IAB/W3C Workshop Position Paper on consumer choice for Online Age Assurance

The rapid evolution of the digital landscape necessitates robust and ethical standards for online age assurance to safeguard individuals, particularly children, while respecting privacy and facilitating legitimate access to services. As regulatory bodies globally, including those in the UK, EU, and Australia, increasingly mandate age-appropriate access and "highly effective age assurance" for online platforms, the need for clear, practical, and interoperable standards becomes paramount.

This position paper, informed by extensive experience and drawing on detailed insights into the implementation and challenges of age assurance technologies, advocates for an approach that prioritises a diverse range of flexible tools, anchored in transparency, independent auditing, and user-centric design.

The Imperative for Effective and Standards-Aligned Age Assurance

Online age assurance is no longer merely a best practice; it is a regulatory imperative. Jurisdictions such as the UK, with its Online Safety Act, and the EU, through the Digital Services Act (DSA), are implementing stringent requirements for platforms to assess and manage user age. The core objective is to simultaneously protect children from harm, enable adults to access content freely, and preserve individual privacy. Platforms in the UK, for instance, must conduct children's access and risk assessments and implement "highly effective age assurance" (HEAA) for services containing harmful content, such as pornography. The Australian Online Safety Amendment Act also requires platforms to take reasonable steps to prevent children under 16 from creating or maintaining accounts. These regulatory demands underscore the urgent need for internationally recognised and auditable minimum standards.

A Comprehensive Toolkit: Spanning Verification, Estimation, Inference, and Tokenisation

Effective age assurance requires a diverse set of tools, allowing platforms to select the most appropriate method based on risk, regulatory context, and user needs. A "one-size-fits-all" approach is insufficient given the varied nature of online services and global user demographics. A provider's toolkit should therefore encompass multiple methods, offering choice and flexibility.

A (non exhaustive) sample of age assurance approaches is detailed below:

1. Facial Age Estimation: This AI-powered method instantly estimates a user's age from a live selfie. It is particularly inclusive for the over 1 billion people globally who lack identification documents, representing 13% of the world's population.

- How it works: A neural network, trained on millions of diverse facial images, analyses pixels from a live selfie to compute an age estimate. No images should be stored, they should be deleted instantly, and non-identifiable, ensuring user privacy. The system should operate anonymously and not create biometric templates or learn from user images.
- Accuracy and Bias Mitigation: The technology should actively monitor and mitigate bias across age, gender, and skin tone, with published performance data demonstrating fairness. This information should be independently verified by a credible third party and should be made available to the public, for instance via NIST Face Analysis Technology benchmarking.
- Safety Buffers: To account for the probabilistic nature of age estimation and minimise False Positives (where an underage person is incorrectly estimated as overage), configurable safety buffers are essential. For example, for an 18+ check, setting the thresholds to 20 - 25 years can reduce the false positive rate. A 3 or 5 year buffer was defined as acceptable by regulatory bodies in Germany (KJM, FSM)
- Livenness should be undertaken to a strong, recognised industry standard
- Presentation attack detection and Injection detection should be undertaken

2. Digital ID App: Users can share a data-minimised age credential (e.g., "18+") from their reusable digital ID app after a one-time verification process. This method ensures no personally identifiable information is contained in age tokens when shared with a relying party. The user remains in control of what is shared.

3. ID Verification: Users upload an ID document and capture a selfie, which are then verified for authenticity and matched to confirm identity and age. This method offers different "levels of assurance" based on the checks performed:

- Low Assurance: Document authenticity check.
- Medium Assurance: Document authenticity check + proof of ownership (face-to-document match).
- High Assurance: Document authenticity check + proof of ownership + liveness check (to ensure a 'live' person).
- Very High Assurance: High assurance + injection attack protection (to ensure images are genuine and not manipulated).

4. Other Age Assurance Methods (Inference & Checks): Beyond biometrics and traditional ID, a comprehensive approach includes various data-based methods:

- Credit Card Check: Verifies if the user holds a credit card and is 18+.
- Database Check: Verifies name, date of birth, and address with a credit reference agency.
- Social Security Number (SSN) Check (US): Verifies validity and age against third-party databases.
- Mobile Phone Check: Matches details against mobile service provider records.
- Electronic ID Check: Allows sharing age attributes from national eIDs like Swedish BankID, Finnish eID, or Danish MitID.
- Email Check: Matches user email against third-party databases and checks domain for supporting information.
- Double Blind Check: Meets specific regulatory requirements for double anonymity, as seen in France.

5. Tokenisation and Interoperability: Once a user's age has been successfully verified, tokenisation allows for the reuse of age credentials across multiple platforms without repeated full age checks, significantly reducing user friction and compliance costs. Age tokens are pseudonymised and do not contain personally identifiable information. They can be configured to specify accepted criteria, such as maximum validity time, verification method, and age threshold. This strengthens repeat authentication by binding age verification to passkeys saved on devices. The interoperable, tokenised approach, with regular authentication of the current user, is already working at scale with millions of checks performed annually following international standards.

However, in order for age tokens networks to be able to develop; there needs to be clear minimum standards and audit amongst the participants in the network. The first few weeks of the OSA enforcement has shown that not all age checks being undertaken have met the level of HEAA; which would preclude those checks from being able to be recognised in an age token network.

Prioritising Transparency, Independent Audit, and Robust Security and proposing minimum standards

The integrity and trustworthiness of age assurance systems are paramount and depend on their transparency, rigorous independent auditing, and robust security measures.
1. Transparency and Explainability:

- Providers should offer clear, accessible explanations of how their technologies work, what data is processed, and how decisions are made. This includes publicly available white papers on facial age estimation, liveness detection, and injection attack detection.
- Detailed performance metrics, including Mean Absolute Error (MAE), True Positive Rates (TPR), and False Positive Rates (FPR), should be published, broken down by demographics like age, gender, and skin tone, to ensure full transparency to regulators and clients. This allows clients and regulators to make informed, risk-based decisions and set appropriate safety buffers.
- Communications for users, especially minors and parents, should use plain language and visual aids, such as explainer videos, to enhance understanding of how AI works and why age checks are performed. Messaging should be constructive and non-judgemental, particularly when a user does not pass an age check.

2. Independent Audit and Certification:
   - Age assurance systems should undergo regular, independent security audits and penetration tests by reputable firms. Certifications such as SOC2 Type 2, ISO 27001, ISO 27701, ISO 9001, IEEE 2089.1 and PAS 1296:2018 and the upcoming ISO/IEC 27566 (CD) demonstrate adherence to high security and privacy standards.
   - Participation in independent benchmarking or testing programs like NIST's ongoing Facial Age Estimation (FATE) evaluation and the 2025 Australia benchmarking is crucial for validating accuracy, fairness across diverse demographics and technology readiness.
   - Engagement with regulatory sandboxes and expert bodies, such as the UK ICO Sandbox and German Commission for Youth Protection (KJM), fosters trust and helps shape regulatory guidance.
3. Robust Security and Fraud Prevention:
   - Age assurance systems should implement defence-in-depth security measures, including TLS 1.2+ encryption for data in transit and AES-256 encryption for data at rest, alongside strict access controls.
   - Sophisticated anti-spoofing techniques, such as passive liveness detection, are essential to prevent the use of photos, masks, videos, or deepfakes.
   - Injection attack detection is critical to guard against manipulated image feeds and ensure input source integrity.
   - Providers must have a well-established incident response framework with 24/7 monitoring, rapid containment, root cause analysis, and prompt notification to regulators and users if required.
4. Redress Mechanisms and Continuous Improvement:
   - Systems must offer clear redress mechanisms for users to raise concerns, dispute age decisions, or request information. While platforms integrating the service often manage the direct user support, age assurance providers should enable and encourage clear communication of these pathways.
   - Continuous improvement is vital, involving ongoing research, model training, external benchmarking, and ethical oversight to ensure systems remain state-of-the-art, accurate, and fair.

Strategic Placement in the Tech Stack

The placement of age assurance within the tech stack is a critical discussion point, demanding a balance between user convenience, privacy, and operational feasibility.

1. The Principle of Proximity (Point of Access Checks):
   - Age checks should occur as close as possible in time to the relevant activity or service requiring age-appropriate access. This "principle of proximity" ensures clarity for users on why the check is needed.

- This approach is already working at scale, with billions of checks performed annually following international standards. It is particularly effective for services like gambling, adult content, or mixed-age gaming, where immediate age confirmation is essential.

2. Challenges of System-Level (OS/App Store) Checks:
   - While system-level age checks (e.g., at the operating system or app store level) are sometimes proposed, they present significant practical and ethical challenges.
   - Privacy and Data Risks: Centralised collection of sensitive age and identity data at the OS level increases the risk of large-scale data breaches, unauthorised data mining, and corporate dominance over identity markets. Anonymous access to apps may be undermined.
   - Burden on Parents and Device Management: Sole reliance on parental controls has proven ineffective, with industry data showing minimal adoption rates (e.g., only 1% of US parents use Snapchat and Discord parental control tools). System-level checks could shift significant, unmanageable responsibility to parents and complicate device reuse, resale, and sharing within families.
   - Sector Exclusion and Innovation Stifling: Many age-restricted goods (e.g., alcohol, nicotine) are not sold via app stores. Centralisation could also stifle innovation in age assurance technology, which has largely come from independent providers.
   - Cost and Accountability: Determining who pays for these checks and how accountability for failures would be managed are significant concerns, potentially leading to unchecked fees or a shift of liability.
   - Operational Challenges: It would be difficult to make OS-level checks mandatory globally, ensure re-authentication for shared devices, or secure open-source operating systems from manipulation.
   - Proportionality challenges: For an effective OS (or ISP) to be put in-place without risking over-blocking, there need be a mechanism for the system to identify and qualify the risk of all content in real-time. Effectively generating a requirement for all devices to include content scanning without context of the content or its platform.

3. Layered Approaches and Unintended Consequences:
   - While a layered approach combining device setup checks, periodic reauthentication, and service access checks is discussed, implementing more age checks than currently required could lead to unintended consequences.
   - Some apps and services might exclude minors entirely to avoid regulatory burdens, thereby depriving children of valuable online experiences. Increased friction could discourage legitimate users, and small developers might struggle to meet complex certification requirements, stifling innovation and competition.

Balancing Protection, Autonomy, and Accessibility

Effective age assurance must delicately balance the protection of children from harm with respecting their evolving rights, autonomy, and participation in society.

1. Evolving Capacity of Children:
   - Systems should enable platforms to tailor digital experiences based on a child's developmental stage, moving beyond a binary "allowed/denied" model to offer tiered access to features and content. For instance, highly restricted experiences for under 13s, limited functionality with parental controls for 13-15 year olds, and more autonomy for 16-17 year olds.
   - This aligns with frameworks like the UK Age Appropriate Design Code and the UNCRC General Comment No. 25, which encourage considering evolving capacities. Age assurance can support graduated parental consent models and risk-based safeguards.
2. Accessibility and Inclusivity:

- Age assurance solutions must be accessible to all users, including those with disabilities or limited technical literacy. This is achieved by offering a wide range of methods, including non-ID based options like facial age estimation.
- Systems should be designed with screen reader compatibility, high-contrast UI, scalable text, and support for keyboard-only navigation. Clear visual guides, simple language, and support for assistive technologies are crucial.
- For those with limited technical literacy, solutions should offer step-by-step onboarding with visual cues, one-time setup for reusable credentials, and minimal interaction for methods like facial age estimation.
- Cultural considerations, such as engaging with First Nations communities to address barriers to documentation and sensitivities around facial image capture, are vital for truly inclusive design.

3. Support for Vulnerable Children:
    - For children in vulnerable situations (e.g., unaccompanied, in care, or estranged from parents), facial age estimation provides a crucial, non-ID-based option to prove approximate age without adult involvement or account creation. This helps prevent digital exclusion while meeting protection requirements.

Conclusion

Effective online age assurance is a complex but achievable goal, integral to creating a safer and more age-appropriate digital environment. Based on practical experience and adherence to established standards, we advocate for the development of future standards that embody the following principles:

- User Choice and Method Diversity: Standards should emphasise a flexible toolkit offering a range of age assurance methods, including robust verification, accurate estimation, and relevant inference-based approaches, suitable for different risk contexts and user populations.
- Privacy-by-Design: Principles of data minimisation and privacy-by-design must be paramount, ensuring that sensitive personal data is not unnecessarily collected, stored, or linked, and that anonymous options are available.
- Interoperability and Tokenisation: Standards should support interoperable, tokenised approaches to reduce friction for users and compliance costs for businesses, enabling the reuse of pseudonymised age credentials across multiple platforms.
- Auditable Transparency: Systems must be auditable and transparent, with public reporting on accuracy, bias, and security performance, allowing regulators and clients to make informed decisions.
- Configurable Safety Buffers: Standards should allow for configurable safety buffers in probabilistic methods like facial age estimation, enabling platforms to meet varying regulatory risk appetites and significantly reduce false positives.
- Recognition of Evolving Capacity: Standards should recognise the evolving capacity of children, moving beyond rigid restrictions to enable nuanced, age-appropriate experiences and graduated consent models.
- Inclusivity and Accessibility: Solutions must be inclusive and accessible, ensuring that no one is digitally excluded due to lack of traditional ID, disability, limited technical literacy, or cultural barriers.

By adhering to these principles, age assurance providers can foster the development of robust and ethical standards that protect children, empower users, and create a safer, more equitable digital environment for all, leveraging existing, auditable solutions and leveraging what works effectively today

Appendix
Links

OECD Study - Age Assurance Practices (OECD, June 2025)

- 28 services (56%) use some form of age assurance.
- Only 2 services (Yubo, Wizz) use it systematically at sign-up.
- 34 out of 50 services (68%) set a non-overridable numerical minimum age.
- 16 services (32%) allow access below the stated minimum age
- Only 3 services (TikTok, Yubo, Pornhub) publish data on underage accounts.

Yoti Position paper Where in the stack should age assurance happen

Meta explainer video on Yoti facial age estimation

Yoti white paper facial age estimation
Yoti white paper anti injection detection
Yoti white paper liveness detection

Yoti developer documentation https://developers.yoti.com/

How age estimation is built - To learn more about how we built our facial age estimation technology, watch the below three minute video by Be inTouch.

How AI really works - We partnered with Youtuber Be inTouch to help talk about how AI really works. Check out the eight minute and twelve minute explainer videos below to learn more about AI, deep learning, and how it's used to estimate age and create safe spaces online.

[video - part 1] [video - part 2]

Play Verto research: young people's attitudes towards facial age estimation

NIST Benchmark results explainer video, in Q&A format with Erlend, our Head of R&D

https://vimeo.com/963138932/a0b52b9da0?share=copy

If you click on this icon (bottom right of the vimeo video) you can skip through the questions

Why has this benchmark been so hotly awaited?
vimeo.com/manage/videos/963138932/a0b52b9da0#chapter=15411266

Why is sample size so important?
vimeo.com/manage/videos/963138932/a0b52b9da0#chapter=15411291

NIST's four different sets of test data
vimeo.com/manage/videos/963138932/a0b52b9da0#chapter=15411294

How their test data sets compare to the images we use
vimeo.com/manage/videos/963138932/a0b52b9da0#chapter=15411302

Minimum test set
vimeo.com/manage/videos/963138932/a0b52b9da0#chapter=15411332

Why test the technology on over 50s?
vimeo.com/manage/videos/963138932/a0b52b9da0#chapter=15411338

How industry should handle skin tone and ethnicity?
vimeo.com/manage/videos/963138932/a0b52b9da0#chapter=15411347

What the top 3 vendors have achieved
vimeo.com/manage/videos/963138932/a0b52b9da0#chapter=15411368

Are NISTs results consistent with our own results?
vimeo.com/manage/videos/963138932/a0b52b9da0#chapter=15411378

What are the inconsistencies?
vimeo.com/manage/videos/963138932/a0b52b9da0#chapter=15411383

How the industry has improved since the last test
vimeo.com/manage/videos/963138932/a0b52b9da0#chapter=15411388

Why did NIST test both still images and video?
vimeo.com/manage/videos/963138932/a0b52b9da0#chapter=15411389

What are NIST and the vendors going to finetune?
vimeo.com/manage/videos/963138932/a0b52b9da0#chapter=15411415

The key terms used in the report
vimeo.com/manage/videos/963138932/a0b52b9da0#chapter=15411418

The risk of testing on the same person
vimeo.com/manage/videos/963138932/a0b52b9da0#chapter=15411452