⓘ There is currently no anti-spoofing included on this demo.



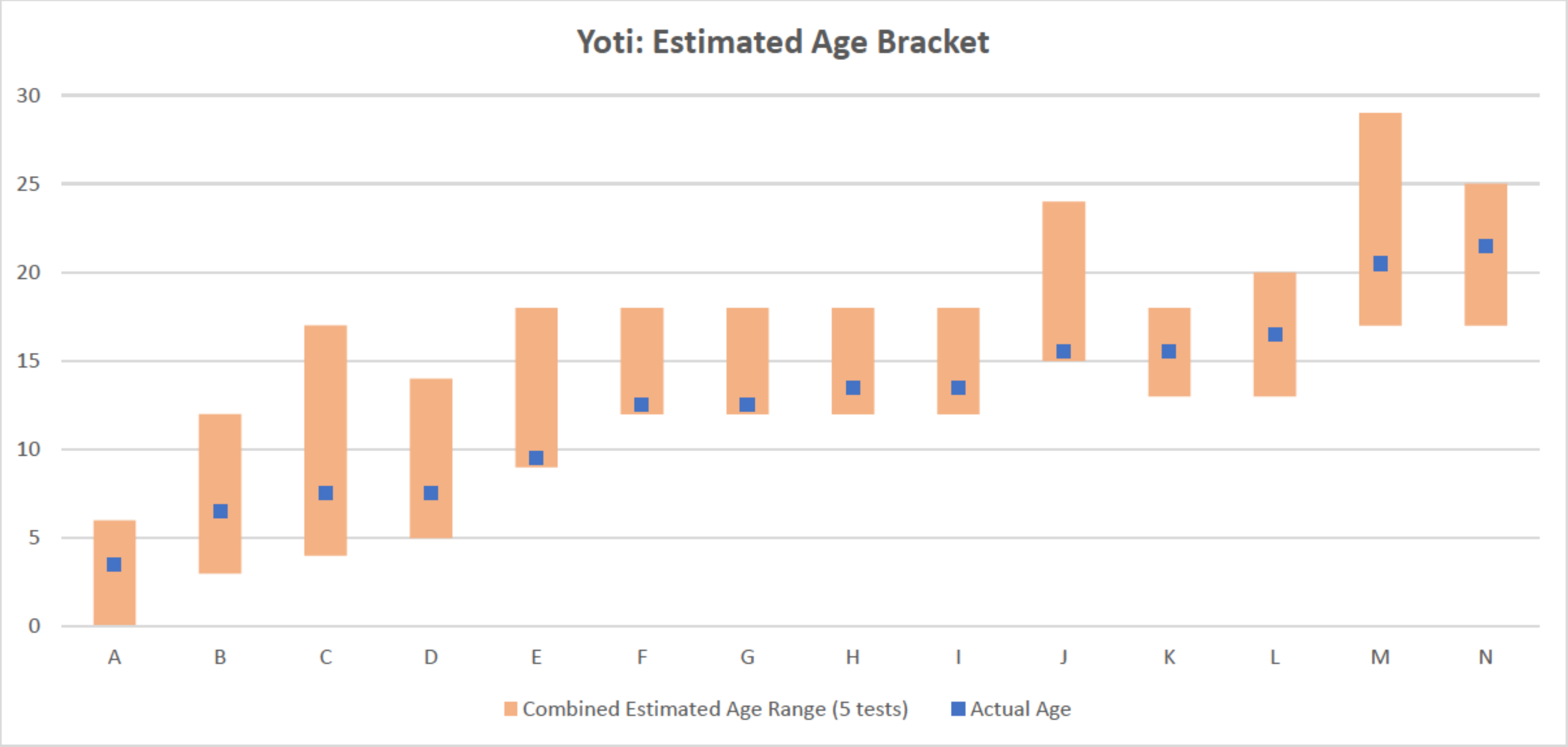
Your estimated age is

56-60

If you want to learn more about integrating Yoti into your business, we'd love to hear from you.

Yoti test results

First name	Current age	Attempt 1	Attempt 2	Attempt 3	Attempt 4	Attempt 5
A	3	1-5	1-5	1-5	1-5	0-4
B	6	7-11	3-7	4-8	4-8	4-8
C	7	11-15	12-16	5-9	4-8	8-12
D	7	9-13	6-10	6-10	6-10	5-9
E	9	12-16	12-16	13-17	9-13	11-15
F	12	12-16	12-16	13-17	13-17	12-16
G	12	12-16	13-17	12-16	13-17	13-17
H	13	12-16	13-17	12-16	13-17	13-17
I	13	12-16	12-16	13-17	13-17	13-17
J	15	16-20	16-20	19-23	15-19	16-20
K	15	13-17	13-17	13-17	13-17	13-17
L	16	13-17	14-18	15-19	14-18	15-19
M	20	24-28	22-26	20-24	17-21	23-27
N	21	17-21	18-22	17-21	20-24	19-23

Yoti: Estimated Age Bracket
30
25
20
15
10
5
0
A
B
C
D
E
F
G
H
I
J
K
L
M
N
Combined Estimated Age Range (5 tests)
Actual Age

Assessment of Yoti

Design

Item	Assessment
Feasibility – is the technology ready to be rolled out?	Technology is in operation but has limited client feedback to draw any conclusions. Online publishers including Meta and Yubo have incorporated Yoti into their services.[35]
Extent and sensitivity of the data required for the technology to operate	Other than an image, no data is required for this form of age identification. Once an image is processed, Yoti claims to erase the image.
Security and technical integrity of the technology	As no sensitive data is retained, security does not pose a problem. Accuracy is not perfect, and some children may be able to pass unacceptable age gates. Yoti was certified to ACCS 1:2020 (Technical requirements for Age Estimation technologies) on 25 November 2020.
Barriers to inclusion – what identification is required for verification?	None significant
Potential for bias and accessibility	Research literature has received age accuracy biases when comparing women to men and for some ethnic groups.[36]
Accreditations	AVPA, Age Check Certification – Challenge 25, e-IDVT, KJM, FSM[37]
Cost	A potential client of Yoti who was interviewed claimed the cost of implementing this form of age gating for a website can be around $1 per app downloaded by each user. One interviewee during our research said this cost could not be sustained by smaller companies with a modest number of paid subscribers.

Implementation

Item	Assessment
Transparency over decision making, use and appeals, appropriate oversight, and governance	Yoti provides no transparency over the decision making leading to the age range given. The service implementing Yoti can elect to allow another form of age assurance if facial analysis does not grant access.
Can the technology be implemented flexibly for different business models?	APIs, SDKs and assortment of developer documentation is available. Yoti has solutions for all types of age verification.
Can unintended consequences of the technology design be mitigated through the way it is implemented?	This depends on the service implementing Yoti and what it allows. Yoti has many age verification solutions, and the service can decide which ones and how many to implement to best suit its needs.

[35] Yoti is a member of AVPA

[36] Kärkkäinen, K, Joo, J (2021) FairFace: Face Attribute Dataset for Balanced Race, Gender, and Age for Bias Measurement and Mitigation, Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, January 2021

[37] Please see glossary for accreditation information

Privately

All testing was carried out over the Web at Privately's online showroom using the Multimodal Age Test; face image and voice analysis. This is purely a demonstration site. In practical application, users need to load the app into their browser or device. The app performs all face and voice analysis on the device – Privately claim no image or voice data is uploaded.

In these circumstances, we can't draw any conclusions about the accuracy of the app on a user's device. The comments below relate to the online demonstration.

The screenshots on the following page demonstrate the privacy disclaimer and the testing process for the demo site.

The demo appears to resolve four age ranges for both face and voice: 0–13, 13–17, 18–24 and 25+. However, the 0–13 category was not implemented for face in this demonstration. These ranges fit with the age limits of 13 for most social media and 18 as a generally accepted age of adulthood.

The voice analysis appears less developed than the face analysis. Voice analysis usually took significantly longer to process than the face image and at times it would not return a result – it would just say 'processing' after the face analysis returned its result (see screenshot). A further indication that the voice analysis could do with further development is that one subject, instead of reading the prompted text, simply made a short beeping noise – which resulted in a voice age estimate of 0–13.

Our test results showed the voice analysis less accurate than the face analysis in the higher age ranges. It accurately judged 0–13 for some of the younger children, but children as young as six were identified as adults by the voice analysis.

Face analysis also produced some inaccurate results in the younger age groups. A 3-, a 6- and a 7-year-old were judged to be 13-17 for every face analysis, which is quite worrying.

Privately in summary

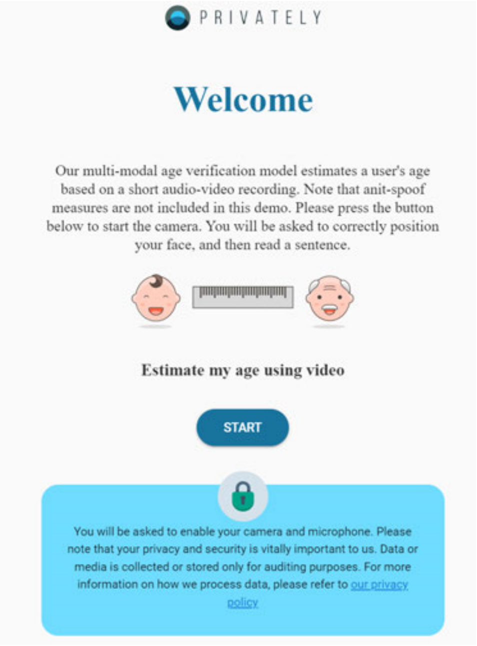
Face:

- 34% of tests classified test subjects in the correct age bracket
- 29% of results averaged over multiple tests classified the correct age bracket
- On average, the actual age was nearly three years outside the estimated age bracket
- Over five tests per person, the results were often consistent with each other

Voice:

- 44% of tests classified test subjects in the correct age bracket
- 31% of results averaged over multiple tests classified in the correct age bracket
- On average, the actual age was about four years outside the estimated age bracket
- Over five tests per person, the results were often highly variable – for example, a single 12-year-old was classed as 13–17, 18–24 and 25+ in successive tests
- These results exclude the case of one seven-year-old child who remained silent during the voice test and was classified as an adult

Privately screenshots

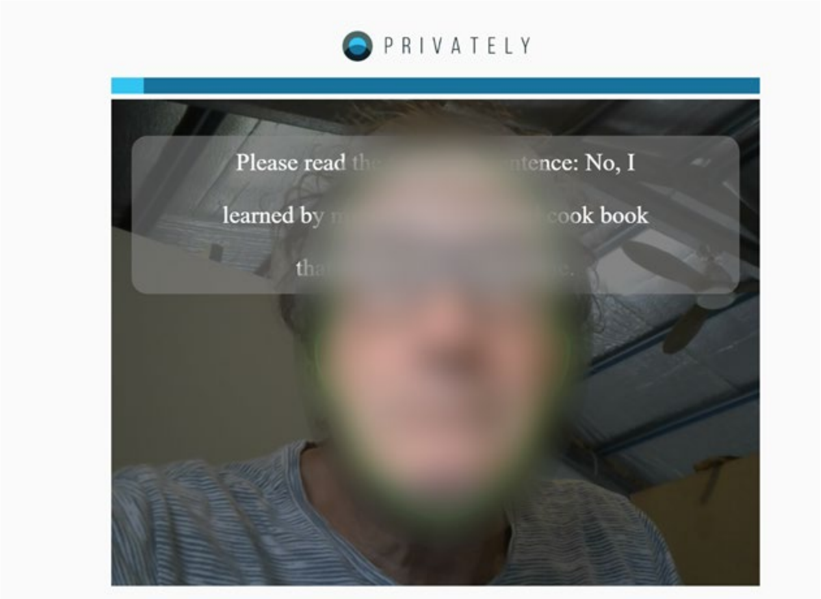


Welcome

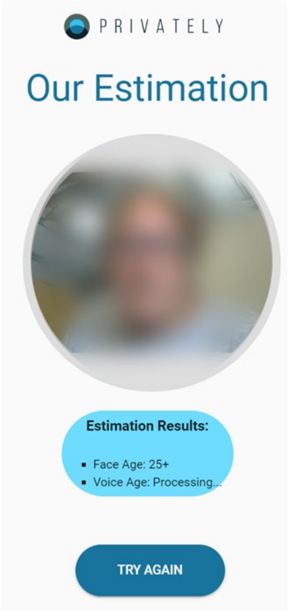
Our multi-modal age verification model estimates a user's age based on a short audio-video recording. Note that anit-spoof measures are not included in this demo. Please press the button below to start the camera. You will be asked to correctly position your face, and then read a sentence.

Estimate my age using video

You will be asked to enable your camera and microphone. Please note that your privacy and security is vitally important to us. Data or media is collected or stored only for auditing purposes. For more information on how we process data, please refer to our privacy policy



Please read th ntence: No, I
learned by m cook book
tha



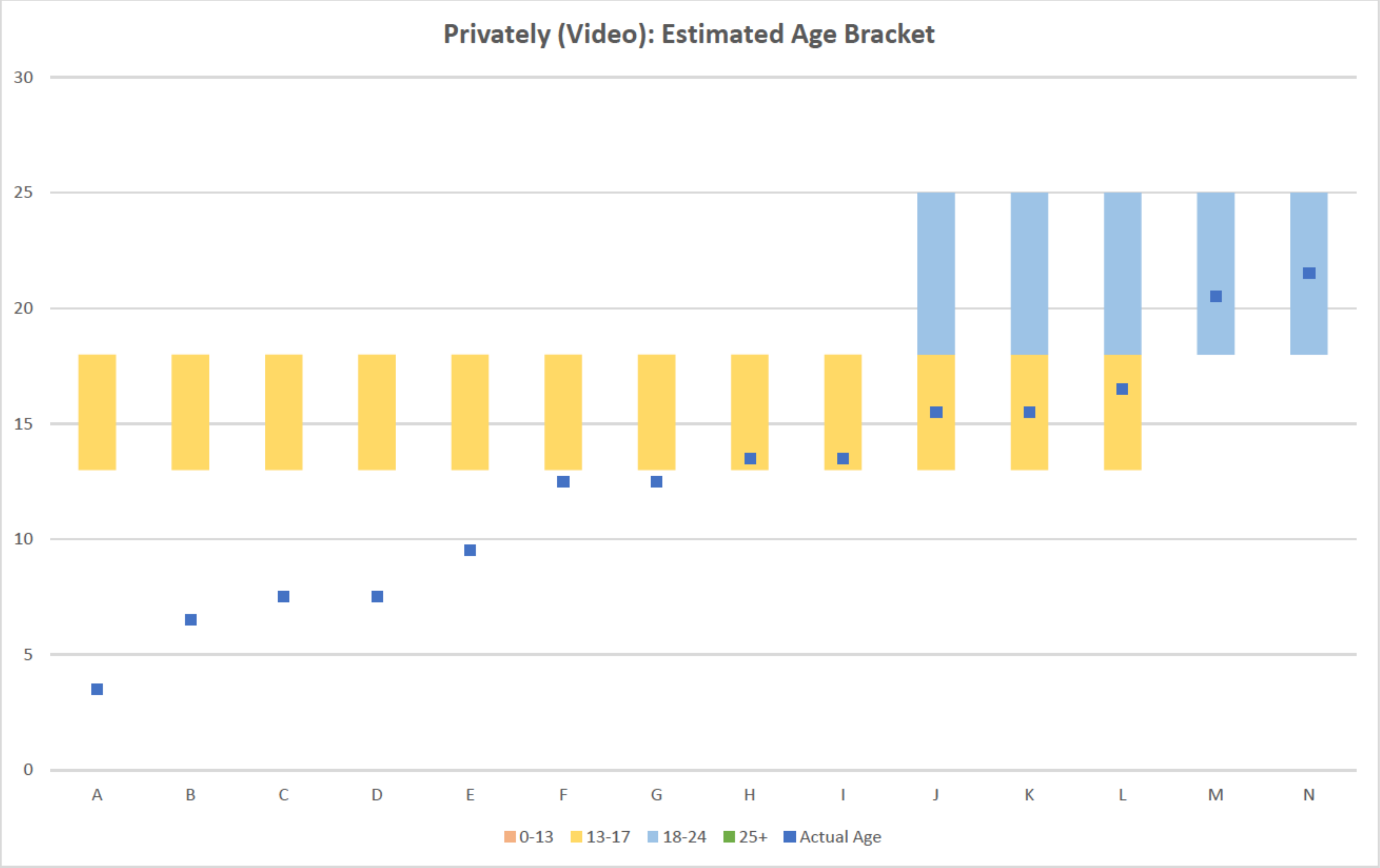
Our Estimation

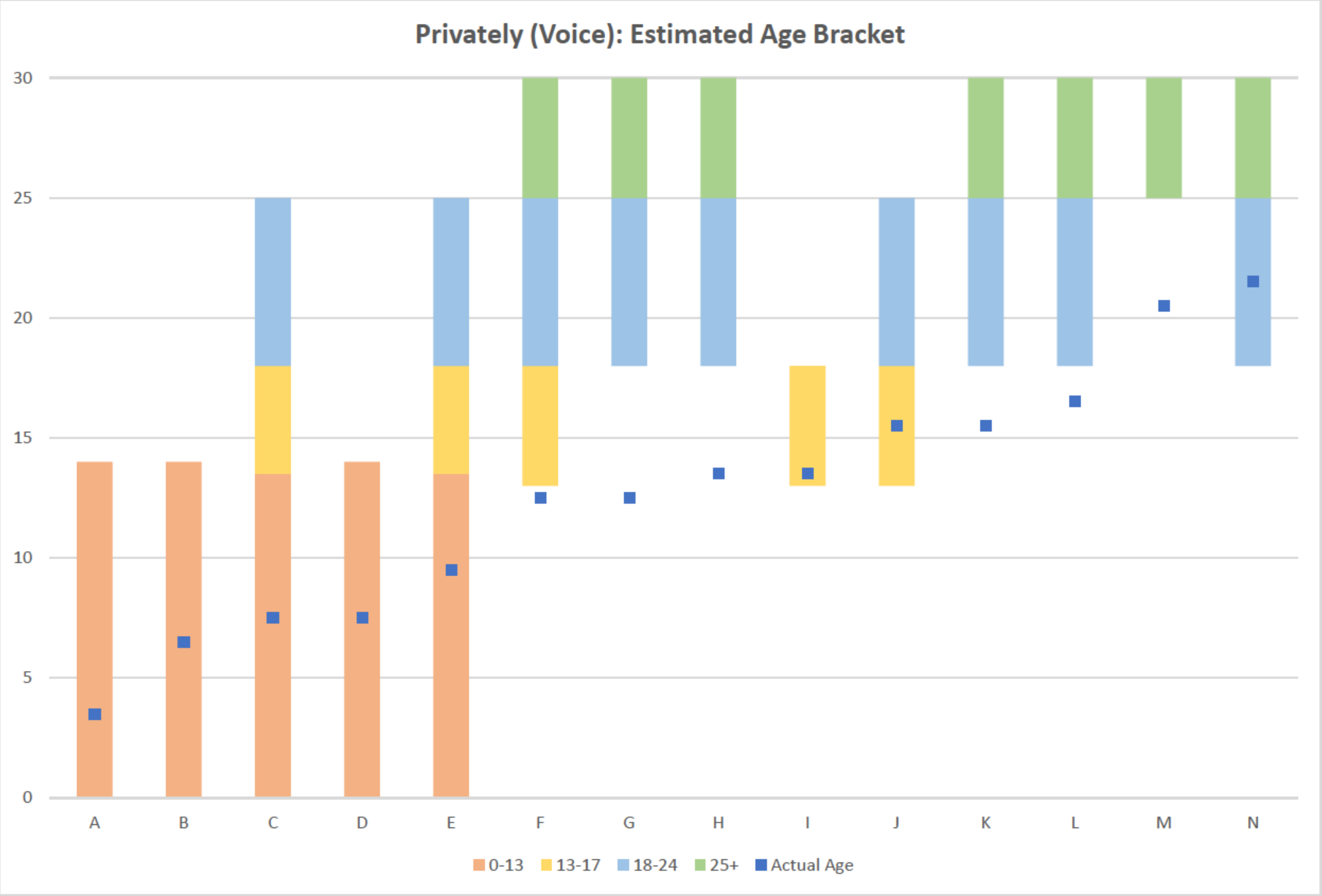
Estimation Results:

- Face Age: 25+
- Voice Age: Processing...

Privately test results

First name	Current age	Video	Voice	Video	Voice	Video	Voice	Video	Voice	Video	Voice
A	3	13–17	0–13	13–17	0–13	13–17	0–13	13–17	0–13	13–17	0–13
B	6	13–17	0–13	13–17	0–13	13–17	0–13	13–17	0–13	13–17	0–13
C	7	13–17	18–24	13–17	0–13*	13–17	18–24	13–17	18–24	13–17	**
D	7	13–17	0–13	13–17	0–13	13–17	0–13	13–17	0–13	13–17	0–13
E	9	13–17	13–17	13–17	13–17	13–17	0–13	13–17	18–24	13–17	0–13
F	12	13–17	25+	13–17	18–24	13–17	18–24	13–17	13–17	13–17	18–24
G	12	13–17	18–24	13–17	25+	13–17	25+	13–17	25+	13–17	25+
H	13	13–17	18–24	13–17	18–24	13–17	25+	13–17	18–24	13–17	18–24
I	13	13–17	13–17	13–17	13–17	13–17	13–17	13–17	13–17	13–17	13–17
J	15	13–17	18–24	18–24	13–17	18–24	13–17	18–24	13–17	18–24	13–17
K	15	18–24	18–24	18–24	18–24	13–17	25+	13–17	25+	18–24	25+
L	16	18–24	18–24	18–24	18–24	18–24	25+	18–24	25+	13–17	25+
M	20	18–24	25+	18–24	25+	18–24	25+	18–24	25+	18–24	25+
N	21	18–24	25+	18–24	18–24	18–24	25+	18–24	18–24	18–24	25+


Privately (Video): Estimated Age Bracket
30
25
20
15
10
5
0
A
B
C
D
E
F
G
H
I
J
K
L
M
N
0-13
13-17
18-24
25+
Actual Age


Privately (Voice): Estimated Age Bracket
30
25
20
15
10
5
0
A
B
C
D
E
F
G
H
I
J
K
L
M
N
0-13
13-17
18-24
25+
Actual Age

Assessment of Privately

Design	
Item	**Assessment**
Feasibility – is the technology ready to be rolled out?	For the production version, users download a web app and all processing is carried out via the browser on the device. We did not have a chance to evaluate the effectiveness of the on-device processing; its performance could depend on the device's processing power.[38]
Extent and sensitivity of the data required for the technology to operate	Privately claim no voice or video data leaves the user's device in their implementation.
Security and technical integrity of the technology	Security is reliant on the user's phone setup. Accuracy of the age estimation system needs to improve.
Barriers to inclusion- what identification is required for verification?	Voice assessment requires the ability to read and a language the user can accurate speak being supported.
Potential for bias and accessibility	Unknown, too small a sample size. For example, we would want to test a more representative sample that assessed biometric accuracy performs for those who identified as men, women or nonbinary and for those for whom English was not a first language.
Accreditation	AVPA, Age Check Certification – Challenge 25, KJM
Cost	Unknown.

Implementation	
Item	**Assessment**
Transparency over decision making, use and appeals, appropriate oversight, and governance	Privately provides no transparency over the decision making leading to the age range given. The service implementing Privately can allow another form of age assurance if facial or voice analysis do not grant access.
Can the technology be implemented flexibly for different business models?	Privately is an on-device service and may require a relatively powerful device for facial or voice analysis. This could limit the use cases for specific applications.
Can unintended consequences of the technology design be mitigated through the way it is implemented?	This depends on the service implementing age verification and what it allows. Privately has a few options and the service can decide which ones and how many to implement to best suit its needs.

[38] Privately is a member of AVPA

AGEify

AGEify did not respond to Enex TestLab's requests for product details or a means to test the product. We have produced a basic assessment. AGEify state on their website that their services include, facial analysis, ID document scanning, credit card validation, and human verification (in-person or via video call)

Assessment of AGEify

Design

Item	Assessment
Feasibility – is the technology ready to be rolled out?	Enex Test lab cannot comment on feasibility of AGEify – the vendor did not respond to our request to assess its technology.[39]
Extent and sensitivity of the data required for the technology to operate	AGEify's website claims that once it processes an image or ID document, it deletes the images and creates a random untraceable identifier that it assigns to the user for sharing their age or passing an age gate.
Security and technical integrity of the technology	AGEify's website claims it complies with GDPR (but does not mention other international standards) and provides 100% anonymity. It says digital copies of ID documents are deleted after processing.
Barriers to inclusion- what identification is required for verification?	None significant for facial age estimation. If government ID documents are required, this poses the usual barriers to full inclusion.
Potential for bias and accessibility	Unknown. For example, we would want to test a more representative sample that assessed biometric accuracy performs for those who identified as men, women or nonbinary and those where English is not a first language.
Accreditations	AVPA
Cost	Limited free trial available and €0.05 per credit tier without limits available. Definition of credits was unclear.

Implementation

Item	Assessment
Transparency over decision making, use and appeals, appropriate oversight, and governance	AGEify provides no transparency over facial analysis. It can perform secondary verification with government identity documents.
Can the technology be implemented flexibly for different business models?	The technology is implemented in the form of a QR code that users must scan to be authenticated and proceed past the checkpoint.
Can unintended consequences of the technology design be mitigated through the way it is implemented?	Users may not be able to scan QR codes on some devices – this is the only access method currently shown on AGEify's website. Multiple forms of ID are accepted for age verification.

[39] AGEify is a member of AVPA

AgeChecked

AgeChecked had an initial short video call with Enex staff and agreed to provide information on accessing the product for testing. However, this was not forthcoming within the timeframe of the project. We have produced a basic assessment.

AgeChecked claims to offer a variety of age verification packages to allow age verification of customers.

Assessment of AgeChecked

Design	
Item	**Assessment**
Feasibility – is the technology ready to be rolled out?	AgeChecked technology is actively used by businesses dealing with adult products. It's available for use on Shopify and WooCommerce platforms.
Extent and sensitivity of the data required for the technology to operate	AgeChecked's website claims it deletes any collected identification documents from its database once it has verified a user's age.
Security and technical integrity of the technology	AgeChecked's website claims it is compliant with GDPR, BSI and PAS 1296.
Barriers to inclusion- what identification is required for verification?	AgeChecked relies on government identity documents, so could exclude users without identity documents.
Potential for bias and accessibility	Unknown. For example, we would want to test a more representative sample that assessed biometric accuracy performs for those who identified as men, women or nonbinary and for those for whom English was not a first language.
Accreditation	AVPA, Age Check Certification - PAS1296,
Cost	From £35 per month for 100 checks.

Implementation	
Item	**Assessment**
Transparency over decision making, use and appeals, appropriate oversight, and governance	We could not assess implementation.
Can the technology be implemented flexibly for different business models?	
Can unintended consequences of the technology design be mitigated through the way it is implemented?	

Mastercard ID

Mastercard ID is a platform that allows users to upload hard identifier documentation once to verify their identity and age at the same time then re-use this identity across multiple other platforms within Mastercard's digital wallet which supports Mastercard ID. The information collected may include government IDs, payslips and date of birth.

Mastercard ID is currently accredited in Australia under the Trusted Digital Identity Framework (TDIF) scheme since 21 July 2022. Telecommunications company Optus uses it for identification checking [redacted]

Assessment of Mastercard ID

Design	
Item	**Assessment**
Feasibility – is the technology ready to be rolled out?	Technology is rolled out in multiple markets globally. Limited client feedback prevents us drawing precise conclusions.
Extent and sensitivity of the data required for the technology to operate	Mastercard ID requires the user to upload and verify whichever ID documents they would like to be re-used with the service.
Security and technical integrity of the technology	The Mastercard ID website claims the service is "accredited under Trusted Digital Identity Framework (TDIF) as an Identity Provider (IP1+), Credential Provider (CL2), and Identity Exchange." It claims all personal data and ID documents are stored with encryption. Enex TestLab has not tested this claim.
Barriers to inclusion- what identification is required for verification?	No obvious barriers to inclusion other than lack of sufficient identification documents for the required identity check.
Potential for bias and accessibility	To use the service, you must have ID documents and the ability to upload and verify them
Accreditation	TDIF (Identity Provider, Identity Exchange, Credential Provider)
Cost	Unknown.

Implementation	
Item	**Assessment**
Transparency over decision making, use and appeals, appropriate oversight, and governance	Mastercard ID provides no transparency over decision making. Users may supply other forms of identification to ensure they can access resources that use Mastercard ID for verification.
Can the technology be implemented flexibly for different business models?	Mastercard ID removes the burden of storing PII and verifying ID documents from website providers. It allows users to provide the least amount of information necessary to confirm their identity.
Can unintended consequences of the technology design be mitigated through the way it is implemented?	Technology has been accredited in multiple categories. It can verify many forms of ID documents and PII.

Internet content filters

The range of internet content filters on the market offer a variety of options including:

- Blocking according to age only – the filter determines the types of content blocked for each age group.
- Block according to category only – parents or administrators determine which categories are appropriate for their users.
- A combination where the selected age determines which categories of material to block but the parent or administrator can adjust these.

Some filters allow fine tuning with the use of blocklists to prevent access to specific sites that aren't blocked by the selected filter and allow lists to enable access to URLs that were blocked under the chosen filter. Some allow children to request access to sites blocked by the filter, which a parent must approve. This gives additional opportunity for parents and children to exercise their respective rights. Some parents may find difficulty with setting up or managing the ICF in a day-to-day setting and adapting the blocking categories based on the evolving needs of growing children.

Enex has a wide range of control URLs tested to ensure excessive blocking is not occurring and these include websites that are age appropriate for children, government services, LGBTQIA+ organisations and resources, etc.

Filters that provide reasons for blocking a particular URL give parents an opportunity to better understand the operation of the filter and adjust the settings to suit their parental and child needs.

Remember that filter vendors' categorisation of URLs can be highly subjective – it's often based on the use of keyword searches or AI assessment of images and words. Also, the content of a URL may change significantly after a vendor has classified it. The use of RTA metatag could help ICF's with classification of URLs once a larger majority of websites implement it.

Filter vendors can't possibly classify every URL – they are huge in number and constantly changing. However, filters can be set to block or allow unknown URLs – typically the default is to block, especially for younger users. Parents should have the opportunity to specify how their filter manages unknown URLs.

ASUS ZenWiFi Pro XT12

The ASUS ZenWiFi Pro is a router with ICF functionality provided by Trend Micro. The router requires an iOS or Android device to set up. Once the hardware is configured, parental controls can be set up and adjusted by mobile device or a PC on the network.

Under parental controls, the web filter provides four category choices, each with at least two subcategories:

- Adult and mature content (presumably over 18) and sites that contain material of a sexual, violent or illegal nature
- Instant messaging and communication
- Peer to peer and file transfer services
- Streaming and entertainment sites.

The router can also manage a user-specified deny or allow list but not both, with a maximum of 64 sites. This item is configured under the firewall settings as URL Filter.

Time scheduling can be configured under parental controls, with start and end times for weekdays and weekends or individual days in 15-minute increments.

This device did not fare as well as the other ICF software and devices tested. It blocked:

- 3.6% of URLs on the control list
- 58.2% of URLs on Enex TestLab's RX list – only 32.1% of the racism/hate category but 76.9% of the pornographic category
- 94.2% of URLs on the Office of eSafety Commissioner (OeSC) Prohibited URL Filter (PUF) list; 58 of 1000 tested URLs were visible – this is below the level required for family friendly filter certification.

McAfee Safe Family

McAfee Safe Family app is available on Windows, Android and iOS. Features vary depending on the platform. The filter is managed by installing the app on a parent's device (when installing the app on each device, the user indicates whether it is a parent's or child's device.) The installation process is guided with the installation wizard, upon completions set up the child's age profile and select any extra categories to block or allow.

As well as filtering web content, the app can track the location of a device (and thus the child) and when and what usage is occurring on the device.

This filter blocked:

- 7.7% of URLs on the control list
- 81.8% of URLs on Enex TestLab's RX list – 62.6% of the racism/hate category and 100% of the pornography category
- 99.2% of URLs on the OeSC PUF list; 8 of 1000 tested URLs were visible – this is above the level required for FFF certification.

McAfee Secure Home Platform

McAfee Secure Home Platform is a router-based filter. It requires an iOS or Android device to complete setup. The setup process involves setting the child's age and blocked categories and assigning these to the child's devices.

To increase protection, parents can also install a secondary application named McAfee Web Advisor on Windows devices. McAfee Web Advisor is a browser extension that works with the Secure Home Platform router to ensure sites are blocked. Enex TestLab tested with a Secure Home Platform router and Web Advisor installed.

As we noted before, Secure Home Platform can't prevent access to content from a device using a mobile network connection. It also appeared blocking was not effective on a Windows device with McAfee Web Advisor installed when it was using a cellular connection. We concluded McAfee Web Advisor only works when connected through the ICF on the router.

This filter blocked:

- 6.8% of URLs on the control list.
- 76.9% of URLs on Enex TestLab's RX list – 61.3% of the drug advocacy category and 97.3% of the pornography category. We suspect these results could be improved by tightening the filtering criteria, but this would likely increase the number of innocuous URLs blocked.
- 97.8% of URLs on the OeSC PUF list; 22 of 1000 tested URLs were visible – this is above the level required for FFF certification.

Norton Family

Norton Family is a software filter. To set it up, a parent registers online and then downloads the software from a link provided. The straightforward parental control setup process involves establishing the child's profile, blocked categories and time restrictions. Enex TestLab evaluated Norton Family with the default installation and blocked categories.

Norton Family is available for Windows PCs and iOS and Android mobile devices. We did not test any mobile devices. We didn't evaluate the browser extension because it might not be installed by all users and could be easy for a child to circumvent by installing another browser.

This filter blocked:

- 6.4% of URLs on the control list.
- 81.1% of URLs on Enex TestLab's RX list – 56% of the racism/hate category and 100% of the pornography category. We suspect these results could be improved by tightening the filtering criteria, but this would likely increase the number of innocuous URLs blocked.
- 97.1% of URLs on the OeSC PUF list; 28 of 1000 tested URLs were visible – this is above the level required for FFF certification.

Internet content filter product features and test results

Product	Safe Family	Secure Home Platform	Norton Family	ZenWiFi Pro XT12
Vendor	McAfee	McAfee	NortonLifeLock	ASUS
Version	2.9.3.144	3.24	3.8.6.41	3.0.0.4.388_21678
Type	Application	Router	Application	Router
Compatibility	Windows, Android, Apple iOS	Requires Android or iOS device for configuration	Windows, Android, Apple iOS	Requires Android or iOS device for configuration
Cost	$49.99/year	Price varies per router	$54.99/year	$1249[40]
Lists				
Predefined	16	15	47	4
Allow	Yes	Yes	Yes	Partial
Block	Yes	Yes	Yes	Partial
Features				
Limit by total hours	Screen time can be defined as a range of time on specified days. Separate schedules can be set for weekdays and weekends.	Screen time can be defined as a range of time on specified days. Separate schedules can be set for weekdays and weekends.	Screen time can be defined as a range of durations on specified days. Separate schedules can be set for weekdays and weekends.	Screen time can be defined as a range of durations on specified days. Separate schedules can be set for weekdays and weekends.
Limit by schedule	No	No	Yes	No
Configure by age	Yes	Yes	Yes	No
Multiple user profiles	Yes	Yes – one per device	Yes	Yes, one per device
Override block	Child can request access	Child can request access	Yes, with parental approval	Yes, with parental approval
Warn option	No	No	Yes	No
Reasons for blocking provided	No	Partial	Yes	Yes
Classification results				
% blocked – control list	7.7%	6.8%	6.4%	3.6%
% blocked – RX list	81.8%	76.9%	81.1%	58.2%
% blocked OeSC PUF list	99.2%	97.8%	97.1%	94.2%
Filter rating[41]	**Class 3**	**Class 2**	**Class 3**	**Fail**

[40] JB Hi-Fi, Asus ZenWiFi Pro XT12 Tri Band Wi-Fi 6 Mesh System [2 Pack], accessed on 20 December 2022, no ongoing fee for ICF

[41] See Appendix B: Family friendly filter test methodology on page 63

Assessment of software internet content filters

Design	
Item	**Assessment**
Feasibility – is the technology ready to be rolled out?	Mature products; technology has been in service for over a decade and many vendors supply and support ICFs.
Extent and sensitivity of the data required for the technology to operate	No personal ID data is shared.
Security and technical integrity of the technology	Savvy children can bypass filters in many ways including: • Alternative browsers • VPNs • Writing code or scripts • Alternative user profiles or devices • Alternative network or DNS settings While each technology may be subject to differing circumvention techniques, the same process may or may not work on each filter, or even every version of the filter.
Barriers to inclusion- what identification is required for verification?	Cost, parental computer literacy, etc.
Potential for bias and accessibility	None, however parental supervision options will work better for families that have strong levels of communication and trust.
Cost	Moderate cost per device should be within most household budgets. See above table for individual pricing.

Assessment of hardware internet content filters

Design	
Item	**Assessment**
Feasibility – is the technology ready to be rolled out?	Dedicated filter routers are now very good but not as mature as software filters. Many new Wi-Fi routers include parental controls, but their blocking definitions are typically not granular or broad enough to be useful.
Extent and sensitivity of the data required for the technology to operate	No personal ID data is shared.
Security and technical integrity of the technology	A savvy child may be able to VPN through the device or connect to another internet source such as a mobile network, bypassing the hardware. Dedicated filter routers are generally more secure than generic Wi-Fi routers and the latter can have very poor filtering capabilities.
Barriers to inclusion- what identification is required for verification?	Cost, parental computer literacy, etc.
Potential for bias and accessibility	None.
Cost	Relatively expensive. For the average household, not an attractive option compared to a far cheaper software solution. See above table for individual pricing.

Spectrum Labs

Spectrum Labs is an artificial intelligence content moderation service. We could not test Spectrum Labs product because it needs to be installed in a website ecosystem and trained to work with the characteristics of the site it's employed to monitor.

Enex TestLab's technical interviewed three clients of Spectrum Labs who use it to help moderate their online ecosystems.

Spectrum Labs can be configured with a large range of AI modules to detect and moderate against behaviours including: bullying & harassment, hate speech, insults, profanity, radicalization, self-harm, sexual content, solicitation of sex, spam, threats, underage and child sexual abuse material and violence.

IMVU (Jeff Hanlon)

IMVU is a social media chat site that uses 3D graphics and personal avatars to increase participants' involvement.

The site has three levels of participation, each with greater levels of security. Each level exists in the same virtual world but at the lower levels, participants can't enter levels or locations where they would be exposed to more adult content. The levels are:

- The *general* audience level allows participation of anyone over age 13 – adults and children may interact.
- The *adult* level allows more freedom of expression, but chats are monitored and can be blocked to ensure no explicit adult language is used. This level is protected by a simple honour-system age gate; the participant simply provides a birthday to show they're over 18.
- The *access pass adult* level requires age verification in the form of a government ID to register and gain access. Explicit adult language is permitted in this level.

IMVU uses 12 Spectrum Labs contextual modules to actively monitor the site, including the age verification, hate speech, bullying and explicit sex modules.

IMVU initially approached Spectrum Labs to protect minors on their site but has expanded its use of the platform since. Before purchase, IMVU provided Spectrum Labs a random log of chat streams to run a health check. Spectrum Labs returned a list of potential underage participants and examples of bullying. IMVU was impressed with the results and purchased the product.

When the Spectrum Labs modules detect a problem, they send the information to human moderators to vet and act upon.

Hanlon claims the age verification module performed extremely well straight out of the gate while Spectrum Labs had to tweak other modules to better fit the IMVU audience. Currently the modules run at an accuracy of greater than 80% with some, for example hate speech, approaching 100%. Accuracy is measured by human moderators assessing problematic content flagged by the AI. It doesn't measure false negatives – problematic content the AI didn't find.

On average, the system identifies 5–10 minors per week who have stolen an adult's ID and joined the explicit access pass adult level. Moderators immediately remove their membership. In the US, Spectrum Labs provides an appeal mechanism for adults incorrectly categorised as minors to prove their age.

The AI vets all chats in real time – inappropriate material is assessed and stopped in a few milliseconds – and there are around 6 million active users on the site. If the AI processing ever slowed to 1 second, the chat will be released unprocessed, but Hanlon says this has never happened.

Spectrum Labs is currently developing modules for severe toxic speech and severe insulting speech – these expand on the hate speech module which takes a black-and-white interpretation of speech. One of the difficulties with training these modules stems from the diverse community using the site. What is a terrible insult to one ethnic group will be commonly used and acceptable to another; it can also depend on the context the phrase is used in.

IMVU uses the child sexual abuse material module to immediately flag an adult talking to a minor in a sexually explicit way. It's also training the module to detect child grooming – a far more difficult task that involves correctly interpreting long discussions.

Hanlon says the moderator tool for viewing and moderating AI hits is "pretty clumsy". He notes the AI is only used for punishing negative behaviour. As a member of the Spectrum Labs community, he has suggested modules that reward positive behaviour in the online community.

Wildlife Studios (Aoife McGuinness)

Wildlife Studios is an online gaming platform with games that cater for young children up to adults. McGuinness wrote the Wildlife Studios gaming protection scheme from scratch.

Wildlife Studios uses two systems to help monitor the gaming chats: Spectrum Labs and Hive. McGuinness says Spectrum Labs offers "extraordinary" support compared to Hive. For example, when Wildlife Studios needed to provide a submission to the UK Government for regulatory purposes, Spectrum Labs provided a great deal of material to support the submission.

The Spectrum Labs system is configured to provide three levels of oversight for the online gaming platform: relatively low for adults-only games; moderate for mixed adults' and children's games; and strict for children's games. Wildlife Studios uses a graduated ban method: a warning for first offence, a timed suspension for a second offence and deregistration for a third offence. (Serious transgressions such as child grooming result in an immediate ban.)

McGuinness says the Spectrum Labs technology has become very accurate after tweaking – most models now achieve over 75% accuracy, which has reduced the workload on moderators. With the child sexual abuse material module, Wildlife Studios errs on the side of caution, preferring to have more false positives to ensure fewer real cases slip through.

Grindr (Alice Hunsberger)

Online dating site Grindr has had similar overall experiences to IMVU and Wildlife Studios. However, the notable exceptions reinforce that one size does not fit all. Depending on the type of site, a differently tweaked combination of modules may be needed to provide satisfactory results.

As an example, the age detection module was initially flagging up to 50% false positives. Grindr's moderators found the reason was the way participants constructed their chat messages. Typically, the age detection module flags numbers less than 18 – a message such as 'I'm 12' would be flagged as an age admission. However, in the abbreviated style of Grindr chats, 'I'm 12' could be a distance to a location or a claim about the size of a user's anatomy. Spectrum Labs tweaked the module to examine comments more rigorously and ascertain the contextual meaning of the number. This has reduced false positives to around 25%.

While the other two companies used in-house moderators, Grindr contracts around 150 outsourced moderators and has created its own custom interface tuned to its needs.

Assessment of Spectrum Labs

Design

Item	Assessment
Feasibility – is the technology ready to be rolled out?	The product is successfully deployed in the marketplace. User feedback shows it provides very good results.
Extent and sensitivity of the data required for the technology to operate	The system uses machine learning alone – no personal ID is stored or needed, and the modules appear to flag unacceptable behaviour and not store any other data. Human moderators only view system infraction data provided by the AI. Spectrum Labs is developing a personally identifiable information module to protect the anonymity of individuals in chat rooms and to intercept identity data before it is passed on.
Security and technical integrity of the technology	The Spectrum Labs AI software is cloud based. Enex TestLab has not investigated how secure this is. SOC2 certified and GDPR/CCPA compliant.
Barriers to inclusion- what identification is required for verification?	None obvious.
Potential for bias and accessibility	Some of the experiences we discussed with Spectrum Labs clients showed that the system could develop or reflect biases. However, Spectrum Labs can tweak the AI to work around these issues.
Cost	From the interviews, we believe Spectrum Labs has a substantial cost. One interviewee said it would have been impossible to use the product when the company was a small start-up; only a reasonably sized company with a strong paid subscriber base could afford Spectrum Labs.

Implementation

Item	Assessment
Transparency over decision making, use and appeals, appropriate oversight, and governance	Spectrum Labs does not provide transparency over the decisions its AI makes but each service can have human moderation teams to validate the AI's decisions.
Can the technology be implemented flexibly for different business models?	The technology is modular and can be implemented with whichever modules meet a service's needs.
Can unintended consequences of the technology design be mitigated through the way it is implemented?	As each client decides how to use the service, this report can't easily discuss the unintended consequences.

Phase 2 d) i) Analysis: Development of the AV industry – positive developments

Very few of the available age verification technologies have undergone rigorous independent testing to determine their accuracy or bias. There are some exceptions; Yoti has undergone in-depth testing by an accreditation lab. The absence of testing could stem from a shortage of relevant resources to conduct tests and the cost of obtaining tests.

We feel that facial age assessment is a front runner – it provides a reasonably accurate assessment but maintains user privacy because it doesn't permanently store unique personal data.

Contextual filtering by AI is another strong contender. While it's not an accurate method for initial age gating, this technology is getting better at identifying underage children in adult sites by analysing chat streams. Furthermore, the technology can filter out a range of undesirable online behaviours such as bullying, hate speech and child grooming. This technology can monitor adults' and children's sites to ensure a safe environment. Based on interviews with clients of Spectrum Labs, there are good indications that this technology is accurate.

Internet content filtering is a very mature technology. Through our experience testing many ICF products for inclusion in the Family Friendly Filter scheme, we recognise the high accuracy of the latest software filters. Hardware based ICFs have lagged their software counterparts until relatively recently. However, one of the hardware filters we tested for this report performed on par with best software counterparts. Unfortunately, smart children – especially older ones – can circumvent ICFs but the majority of children will be protected from unsafe sites.

An internationally defined age token that can be linked to various age verification technologies would be extremely helpful. The most accurate way of creating reliable tokens is to use trusted documents such as birth certificates, passports and other photo IDs to verify a person's age. However, once the token is created, it's essential to purge any records of the ID documents to prevent traceability back to an individual. An age token record should be stored with a unique identifier which, once presented to an age gate, determines if the age gate has been passed or not. It should release the least possible amount of personally identifiable information. Age tokens should be secured at a device level.

The UK Pass Card process seems like a good starting point to develop an evidence-based methodology for creating and processing digital age tokens, which ultimately will be more convenient that any physical card. The token can reside in your phone or your watch, for example. When combined with face recognition security processes available in most modern mobile devices, it becomes a simple and strong multifactor age gating system. This would be hard to circumvent, since a likeness image that could fool the mobile device's facial recognition would be needed to confirm the user's ownership of a token.

Phase 2 d) ii) Analysis: Development of the AV industry – challenges and gaps

The age verification industry and its associated technologies are relatively new and still evolving. This is an immature market globally with a small number of participants providing a diverse range of solutions to address the need. With this context, it is not a straightforward task to compare products head-to-head and provide an accuracy ranking.

Current published standards are not comprehensive and mainly involve high-level principles – they can't be applied to address domestic requirements.

Casting an Australian lens on the subject, any standards would need to be aligned with existing frameworks such as the Digital Transformation Agency's (DTA) Trusted Digital Identity Framework[42] and the Department of Social Services' (DSS) National Consumer Protection Framework for online wagering.[43] Such coordination would ensure fit-for-purpose coverage and enable age verification vendors to ensure their claims could be measured and verified for compliance.

We note that recommendation 1 (2.143) of the Australian Parliamentary Inquiry into age verification and online pornography (Protecting the age of innocence) was that DTA in consultation with Australian Cyber Security Centre (ACSC) develop standards for online age verification for age-restricted products and services in Australia.[44]

Australia lacks mandated age checking for high-risk sites. Buying alcohol online in most cases is an honour system that asks users to enter a birthday. To be blunt, these sites could afford to implement rigorous age gating systems. On the other hand, small or start-up companies could find rigorous age gating too expensive until they grew a substantial paying subscriber base. Then, of course, there is the proliferation of illegal adult sites that can be geographically relocated to dodge any regional mandates for age gating.

From an admittedly cursory inspection, voice age estimation is unreliable, certainly lagging behind the accuracy of facial analysis. Perhaps this is expected as almost all children's physical attributes age in a similar manner. The same is not true of linguistic skills, which can widely vary from child to child. A well-educated child could have the enunciation, grammar, conceptual understanding and vocabulary of an adult and it's easy to imagine an AI attributing a higher age. Of course, there are adults who could be mistaken for children by the same yardstick.

While facial age assessment technologies are typically more reliable, they often display biases such as being more accurate for males than females or for lighter skinned people more than those with darker skin. Closing this gap will take time and increased variety of the AI's training sets. For example, obtaining statistically valid results with facial age recognition would require many photos of people at various ages. It would also need a diverse training set to ensure it was relevant to the demographic of users who would rely on this technology. Given Australia's diverse population, it behoves us to ensure that any solution treats everyone without appreciable biases.

[42] Digital Transformation Agency, Trusted Digital Identity Framework

[43] Department of Social Services, National Consumer Protection Framework for online wagering

[44] Australian Parliamentary Inquiry into age verification and online pornography, recommendation 1 (2.143)

While internet content filters are adept at blocking inappropriate sites, they can be bypassed or circumvented by equally adept children.

It does not help that a large proportion of parents are not comfortable or skilled in IT and find it challenging to install and configure some ICF software or hardware products.[45] Many parents rely on their children to deal with technology problems and some grant administrator privileges to their kids so they can solve tech issues on computers, mobile devices and home appliances. Bypassing an ICF is a trivial problem for a child in these circumstances.

VPNs are a particularly problematic method used to circumvent ICFs. Even the high-end consumer smart modem we tested, the ASUS ZenWiFi Pro XT12, can’t reliably block VPNs as it does not have a robust packet filter.

[45] For example, Pyburne, P, Jolly, R, *Australian Governments and dilemmas in filtering the Internet: juggling freedoms against potential for harm*, NetAlert: foiled or bypassed?, Australian Parliamentary Library, 8 August 2014

Conclusion

The age verification industry and associated technologies are relatively new and still evolving. This is a low-maturity market, comprising a small number of participants globally with diverse range of technical solutions to address the challenges of age verification.

As such, no solution is bulletproof in identifying age or protecting children from all harmful online content.

Some technology solutions perform very well in their particular task. However, some, such as internet content filters, can be circumvented by *clever* children who are often more IT savvy than their parents. Parents have a critical role to play in protecting their children's online safety but often lack the technical expertise or the time to properly understand and mitigate the risks children face online. We don't have any easy answers.

Other solutions perform very well and are not easy to circumvent, for example Spectrum Labs' contextual AI. Unfortunately, these solutions can't cover all the bases when it comes to a child's online safety.

With the current state of technology, it still takes multiple solutions to adequately cover the majority of avenues.

The protection starts with a method for detecting or closely approximating a child's age. Current biometric solutions are quite good, but even the best are far from foolproof. To be fair, the limited testing carried out may have cast the products in a worse light than they deserve (or perhaps a better one). Only rigorous and exhaustive testing of these products, once mature, can deliver statistically valid conclusions.

Even if biometric products were accurate enough to verify age, there is a lack of industry standards to ensure seamless implementation.

Robust international standards could provide guidelines for online entertainment providers to understand and implement the required level of protection on their sites. Accurate, reliable and reproducible methods for assessing vendor product compliance and comparative baseline benchmarking is required, including a statistically significant independent (non-vendor/industry) qualified database of diverse (gender, ethnicity, age-range) age verification test data (image, video, voice).

Based on the results of the euCONSENT trials in Europe, any Australian trial of age verification technology would need to:

- Include a broad range of sites (such as alcohol sales or dating sites) to avoid stigmatising an individual category such as pornography
- Test a representative sample of Australians including multicultural, those who identified as men, women or nonbinary and those for whom English was not a first language
- Provide multiple options for proof of age to help users who don't have access to a particular method of verification
- Consider those whose live identity did not match the name or gender on their ID documents
- Maintain user privacy by ensuring AVP services could not pass user identity to website providers and website providers could not pass to AVPs what content users have accessed.

One thing in particular must be clearly defined in any standard – a universal age token. Such a token could be used and shared industry wide. It could reside in a user's device and provide age verification anywhere in the world.

Appendix A: NSW Government's Liquor Amendment (24-Hour Economy) Bill 2020

Recently, the NSW Government legislated new restrictions relating to sale of alcohol online with same day delivery.[46]

Previously it was required only that delivery drivers verify the age of a person upon delivery of the product. The new rules included a requirement for online age verification at the time of ordering.[47] The legislation came into effect on 1 July 2021.

Businesses reported having difficulty conforming with the new regulations.[48]

The law required suppliers to verify age using a commonwealth government–accredited digital identity service provider or an alternative artificial intelligence system to verify evidence of age documents.

However, a spokesman for Liquor and Gaming NSW said delivery businesses faced "unforeseen technical issues in meeting minimum age verification requirements using the [national framework], which were outside of the direct control of the licensees".

One problem involved matching purchasers' ID documents to the federal government's Document Verification Service. "This included significant processing times and/or large failure rates where the matching process was unable to be completed," the spokesman said.

As a result of these issues, the NSW Government added a third verification option on 1 June 2022. This option involved requiring the purchaser to provide their name and date of birth when they ordered same-day delivery and to state that were able and willing to produce, at the time of the same day delivery, an evidence-of-age document that verified their identity and age. This option is available until 31 May 2023.

The Liquor & Gaming NSW expects to consider the effectiveness of this legislation after 1 June 2023.

Clearly, any age verification method used to protect children in Australia will have to be assessed for accuracy, reliability and efficiency. Showing an identity document to a delivery person is hardly an option when giving a child access to online lessons or preventing them from accessing a pornography website.

[46] Liquor and Gaming NSW, Same day delivery age verification requirements, 2 September 2022
[47] Ioni Doherty, New laws for same day alcohol delivery, drinkstrade, 29 June 2021
[48] Michael Koziol, 'Technical issues' hobble age verification for same-day alcohol delivery, *Sydney Morning Herald*, 14 June 2022

Appendix B: Family friendly filter test methodology

Here are the details of the methodology Enex TestLab uses to test family friendly filters:

- Testing involves a minimum of 1500 URLs.
- Where possible, the filter under test is set to filter internet traffic on a PC running Windows 10 or 11.
- Where possible, we select vendor blocking categories that coincide with the undesirable categories found in the Enex RX and OeSC PUF lists. We deselect other categories.
- Where possible, we use an automated test script that attempts to access the URLs in each category (we do this to improve efficiency and to reduce Enex TestLab staff exposure to unpleasant material). We record which URLs returned the appropriate website, a filter block page or a page-not-found notice (including parking pages). Where the result is unclear, we manually assess the URL.
- When calculating the percentage of sites blocked, we exclude URLs that return an invalid site such as not found or unavailable. For example, if 18 URLs are visible, one is blocked, and one is not found, we report the block percentage as 5.3% (1/19) rather than 5% (1/20).
- Where possible, we use a script to compare returned URLs. This is particularly useful where the filter returns a URL with a distinctive phrase such as 'Blocked by Filter X' – the tester immediately knows that a successful block has taken place.
- Unfortunately, filters employ such a variety of filtering methods and methods for reporting blocks, we often need to manually check URLs not reported as being blocked to determine if they were visible and valid, not found, redirected to a parking page, blocked or sufficiently changed that they were not appropriately placed in the current category. (We don't pass judgement on the categorisation of OeSC URLs – only those collected by Enex TestLab.)
- HTTPS URLs can be a problem – some filters fail to block these URLs or block them but only return an error such as 'secure connection failed' (as often seen under Firefox).

Classification (scoring)

Here are details of the methodology we use to classify filters:

- Any filter which does not block at least 97% of the OeSC PUF automatically fails and does not receive classification.
- Classifications are determined by comparing the rate of blocked URLs on the Enex TestLab RX list and the blocked URLs on the control list (regular websites that should not be blocked). Where a greater percentage of RX material is blocked, a proportionally larger amount of blocking of control material is allowed.
- Where a product adequately blocks the OeSC PUF list but no other rating can be granted, the product receives an Unclassified rating.

	RX Minimum	Control Maximum	PUF Minimum
Unclassified	0%	100%	97%
Class 1	55%	5%	97%
Class 2	70%	10%	97%
Class 3	80%	20%	97%

- Unclassified: These filters block websites on the eSafety Commissioner's Prohibited URL Filter (PUF) list, and are recommended for 18+ years of age

- Class 1: Recommended for children over 15 years of age
- Class 2: Recommended for children between 10 and 15 years of age
- Class 3: Recommended for children under 10 years of age

Appendix 9: Response to independent assessment of age assurance and safety technologies (Yoti)

Attachment

1.Sample Size

We would recommend that testing/benchmarking should only be undertaken with a statistically significant sample size. Using one or test subjects for each year of age; is highly unusual - given that there is no way to ensure that the person selected is representative of an average person of that year of age.

We suggest that liaison with NIST could be instructive going forward, as their facial age estimation benchmark will open in the coming months and should be at scale.

In terms of the questions in the report; please see our comments to the points in blue below:

2. Feasibility – is the technology ready to be rolled out

The report does not mention that Yoti has undertaken over 600 million facial age estimates, for a wide range of leading global companies.

3. Client feedback

The report erroneously states, based on its view of a demo site, rather than a production site, that the Yoti solution 'has limited client feedback to draw any conclusions'.

All Yoti clients receive an age estimate, and the uncertainty score, so that they can decide if that is adequate for their purposes; e.g. if the image quality presented or lighting was sufficient.

> To quote client SuperAwesome,
>
> SuperAwesome has integrated Yoti's facial age estimation as one of the available age verification methods for parents going through the VPC (verifiable parental consent) process in certain countries.
>
> "SuperAwesome is excited to be working with Yoti to provide parents with a safe and trusted way to prove their age during the parent verification process. Given not everyone owns or has access to an ID or credit card, Yoti makes KWS more inclusive. In the time since integrating Yoti into KWS, over 60% of parents are choosing facial age estimation in the countries where it is available," said **Paul Nunn, Chief Strategy Officer at SuperAwesome**.
>
> To quote Only Fans
>
> Assuring minors cannot access our platform as fans or creators is central to our business strategy. We are very proud to work with Yoti to deliver age assurance for UK based fans on OnlyFans. As a #privacy nerd getting the balance right between ensuring those using our site are 18+ and protecting user privacy is essential. This is key to protecting children online. Our teams have thoroughly enjoyed working with the tech, legal and strategy teams at Yoti to implement this solution. This is a great example of what can be done when responsible businesses focus on solving the problems rather than saying "it's all too hard".

> *"Ensuring our users are over 18 is a priority for OnlyFans and an important element of protecting our community. We work with Yoti because their market leading age assurance technology provides the right balance between accurately assessing users' ages and respecting their privacy.."*
>
> **Keily Blair,** Chief Strategy and Operations officer at OnlyFans

4. Image erasure

The report states 'Once an image is processed, Yoti claims to erase the image'. To clarify, this is audited independently within Yoti's PAS 1296:2108 alongside its SOC2 audit.

5. Accuracy

The report states, 'Accuracy is not perfect, and some children may be able to pass unacceptable age gates'. We would clarify that where it is illegal to access a good or service under a certain age, it is advised that an age buffer is used. The key differentiator with this technology is that it is inclusive for people who do not own or have access to an identity document.

6. In terms of bias and accessibility we would add - that Yoti has undertaken an external bias review with external IEEE expert, Dr Allison Gardner. Research literature has revealed that humans assess age with 6-8 years of accuracy and age accuracy biases in facial age estimation also exist for in person age estimation e.g. humans ability to estimate age for women and men and for different ethnic groups brings individual bias.

7. Accreditations should detail- ACCS, FSM, KJM[37],

   Yoti's facial age estimation has been certified by the Age Check Certification Scheme for use in a Challenge 25 policy area. The ACCS report is available at: https://www.accscheme.com/registry.

8. In terms of cost - the cost of implementing this form of age gating for a website at low volumes can be around AU$0.40 per check; significantly lower at high monthly volumes. In cases where a business has millions of customers who may be anonymous / have no customer account, (such as shoppers buying age restricted goods at self checkouts or visitors at free adult content sites), these businesses can pay a minimum annual license fee where the cost per check can fall below AU$ 0.01

   There is also a minimum fee option for small adult sites which is £2k per annum, for up to 1 million repeat 'free visitor' age checks; ie. these are people who are window shopping and not setting up an account on an adult site. Yoti has also recently stated that checking of age via the Yoti app is delivered free of charge[14].

   (To qualify for our free digital ID an organisation must be integrated into our age portal with an annual minimum commitment and offer at least either facial age analysis (age estimation) or age from iD document (Auto doc inc. data extraction, face match, and liveness)

   Other age checks methods are available in certain geographies:

   - Credit Card
   - Mobile

[14] https://www.yoti.com/blog/free-digital-id-age-checks/

- Database
- e-ID

Pricing is on application and depends on countries and coverage / cover in the chosen jurisdictions; it also depends on whether a check is a one-off or an account based check. Yoti is working with very small organisations requiring low numbers of checks and very large global companies who avail of volume discounts.

In addition, Yoti can offer interoperable age tokens, with regulatory approval.

9. In terms of transparency over decision making, use and appeals, appropriate oversight, and governance - Yoti allows each company to set its required age or age range and to describe to its users the options available. Yoti may provide one or more of the options offered to the public by any company. The company will decide what their age requirement and uncertainty tolerance is. The company will also decide what support to offer to consumers. The service implementing Yoti can elect to allow another form of age assurance if facial analysis does not grant access or provide access to its customer support functions.

10. Decision making

Decision making is in the client's hands. Yoti provides the client with its assessment, but does not decide on if the user is granted or denied access to services. This is client's decision to make based on the full set of information and indicators that they have available to them

While Enex Testlab noted Yoti provides its clients with insights as to the decision making, it is Yoti's clients who elect to provide users with insight into the decision-making process leading to the age range given.

11. **Appeals process** **-** Yoti provides facial age estimation as a business to business SAAS (software as a service) service. Yoti acts as the data processor; it is the clients which offer a range of age assurance methods of which Yoti may provide one or several, potentially alongside homegrown methods and other providers' methods. The expectation is that the client using age assurance tools is the one that provides the appeal or alternative customer support services (as they elect) and in turn it is the responsibility of the client to offer an appeals process to their consumers. Some online services which have incorporated Yoti's age estimation tools provide a range of methods and fallback have provided users with alternative age assurance options where users believe Yoti has incorrectly estimated their age.[15]

12. Risk Management

Yoti's ISO 'Statement of Applicability' is available on request to interested third parties who are looking for more detail; typically these are provided. In Yoti's FAQs[16] details are given about relevant security and risk frameworks

(https://yoti.my.site.com/yotisupport/s/article/Has-Yoti-been-approved-by-any-regulators-andorganisations)

13. Different business models

[15] Epic Games, Parental Consent for Epic Accounts, Epic Games website, n.d.
[16] (https://yoti.my.site.com/yotisupport/s/article/Has-Yoti-been-approved-by-any-regulators-and-organisations)

Can the Yoti technology be implemented flexibly for different business models?

Yes. APIs, SDKs and assortment of developer documentation is available. Yoti offers a range of configurable age assurance optionshas solutions for all types of age verification. These can be integrated directly into an organisation’s flow. Alternatively an organisation can integrate Yoti’s Age Verification Service and access the full range of options. Companies have access to a portal where they can undertake A/B testing, implement different solutions in different geographies and configure the age buffers.

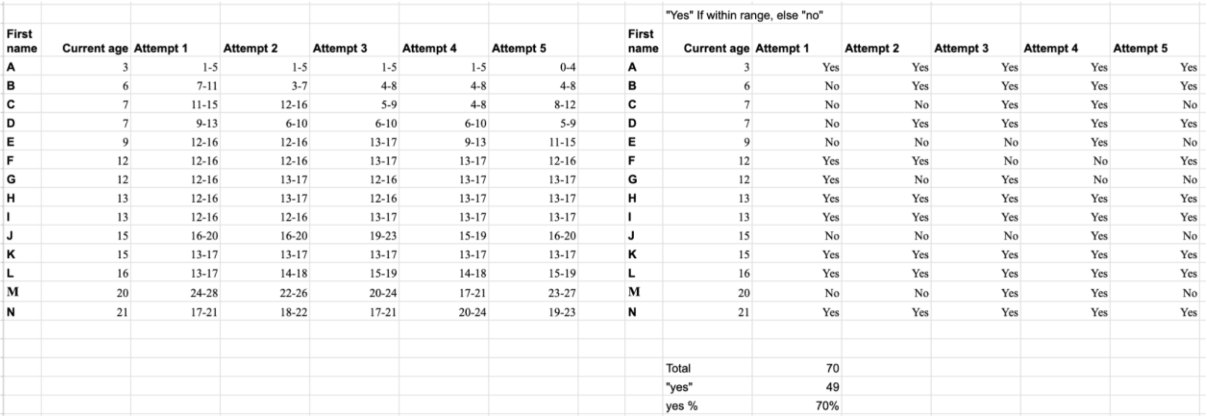
14. In terms of sensitivity of data - Yoti collects a scan of a user’s face using a camera. Yoti also states it deletes all biometric data as soon as the age check is conducted. Yoti claims that if you put the same face into the model several times, it would not be able to identify it as a face it has seen before and would provide **a new estimation each time.**

15. **The ICO sandbox exit report states,**[17] ‘

3.15 Having reconsidered our guidance in the context of our engagement with Yoti, we have concluded that the above guidance needs to be updated. This is because Yoti’s age estimation tool has demonstrated that it is, in some contexts, possible to use biometrics to make a decision about an individual or treat them differently without using that biometric data for the purpose of uniquely identifying that person. We have therefore concluded that it will be appropriate to revise the guidance above to make this clear.4.2 ‘ Yoti’s age estimation tool will not result in the processing of special category data.’

16. **True age fell within the estimated age range in 70% (49/70) of the attempts** In terms of the calculations in the draft report, the true age fell within the estimated age range in 70% (49/70) of the attempts - we attach a spreadsheet of the workings

https://docs.google.com/spreadsheets/d/15mIawRCbQ5TcxeTYahv1gepcjROxNDzmfp_lIlMSVn8/edit?usp=sharing



First name	Current age	Attempt 1	Attempt 2	Attempt 3	Attempt 4	Attempt 5
A	3	1-5	1-5	1-5	1-5	0-4
B	6	7-11	3-7	4-8	4-8	4-8
C	7	11-15	12-16	5-9	4-8	8-12
D	7	9-13	6-10	6-10	6-10	5-9
E	9	12-16	12-16	13-17	9-13	11-15
F	12	12-16	12-16	13-17	13-17	12-16
G	12	12-16	13-17	12-16	13-17	13-17
H	13	12-16	13-17	12-16	13-17	13-17
I	13	12-16	12-16	13-17	13-17	13-17
J	15	16-20	16-20	19-23	15-19	16-20
K	15	13-17	13-17	13-17	13-17	13-17
L	16	13-17	14-18	15-19	14-18	15-19
M	20	24-28	22-26	20-24	17-21	23-27
N	21	17-21	18-22	17-21	20-24	19-23

	"Yes" If within range, else "no"					
First name	Current age	Attempt 1	Attempt 2	Attempt 3	Attempt 4	Attempt 5
A	3	Yes	Yes	Yes	Yes	Yes
B	6	No	Yes	Yes	Yes	Yes
C	7	No	No	Yes	Yes	No
D	7	No	Yes	Yes	Yes	Yes
E	9	No	No	No	Yes	No
F	12	Yes	Yes	No	No	Yes
G	12	Yes	No	Yes	No	No
H	13	Yes	Yes	Yes	Yes	Yes
I	13	Yes	Yes	Yes	Yes	Yes
J	15	No	No	No	Yes	No
K	15	Yes	Yes	Yes	Yes	Yes
L	16	Yes	Yes	Yes	Yes	Yes
M	20	No	No	Yes	Yes	No
N	21	Yes	Yes	Yes	Yes	Yes
	Total	70				
	"yes"	49				
	yes %	70%				

17. Here are links to Yoti’s current white papers 2023 1. for facial age estimation and 2. for liveness. We would ask that these be referred to in the report, rather than the previous 2022 white paper.

[17] https://ico.org.uk/media/for-organisations/documents/4020427/yoti-sandbox-exit_repo rt_20220522.pdf.

- True Positive Rate (TPR) for 13-17 year olds correctly estimated as under 25 is

99.93%.

- There is no discernible bias across gender or skin tone for 13-17 year olds. The current TPRs are:
- 99.90% and 99.94% for females and males respectively.
- 99.93%, 99.89% and 99.92% for skin tones 1, 2 and 3 respectively.
- TPR for 6-11 year olds correctly estimated as under 13 is 98.35%.
- The current accuracy rates (Mean Absolute Errors) are:
- 2.9 years for 6-70 year olds.
- 1.4 years for 13-17 year olds.
- 1.3 years for 6-12 year olds