Paper for Submission to the IAB/W3C Workshop on Age-Based Restrictions on Content Access (agews)

Title: Effective control of age restricted material for young people
Authors: Tom Newton (tom.newton@qoria.com) & Tim Levy (tim.levy@qoria.com)

Abstract

This paper explores the multifaceted challenges associated with age-based restrictions on online content access for young people, encompassing technological, ethical, business, privacy and practical considerations.

Contemporary views on age based restrictions view platforms as the gatekeepers of content and so aim to fix safety at that level. Such a view places the onus on publishers but is awash with issues including imprecision and error rates, susceptibility to avoidance, concerns around privacy, the prohibitive costs for small platforms and the removal of practical choice by parents.

Alternative network-based filtering models for age-gating access have different but similarly challenging issues including misalignment with the IETF's core mission of end-to-end privacy enhancing Internet protocols.

As an alternative, this paper proposes a reliable and secure alternative. An architecture whereby guardians (whether parental, or in loco parentis, e.g. schools) can, and are indeed expected to, install or configure reliable safety technology on devices.

This technology is standards based, supported by the device ecosystems and can apply content filtering and app controls through supported device and application level APIs. All of the components of this architecture exist in-market today.

Through this model a number of important opportunities arise:

- **Guardians are empowered.** They have choice - not only in how their child/student is protected, but also in who is involved in that protection.
- **Verification & consent is streamlined:** Ed-tech providers already routinely interact with school systems to leverage the trusted school <> parent relationship. This is a tremendous asset, allowing guardians to consent to platform access and set maturity or other policy directives - and all without having to expose identity data to the internet.
- **Security is preserved:** In this model, the objectives of the IETF remain supported. Internet access need not be interfered with to preserve safety.
- **Age assurance:** Age assurance can be reasonably effected, without excessive concern over pushing smaller organisations out, or moving access to less well regulated parts of the internet.

Introduction

It is an inevitable result of the digital environment that various governments and other bodies are going to require that young people cannot access “mature content” online. There are a number of challenges with this objective, including those of technology, ethics, business, privacy and practicality. This position paper seeks to address some of these problems, and propose a route to a safer internet, while preserving choice, privacy and security.

The Challenges of Age Verification

Contemporary views on age based restrictions view platforms as the gatekeepers of content and so aim to fix safety at that level. But platform level gating of access has a number of significant challenges and as a singular measure, unreliable:

- **Cost:** The cost of age gating through verification or assurance measures is going to be significant and create barriers for innovation and for smaller platforms and hobbyists.
- **Privacy:** In order to prove age, some platforms will require access to personal information, likely to generate real and perceptive concerns.
- **Avoidance:** If you can pretend to be in a non-AV territory, you can avoid this entirely - or indeed if you spend any time on holiday in a non-AV territory.
- **Cheating**: Users can game the system and pretend to be older. This is a balance - it behoves the publisher to only install cheaper techniques that’s good enough to please the regulator, not that a parent or guardian would consider robust. AV is in conflict with cost, and friction.
- **Non-Compliance:** Sites don’t have to comply, as it may be very challenging to sanction a site which is not based in the same jurisdiction as the regulator.
- **Bluntness:** Trials in Australia show 15% error rates in age assurance technologies when applying an 18 month tolerance. The imprecision is beyond what will be acceptable to the community which will drive false expectations, confusion and civil disobedience .
- **Unintended consequences:** Platform based age gating will likely result in unintended consequences such as encouraging use of VPNs (which is the experience in all jurisdictions trialling so far); disabling geo-location capabilities and moving children to darker parts of the internet.
- **Global applicability:** Platform level age gating might need to be applied differently under different legislative regimes. Contrary to popular belief, it is not straightforward to geolocate a user, and this may become more difficult in future.
- **Insufficiency:** Platform level age gating might work for the types of content governments are wont to legislate - but parents and guardians are likely to want finer grained controls. This is likely to lead to some users having additional controls by other means.

Having said that, publisher-side-age gating could still have a part to play, even if it is as part of a larger ecosystem.

A safety solution which empowers choice and preserves privacy

Whilst sites that might be suitable candidates for age restriction are almost uncountable in number, the average young person uses a relatively small number of devices. Additionally, access to these devices is almost certainly governed and monitored by a guardian (parent or school).

Accordingly, provided installation and management is simple & reliable, it is possible to expect guardians to install or configure safety technology on all devices (or accounts on shared devices) used by children.

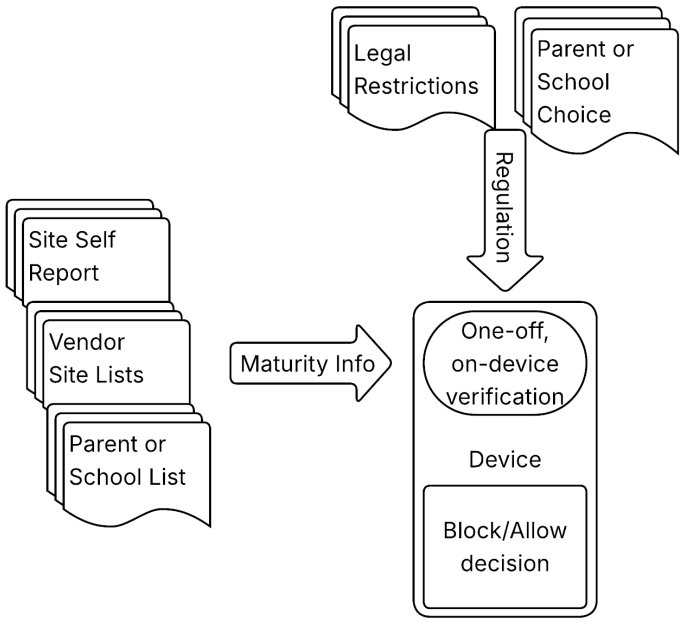
With brokering the decision on-device, we no longer have to be concerned that the network we are on does not support age restrictions or that the site we are visiting has correctly recognised our location and is treating us correctly.

On-device safety technology can communicate with online platforms & sites serving age-tokens or other policy settings, without necessitating the sharing of personal data. This allows for age-moderated experiences (eg. social media sites might change their timeline algorithm for younger audiences, or alter their advertising mix, rather than risk an outright ban).

A proposed architecture might look like:

1. 1st or 3rd Party safety tech running on end-user devices captures guardian's policies e.g. maturity and consent.
2. Safety tech vendors are expected to adhere to local government legislation - and therefore have some immutable rules for certain maturity levels in certain territories.
3. Online platforms voluntarily support rating tags, consent, device handshakes, content moderation, algorithm transparency etc.
4. Safety tech running on end-user devices filters the internet where possible (eg in browsers).
5. Native apps are able to work with the installed safety tech vendor to acquire maturity settings, and "guarantee" their suitability.
6. Guardians can configure safety tech on devices to limit access to certain online platforms and / or block access to platforms which don't publicly comply with all of certain standards.
7. Online platforms are expected to comply with standards and work with device level safety techniques.

8. Where users of online platforms are not “protected” (ie the devices they are using do not have safety tech active) then the platform would be expected (through safety regulations) to apply age assurance of age verification techniques depending on the content and service.



With such an approach:

- **Guardians are empowered.** Schools & parents are empowered with a simple instruction. Install safety technology on your ward’s devices. Doing so will provide the best safety coverage available.
- **Verification & consent is streamlined:** Ed-tech Safety technology providers routinely interact with school systems to leverage the trusted school <> parent relationship. This is a tremendous asset of the community, allowing guardians to consent to platform access and set maturity or other policy directives - and all without having to expose identity data to the internet.
- **Security is preserved:** In this model, the objectives of the IETF remain supported. Internet access need not be interfered with to preserve safety.
- **Publishers can build their business models:** Online platforms can operate with confidence that their users are appropriate and can therefore totally focus on the experiences and services their customers desire.

Of course this brings its own challenges. Safety software on the device is necessarily a marketplace. There will be safety software for different OS’s, and some which work together. Some will be free, others, at cost. How a government mandates that devices given to young

people is arguably less challenging than it seems. In most AV-leaning regulatory environments this type of control is already the case in Schools (CIPA, KCSIE).

For mobile devices, networks frequently force install software, and they already know if a device is in use by a child.

And it’s rare that internet security technology isn't running on end-user devices. This is because it's supported by OEMs and embedded in set-up. If parents typically protect their data, then we should expect them to protect their kids (if similarly easy).

Some may argue that on-device techniques can fail because:

1. Kids can remove them: This is only true because of the commercial choices of Google, Apple and Microsoft, it is not a technological issue.
2. Kids may use shared or parent devices: This may be true but this creates the same issues for platform level age gating.
3. Kids may use burner devices: This may be true but at least parents are strongly empowered through a simple instruction: “protect the device to protect your child”

Additionally, as we have stated, platform age gating is a problematic and blunt instrument. Even if it works for a particular platform; the internet is not safer. It just moved the issue.

Ultimately an holistic model, as we propose, recognises stakeholder roles and incentives. It is inevitable that two systems will be in play - “publisher side” which is largely implemented to shield the publisher from liability, and guardian-side, which is more focussed on protecting the young person.

What’s needed to get us there?

There are actually almost no technical hurdles to achieve this model; this is because it already exists, and works reliably in enterprise safety (e.g. schools) today.

Today, in enterprise environments device owners can seamlessly install safety technology on end-user devices through OS enabled MDM tools. Such safety technology is robust and can reliably apply personalised policies, filter internet content (in browsers) and control online apps (controlling app access). It can also direct users to age-appropriate versions of online platforms where suitable APIs exist with the online platform (e.g. YouTube).

Consumer app developers do not however have the same or interoperable access to these capabilities which has stifled competition and adoption.

Accordingly the key missing piece is device ecosystem interoperability which is essential to enable ubiquitous on-device safety technology and to enable competition. It is also the necessary precursor to encouraging (or mandating) sites and online platforms to interoperate with them.

A more complete solution requires these technical components:

1. **Device ecosystem interoperability:** Device ecosystems should be required to open up their platforms to enable access to all necessary features to permit guardians to manage and monitor device and internet access.

   As stated above, for the most part, such capability already exists but is only fully available for enterprise app developers.

   Improvements here should include better access to browsing data (it is clear attempting to filter browsing at the network level is no longer feasible), and a clear, standard set of APIs for mediating control with apps whose traffic can't be easily classified.

2. **User verification and/or age assurance & consent:** User devices need to know what policy settings to apply e.g. jurisdictional age restrictions and parent/school policies and the ability to set their rules.

   As stated above, for the most part, such capability already exists. Parents can install / configure safety settings and age tokens on devices and schools do make available SIS data for ed-tech providers for parent consent and so on. What is required is industry standards and mandatory interoperability. OS vendors are somewhat deficient here in allowing adults to easily and flexibly maintain control of the devices used by young people in their care. Addressing this is of utmost importance regardless of routes taken by AV.

3. **A standard way to identify adult sites**: We already have the RTA tag[1], this is trivial for sites to implement.

Conclusion

There are many reasons why age verification is a hot topic and we believe the wider standards community can help lead industry to a sensible and safe way forward with the architecture proposed.

[1] RTA - Parental Control Software - Website Label