# Papers, Please: A First Look at Age Verification on the Web

Shreyas Minocha,[†*] Isaac Sheridan,[†] Harry Oppenheimer,[†] Paul Pearce,[‡] and Michael A. Specter[†*]

[†]*Georgia Institute of Technology,* [‡]*University of California, Irvine*

***Abstract*—Since 2022, twenty-five US states covering more than 40% of the US population have adopted laws compelling websites with content "harmful to minors" to verify their users' ages. Many websites that comply with these laws are widely reported to rely on third-party services, effectively outsourcing the age verification process. However, little is known about how these services are shaping the web and affecting user privacy.**

**In this work, we conduct the first large-scale exploration of age verification providers on the web. We begin by exploring the CrUX top one million websites from three different states—two with legal age verification mandates and one without—to identify the prevalence and composition of age verification services. We then reverse engineer Yoti, the dominant age verification provider, and provide an in-depth privacy analysis.**

**Our findings show that age verification services can be ineffective in restricting minors, create significant new privacy risks for end users, and are causing the first instance of cross-state balkanization of the web in the US. We find that Yoti often requires end users to share sensitive data—photos of their face, government IDs, credit card details, browser fingerprinting data, the website being accessed, and more. Such data may be entrusted not only to the contracted provider, but also to several "fourth parties" that are significantly less visible to users. We identify security and privacy issues, and connect these findings to key assumptions underlying recent Supreme Court precedent.**

## 1. Introduction

> "Users only have to submit verification to the covered website itself or the third-party service with which the website contracts... Both those entities have every incentive to assure users of their privacy."
>
> —Clarence Thomas, writing for the 6–3 SCOTUS majority in *Free Speech Coalition v. Paxton*

Twenty-five US states—encompassing over 40% of the US population—have introduced legislation requiring certain websites to verify their visitors' ages using "reasonable age verification methods." Many of these laws are modeled after Louisiana's Act 440, which created "liability for publishers and distributors of material harmful to minors." Texas rapidly followed with a bill that was challenged in *Free Speech Coalition v. Paxton*. In 2025, the Supreme Court upheld Texas' law against free speech challenges in a 6–3 decision [1]. The introduction of these laws follows a global trend of legislation requiring websites to employ age verification, most notably the UK's *Online Safety Act*.

In the US, age verification laws differ by state, leading to differing mandates, enforcement triggers, and technical privacy constraints (see Section 2.1). In most, the law applies to entities that publish content that the state considers harmful to minors. The laws trigger based on where the affected content is viewed, rather than where it is hosted or created; in five states, these laws create rights to recourse specifically for residents of those states. Most states define "harmful content" to mean depiction of nudity and sexual activity that "lacks serious literary, artistic, political, or scientific value for minors" and content that, under contemporary community standards, appeals to "the prurient interest."

To handle the growing complexity of complying with these laws, a growing number of sites are turning toward an ecosystem of third-party age verification providers. Many of these companies are members of the *Age Verification Providers Association* (AVPA), a global trade association for companies that offer these services, that presents itself as "making the Internet 'age aware.' "

Unfortunately, little is known about the prevalence, functionality, security, and privacy of these third-party age verification providers. Though there have been efforts to support privacy-preserving cryptographic protocols for age verification [2], it is unclear if or how these services engage with these tools, or protect user privacy at all. Effectiveness is a concern as well—it is unclear how often these laws are obeyed by websites and whether they put other burdens on the user.

In this paper, we present a first look at the deployment of web age verification in the United States. We find that legislative requirements have created an ecosystem of third-party services whose deployment differs greatly among states, balkanizing the US's web along jurisdictional lines. Additionally, we find that the market for age verification providers is highly concentrated, with one service being used in over 60% of sites detected to use age verification in affected states. We perform a security and privacy analysis of the most popular age verification provider in the US, Yoti, and find that it creates serious privacy risks.

*. Contact: specter@gatech.edu, shreyas@shreyasminocha.me

**Contributions.**

1) We measure the prevalence and composition of third-party age verification providers among the top one million websites. We find that although dozens of third-party providers exist, the five most commonly used providers cover over 70% of the websites that employ age verification in states with mandates.
2) We compare the prevalence of age verification providers across multiple US states and find that the presence of the top providers is concentrated in states with age verification mandates.
3) We compare the prevalence of providers between websites that do and do not label themselves as serving adult content, and find significant differences in the ecosystem for these use cases.
4) We reverse engineer and perform a privacy analysis of the most commonly used age verification provider, Yoti. We find that the service collects significant private information beyond what is strictly necessary to verify age, including high-entropy browser and device metadata, and other granular telemetry. We also discover that Yoti relies on sharing sensitive user information with several less user-visible fourth parties.

**Roadmap.** Section 2 covers legislative background related to age verification laws in the United States, methods of restricting digital adult content, and presents related academic work. Section 3 describes our methods for measuring the prevalence of age verification providers, and for analyzing Yoti's security and privacy. Our results from the prevalence measurement are described in Section 4, and we present an analysis of Yoti's security and privacy in Section 5. Finally, Section 6 discusses takeaways and conclusions from our results.

## 2. Background & Related Work

In this section, we give an overview of the legal background surrounding age verification on the web. We also survey work related to privacy law compliance, geoblocking, adult content controls, and age verification systems.

### 2.1. Legal Background

**US legal history.** There have been repeated efforts in the US to restrict Internet content based on users' ages. The Communications Decency Act of 1996 first attempted to restrict minors access to "obscene or indecent" content. The landmark *Reno v. American Civil Liberties Union* found this law unconstitutional, as it would block constitutionally protected speech. The Supreme Court decision found that there was "no effective way to determine the age of a user who is accessing material through e-mail, mail exploders, newsgroups, or chat rooms" [3]. As a result, websites would have no choice but to remove content entirely because they had no method to distinguish among their users.

Twenty-five years later, legislators argued that new tools have enabled websites to age-gate access, and in 2022, Louisiana passed the first state-level age verification law. In 2023, the Texas Legislature enacted House Bill 1181 (88th RS) with bipartisan support. The act states that any "commercial entity" that publishes a website where more than one-third of the content is "[s]exual material harmful to minors" must use "reasonable age verification methods" to verify that its users are at least 18 years old. Age verification must use "digital identification" or "a commercial age verification system" that uses government identification or financial transactional data. Fines can reach $\$10{,}000$ per day, and $\$250{,}000$ if a minor accesses restricted content. The law did not immediately go into effect due to a lawsuit from the *Free Speech Coalition*, a non-profit trade association founded in 1991.

In 2025, *Free Speech Coalition v. Paxton* reached the Supreme Court, creating an opportunity to set a new precedent for online age verification. For technical information, the Justices relied on amicus briefs filed by Yoti Ltd—a prominent age verification provider—and AVPA. Yoti Ltd argued that online age verification is equivalent to a bartender checking ID. In oral arguments, Justice Barrett asked the Free Speech Coalition to "explain [ . . . ] why [online age verification] is so uniquely burdensome here when it's not been in the real-world context." Justice Alito posited that online age verification providers "have very tough privacy limitations built into them" and specifically cited Yoti as an online age verification provider with high privacy standards.

The Supreme Court's majority decision—written by Justice Clarence Thomas and joined by 5 colleagues—found that online age verification does not violate the First Amendment [1]. The majority decision characterizes online age verification as:

(a) an effective way to distinguish minors (who can be restricted) from adults (who have a constitutional right to access affected content),

(b) a "modest burden" unlikely to restrict speech of adults and similar to methods in place "for decades," and

(c) low-risk and unlikely to trigger individual privacy concerns because age verification providers "have every incentive to assure users of their privacy."

Relying on these assumptions, the court found that age verification laws would only need to satisfy intermediate scrutiny, a significantly lower bar for restricting speech than the strict scrutiny applied previously. However, these claims, which have material importance to the constitutionality of online age verification, remain untested.

**State-level age verification laws.** Table 1 summarizes the twenty-five state laws that have been enacted as of April 2026. Most states apply their laws to commercial websites where more than one-third of content qualifies as harmful to minors, though thresholds vary, from Tennessee's 10% to South Dakota and Wyoming's "regular course of business" standard. Kansas' laws refer to the percentage of pages viewed in a month, rather than a percentage of content hosted. Four states (IN, MI, SD, TN) extend their laws to non-commercial websites. Most states exempt news organizations, ISPs, cloud providers, and search engines.

TABLE 1. AGE VERIFICATION REQUIREMENTS BY STATE.

| | Effective Month | Permitted Methods | Hosted Content | Carveouts | Enforcement |
|---|---|---|---|---|---|
| AL | 2024-10 | A | 1/3 | N, I, C, S | P, A |
| AR | 2023-07 | G, D, I | 1/3 | N, I, C, S | P |
| AZ | 2025-09 | G, M, T | 1/3 | N, I, C, S | P |
| FL | 2025-01 | A | 1/3 | N, I, C, S | P, A |
| GA | 2025-07 | G, D, I | 1/3 | N, I, C, S | P, A |
| ID | 2024-07 | G, D, T | 1/3 | N, I, C, S | P |
| IN | 2024-08 | T, B, M | 1/3 | N, I, C, S | P, A |
| KS | 2024-07 | B, A | * | I | P, A |
| KY | 2024-07 | G, T | 1/3 | I, C, S | P |
| LA | 2023-01 | G, D, T, A | 1/3 | N, I, C, S | P |
| MS | 2023-07 | G, D, T | 1/3 | N, I, C, S | P |
| MO | 2025-11 | G, M, T, A | 1/3 | N, I, C, S | A |
| MT | 2024-01 | G, D, T | 1/3 | N, I, C, S | P |
| NE | 2024-07 | G, D, T, F | 1/3 | N, I, C, S | P |
| NC | 2024-01 | B, A | 1/3 | N, I, C, S | P |
| ND | 2025-08 | G, D, T, B | 1/3 | N, I, C, S | P |
| OH | 2025-09 | G, T, B | ? | N, I, C | A |
| OK | 2024-11 | D, T, B | 1/3 | N, I, C, S | P, A |
| SC | 2025-01 | D, T, B | 1/3 | N, I, C, S | P, A |
| SD | 2025-07 | G, C, F, A | Any | I, C, S | A, C |
| TN | 2025-01 | T | >10% | N, I, C, S | P, C |
| TX | 2023-09 | G, M, T | 1/3 | N, I, C, S | A |
| UT | 2023-05 | D, T, B | 1/3 | N, I, C, S | P |
| VA | 2023-07 | B, A | 1/3 | I | P |
| WY | 2025-07 | C, A | Any | I, C, S | P |

For permitted methods: G=Government-issued ID, D=Digitized ID, I=IAL 2 [4], B=Commercially available database, T=Transactional data, M=Mobile credential/Digital ID, C=Debit/credit card, F=Financial document, A=∼Any.
For carveouts: N=News, I=ISP, C=Cloud, S=Search engine.
For enforcement: P=Private lawsuit, A=AG action, C=Criminal lawsuits.
* Appears on $>$ 25% of pages viewed in a calendar month
? "Substantial" revenue, advertising, or content

Several states have requirements alongside age verification. Texas' HB 1181 and Alabama's Act 2024-97 require websites to display warnings about the psychological harms of pornography and a reference to a mental health helpline. Alabama's law requires adult websites to obtain written consent from every individual depicted in a private image prior to its distribution. Ohio Revised Code §1349.10 requires a "geofence system" offered by a "licensed location-based technology provider" to "dynamically monitor" the location of end users of affected websites. Missouri's law requires search engines to blur or hide affected material from the search engine landing page unless they use age verification.

Most states address privacy by requiring the deletion of personally identifiable information (PII) after age verification. There are two exceptions: Virginia and Tennessee. Virginia has no requirement for deletion of PII. The *Protect Tennessee Minors Act* requires the provider to retain at least seven years of historical "anonymized" age verification data, and requires websites to re-verify the user's age after 60-minute "age-verified sessions." Some states have stronger requirements for data deletion and privacy. Under South Dakota's HB 1053 (2025), selling or retaining PII collected for age verification is a Class 1 misdemeanor, and a subsequent violation is a felony. Florida's HB 3 (2024) requires that affected websites offer "anonymous age verification" in addition to "any commercially reasonable method of age verification." Such anonymous age verification must be conducted by a nongovernmental US-owned, US-controlled third party that may not retain PII or use it for other purposes.

**Enforcing age verification.** Several US states have already seen cases filed under age verification laws. For example, in 2024, the Texas Attorney General brought lawsuits under Texas' HB 1181 against *Multi Media, LLC*, *Hammy Media*, and *Aylo Global Entertainment*—all operators of large pornography sites. The case against Aylo is active as of April 2026, while Multi Media and Hammy Media entered into settlements [5]–[7]. In May 2025, lawyers representing the National Center on Sexual Exploitation filed lawsuits against four companies under Kansas' law for failing to implement reasonable age verification methods. Kansas' Attorney General also brought a case under this law against an adult site operator in January 2025 [8]. In the same month, Indiana's Attorney General sent cease-and-desist notices to pornography websites [9]. Similarly, in August 2025, Florida sued the owners of several pornography sites and an adult advertising network under the state's HB 3 [10].

## 2.2. Related Work

Though we are the first to perform a large-scale measurement of age verification on the web, we follow in the tradition of analyzing the effects of Internet legislation through web measurement.

**Privacy Compliance.** A large body of work has focused on questions about compliance with the European Union's General Data Protection Regulation (GDPR) [11]–[15]. This includes large-scale automated analyses of websites for violations [16], and cookie notice compliance [17]. Others measure the prevalence of consent management platforms on popular websites [15], [18].

Others have measured the impact of Internet legislation in the United States. For example, Tran et al. [19] and Nortwick et al. [20] measured the presence of opt-out links that are mandatory under the California Consumer Privacy Act (CCPA). More recent work analyzed GDPR and CCPA cookie banners for compliance issues [21]. Reyes et al. evaluated the compliance of Android apps with the Children's Online Privacy Protection Act (COPPA) [22].

**Geoblocking.** Another line of work characterizes geographical differences in the availability of digital services. Ablove et al. studied the impact of the United States' embargo against Cuba on geoblocking [23]. Kumar et al. studied the blocking of mobile apps on Google Play [24]. McDonald et al. conducted a measurement study of CDN geoblocking [25]. *FilterMap*, *Augur*, and *Quack* measured application-layer blocking using various measurement techniques [26]–[28]. These works have largely focused on either regionalization for business reasons or government-mandated censorship, and provided evidence for inter-

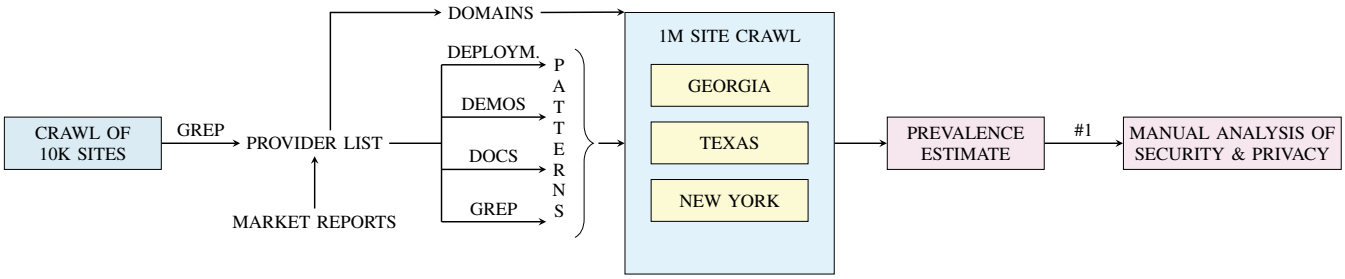


Figure 1. We build a list of age verification providers (AVPs) and patterns to match them, which we use to characterize AVP prevalence. Finally, we perform a security and privacy analysis of the most common AVP.

net balkanization across international borders, rather than among US states.

**Circumvention.** Previous studies find that users may use virtual private networks (VPNs) to circumvent age verification [29], and find significant challenges in censorship circumvention in Chinese and Russian contexts [30].

**Privacy policies.** A number of works study privacy policies of various web-based services. Zhou et al. developed *PolicyComp* to automatically flag overly broad data collection in mobile apps, and conducted an analysis of over 10K privacy policies [31]. Wilson et al. proposed methods for automated analysis of privacy policies [32]. Nokhbeh Zaeem et al. constructed a large corpus of web privacy policies [33].

**Restricted To Adult (RTA) Label.** Although the legal basis for strict verification by websites is relatively new, technological means of restricting access to adult content through parental controls have existed for years.

One such technology is the "Restricted to Adults" (RTA) label, whereby a website can include a metadata element to voluntarily indicate that it serves adult content. Around 5,000 of the top 1 million websites were estimated to use the RTA label in March 2025 [34]. Filtering software, such as parental control software, commonly uses the tag to block access to websites based on this self-labeling [35].

**Facial Recognition.** Facial recognition technology for age verification systems may have fairness and bias-related issues. *Virtual U* found ways to bypass face liveness detection by building virtual 3D facial models in 2016 [36]. Burgess et al. evaluated the accuracy and biases of facial recognition classifiers as used in remote proctoring [37].

**Age verification.** Though we are the first to conduct a large-scale study of age verification providers on the web, age verification has been studied in other contexts. Vallina et al. conducted a study of the privacy practices of pornographic websites, including a manual analysis of age verification use on 50 sites, finding limited use of age verification suites [38]. Others studied the impact of these laws on Google Search trends [29], and performed user studies evaluating responses to online age verification mechanisms [39]. Most recently, a study examined age verification in mobile apps [40].

## 3. Methodology

Figure 1 presents an outline of our methodology. We began by conducting a small-scale pilot study to discover the age verification providers commonly used in the wild. By searching these pages for various terms related to age verification, we were able to discover a handful of common providers. We then expanded this set to include providers associated with a representative industry trade association, and those included in industry market reports. Using information from live deployments, public API documentation, and public demos, we developed a set of detection scripts to infer the presence of each of these providers. We then used these to detect instances of age verification in crawls of the top one million websites from three different states. We additionally performed a security and privacy analysis of Yoti, the most widely seen age verification provider, which is also referred to by name in the *FSC v. Paxton* oral argument.

### 3.1. Data Collection & Measurement

In this section, we describe how we selected candidate age verification providers (Section 3.1.1), crawled the top 1M sites from three states (Section 3.1.2), and checked for the presence of providers on webpages (Section 3.1.3). We also describe the scripts we use to categorize content hosted by several sites with age verification in Section 3.1.4.

**3.1.1. Provider discovery and selection.** To build a set of known age verification providers, we began by conducting a Texas-originating crawl of the home pages of 10,000 websites selected uniformly at random from the April 2025 version of the CrUX list [41]. We queried these webpages for terms related to age verification to build an initial set of popular age verification providers (AVPs).

To complement this set, we also included several age verification providers that advertise their services through various well-known industry channels. We included AVPs listed as members of the *Age Verification Providers Association*, as well as several AVPs listed in a 2025 report about

the Face Liveness Market [42] from *Biometric Update* (a biometrics news source) and *Goode Intelligence* (an identity-focused consultancy firm). Finally, we added many providers certified by the UK-based testing and certification body, *Age Check Certification Scheme* [43].

We limited ourselves to providers that offer online age verification services similar to those mandated by the various US state age verification laws. In total, we considered over 60 providers, as enumerated in Appendix Table 4.

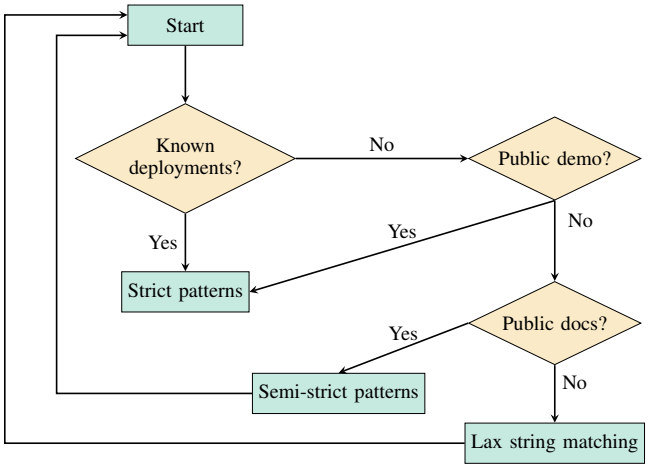


Figure 2. Our process for building patterns to detect providers involved iteratively improving our detection heuristics based on newly detected websites.

**3.1.2. Large scale crawl of 1M websites.** We used the Chrome User Experience Report (CrUX) list of top domains from June 2025 [44] which, being informed by Google Chrome telemetry, is highly representative of human traffic [45]. We used the full set of one million domains for each state's crawl, and we iterated through the domains in a randomized order so that effects like rate-limiting or CAPTCHAs are independent and identically distributed. For each domain, we saved the index page of the website hosted at that domain. Crawls were performed in August 2025.

**Crawling from different states.** We picked Texas because its age verification bill was the subject of *FSC v. Paxton* in the Supreme Court, and Georgia because its age verification law was relatively recent at the time. New York serves as a control, since it did not have any active age verification laws as of 2025.

When necessary to simulate traffic originating from different states, we set up SOCKS5 proxies to route our traffic through servers in those states. To get a faithful view of the content of websites as viewed by users, and to reduce the risk of getting blocked by services like Cloudflare, we used Selenium for our crawls. To further minimize the risk of automatic blocking, we set up xvfb virtual displays to go with our Selenium instances.

**Control webpages.** To validate that our traffic appeared to originate from the intended geographic location, we deployed our own test webpages. We deployed one instance of this webpage behind Cloudflare's free tier proxy, and one directly to a server controlled by us. We configured this page to render the request IP address to approximate how websites would observe our requests. Additionally, we included the scripts, elements, and other patterns we found to be typical of some age verification providers. These test webpages were crawled prior to visits to other sites.

At the end of our crawls, we queried IP geolocation databases for the IP addresses logged by the test webpage on our server to ensure that they originated from the expected regions. We also confirmed that our requests to the test webpage behind Cloudflare were not triggering CAPTCHAs. Additionally, we checked that our heuristics correctly flagged the presence of the providers we emulated on our test webpage.

**3.1.3. Provider detection on webpages.** To provide a baseline view of the prevalence of age verification providers, we searched webpages for domains we know to be associated with each provider. Since we know at least one domain for each AVP, this let us minimize the influence of the availability of demos and other information about the AVP. However, testing for the presence of domains in web source code is a blunt tool, and results in too many false positives.

To get a more accurate estimate of provider prevalence, we developed more precise patterns for their detection. Age verification providers manifest themselves in a variety of ways in the client-side source code. Some providers have their customers include iframes from, or hyperlinks to, the AVP's domains. Some have them include scripts that are hosted by the AVP. Others yet create DOM elements with certain IDs or classes, require initialization of the AVP in the website's inline JavaScript, or load assets from the AVP's domains. We represent these characteristics as BeautifulSoup [46] expressions that represent whether the AVP was detected on a page. We construct each AVP's detection logic as a disjunctive combination of such patterns to minimize false negatives.

In building our provider detection logic, we follow an iterative approach (see Figure 2). We prefer to construct patterns to match an AVP based on the rendered source code of websites that use that AVP. When those are unavailable, we rely on official demos and developer documentation.

For some providers, we were unable to find any public data about their client-side implementation, even after looking through matches for broad queries related to the provider: no known live deployments and no public documentation. We chose to prioritize broad coverage of the age verification ecosystem; for these providers, we therefore resort to regular expression matches for, e.g., known AVP domains. We elaborate on this challenge in Section 3.3.2, and Appendix Table 4 describes the combination of detection methods used for each provider.

To test the precision and recall of our detectors, we sampled 200 websites with the RTA tag and undertook a

manual detector validation of Georgia data. We focused on validation of self-identified adult content as it has a high latent probability of using age verification; randomly sampling the CrUX 1M to identify false positives or negatives would be inefficient, as the probability of sampling positives is very low. We found 32 true positives, 152 true negatives, 0 false positives, and 16 false negatives. 13 of 16 false negatives are for Yoti on websites with a unique inclusion pattern that appear to be operated by one company. Two false negatives are for VerifyMyAge, and one is for Emblem; although we miss these in this sample, our detectors find 100s and 10s of matching websites respectively from the global set of 800K. Of the true positives, all but one (VerifyMyAge) are Yoti websites. We emphasize that our results attempt to establish a lower bound on the prevalence of the top providers. As noted in Section 3.3.2, there are circumstances where we are unable to attempt to detect providers (e.g., post-login, server-side), also a cause of false negatives.

**3.1.4. Web Classification API.** To get a sense of the kinds of websites that use age verification providers, we used a third-party web classification API to retrieve IAB categories [47] associated with domain names performing age verification. In the absence of a well-known, reliable open-source dataset, we chose to use Klazify [48]. We queried this API in November 2025 for websites with age verification providers but without the adult content (RTA) label.

## 3.2. Reverse-Engineering Methodology

To understand the design of Yoti's age verification service, we began by performing static and dynamic analysis on Yoti's public "supermarket demo" [49], which supports, among others, the age estimation and ID verification methods. We also compared these with deployed versions of Yoti on multiple websites to confirm that our findings hold up in the real world. To avoid legal and ethical concerns, we restricted our dynamic analysis to using Yoti like a normal user, following the standard protocol flow. Over the course of our analysis, we did not send any abnormal traffic to Yoti's servers.

Our analysis also involved observing real-time HTTP traffic in browser developer tools via Chromium and Firefox. We also captured HTTP Archive (HAR) traces of browser activity for reference.

To reverse engineer obfuscated JavaScript, we relied heavily on dynamic analysis through the debugger in Firefox's developer tools. As we iteratively recovered strings and code, we recorded these results in a working copy of the code. We used standard JavaScript IDE plugins to rename identifiers while preserving semantics. Through this manual reverse engineering effort, we eventually deobfuscated enough of the relevant code to gain a clear picture of the protocol and the data collected. Finally, we confirmed that our understanding of the code matched the browser APIs being accessed and the HTTP requests being sent and received through dynamic analysis.

Some of the obfuscated code involved in our reverse-engineering effort was re-obfuscated on each page reload, i.e., variable names and control flow obfuscation would change. This made reverse-engineering more challenging; we attempted to do much of this work in a single reverse-engineering session. In the interest of open science, we make available our annotated and deobfuscated version of the relevant parts of Yoti's code (see Reproducibility).

We reverse-engineered version 1.6.0 of Yoti, as denoted in the "preloaded config" loaded on Yoti's age verification portal; this version number was consistent between August and November 2025. In normal operation, Yoti's JavaScript dynamically downloads security-focused code from the server, which holds independent versioning information. For this code, we examined version 2.6.2, the version that was in use in production on major websites in August 2025, when we began our analysis. We further confirmed that these scripts were the same version as those deployed in the wild. Additionally, we have verified our ability to achieve a successful result with a replaced image with version 2.8.0, the latest version as of February 2026.

According to Yoti's End-of-Life policy [50], version 2.6.2 will not be deprecated until May 2026, and will not expire until November 2026. Version 2.8.0 will be deprecated in September 2026 and expire in March 2027.

## 3.3. Limitations

**3.3.1. Crawling.** We were unable to successfully download an average of $86{,}000$ websites across the three states. Therefore, we limit our analysis to the $780{,}391$ remaining websites that were successfully downloaded from all three states. Of these websites, across the three crawls, an average of $1.2\%$ of sites returned Cloudflare error pages, $0.12\%$ resulted in empty downloads, and $0.06\%$ returned pages that could not be parsed by BeautifulSoup. Since we rely on automated visits to thousands of websites, and through VPS IP addresses, we may be caught by bot detection and crawling prevention mechanisms. When our scripts are identified as blocking targets, some sites may simply refuse to connect, but others might serve a seemingly valid webpage without authentic age verification scripts or links.

For ease of setup, we used SOCKS5 proxies in our crawling browser to tunnel our traffic, which may be easier for websites to detect and block than VPNs.

**3.3.2. Prevalence Analysis.** We focus on detecting age verification providers that are observable from the index pages of websites, without any user interaction. We do not detect age verification provider use that is, say, behind a login page, or invoked after a credit card payment.

Accurately detecting the use of age verification providers is challenging. Due to differences in design across providers, it is difficult to design consistent heuristics for programmatically checking for their presence. Providers also vary in their transparency to non-customers: some lack public developer documentation and very few have public demonstration

sites. As a result, our modules for detecting providers vary in their detection techniques and precision.

Many third-party providers of age verification also offer other products, such as identity verification, Know Your Customer (KYC), and fraud detection. Although we make efforts to avoid misclassification using publicly available information and real-world deployments, for some providers this is not possible. In the interest of broad coverage of providers, we develop detectors for these, even though they are unable to differentiate between age verification deployments and identity verification deployments.

To aid in evaluation and reproduction of our results, we describe the methods we use to match each provider in Appendix Table 4, and make our provider detection modules publicly available (see Reproducibility). Since our detectors are expressions that enumerate patterns and return a binary result, they are easy to audit, improve, and update.

We also note that while our methods focus on capturing client-side signs of providers, some age verification products are designed to be used only on the backend without user interaction; this is sometimes advertised as "stealth mode" or "headless" age verification [51]. Our methods would not detect age verification that is performed entirely server-side using previously collected information, such as an email address or credit card information.

**3.3.3. Security and Privacy Analysis.** Our findings are limited to what can be observed from an end user's perspective. We do not have access to Yoti's server-side code, and also lack the level of access that Yoti's age verification customers would have. Nevertheless, we are able to discover many aspects of Yoti's age verification processes by analyzing client-side JavaScript.

## 4. Age Verification Prevalence

In this section, we describe the composition of the third-party age verification provider ecosystem, as well as trends in their presence in Georgia (GA), Texas (TX), and New York (NY). Of the nearly 800,000 websites considered, we detect the presence of age verification providers on 1,650 (0.21%) of them in GA, 1,602 (0.21%) of them in TX, and 619 (0.08%) of them in NY. Of the top 1K websites, we detect providers on 37 (4.61%) sites in GA, 36 (4.48%) sites in TX, and 2 (0.25%) sites in NY (normalized).

We see a clear difference between the prevalence of providers in states with age verification mandates (Georgia and Texas) and in the state without (New York). This suggests that age verification (AV) legislation has a quantifiable effect on the view of the Internet from different US states, with AV-mandated states seeing nearly three times as many websites use third-party age verification providers.

### 4.1. Provider Ecosystem

We observed that a small set of age verification providers is used in a large proportion of websites that perform age verification on their homepage, especially in states with age verification laws. Figure 3 shows the six most commonly used age verification providers in GA and TX; this also includes the top three providers in NY, covering almost 40% of sites with providers in that state.

In Texas and Georgia, approximately 1,000 websites, more than 60% of websites with AVPs, use Yoti. The next four most commonly used providers, together with Yoti, cover 72% of GA websites and 74% of TX websites that have AVPs. There is a long tail of providers in GA and TX, which comprise about a quarter of websites with AVPs.

In crawls from New York, the age verification provider ecosystem is more diverse, albeit less prevalent. The most common provider in NY is ID.me, which is used on close to 90 websites. We suspect that ID.me is likely being used for identity verification, rather than for restricting content that is "harmful to minors": none of the websites using ID.me include the RTA label. The next four most common providers are Veriff, Mitek, AgeVerif, and Yoti in that order. These top five providers account for 338 websites—a little over 50% of the sites using AVPs in NY. In sharp contrast to Georgia and Texas, Yoti is detected only 45 times in our crawls from New York, which is less than 5% of that in GA and TX.

Of the 991 websites detected to have Yoti in Georgia, 701 websites have references to `avsgate.com/age-verify/yoti`. Similarly, 650 of the 994 sites with Yoti in Texas have these references. This subset of websites presents hyperlinks to this endpoint, configured with a set of query parameters, such that accessing them redirects the user to a new Yoti session. Creating new age verification sessions requires access to an organization's private API keys, and we observe `avsgate.com` references on websites widely known to be operated by Hammy Media. This leads us to believe that this subset of 701 sites in GA and 650 sites in TX, or 703 distinct sites, corresponds to sites operated by them.

An additional 214 Yoti-enabled websites in Georgia, and 209 in Texas, have references to endpoints like `/agegate/yoti_av` and `/agegate/yoti_anon`. We observe these on websites operated by Multi Media, LLC. Since this is a highly specific and unusual pattern of loading the AV screen, we speculate that this subset corresponds to websites that are all operated by this same entity.

Almost all sites that are detected to use Incode and VerifyMyAge—the second and third most popular providers in Georgia and Texas—have the `/agegate` pattern associated with Multi Media, LLC sites. In GA, 214 of the 219 Incode sites have this pattern, and in TX, 209 of 214 do. Similarly for the 229 sites in GA and 220 sites in TX that appear to use VerifyMyAge, 214 and 209 sites respectively have the `/agegate` pattern. Additionally, all of these likely Multi Media, LLC sites appear to use Yoti, Incode, and VerifyMyAge in tandem, i.e., all sites that match the `/agegate` pattern have references to all three of these providers. These observations indicate that, despite being among the top providers, Incode and VerifyMyAge likely contract with only a small number of distinct organizations.

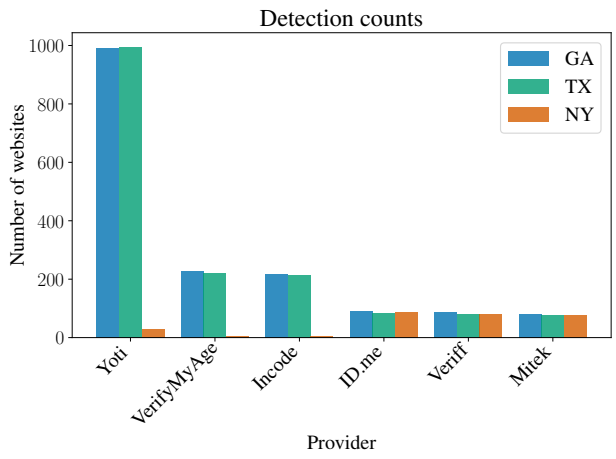


Figure 3. Number of websites detected with age verification providers in each state included in our study.

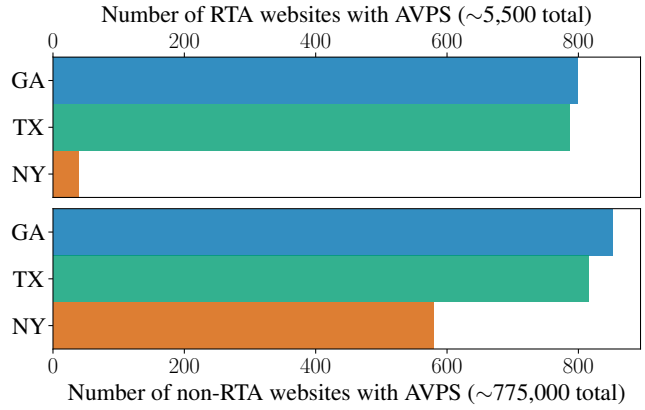


Figure 4. Comparing detection rates across the states, GA and TX have similar AVP detection rates, whereas New York has significantly lower rates. This difference is especially pronounced among the nearly 5,000 RTA-labeled websites.

## 4.2. Geographical Differences in Prevalence

There appears to be a significant difference in adoption between states that do and do not have AV mandates, but not among the states that have AV mandates. In GA, we observe age verification providers on 1,650 sites, and in TX, we observe them on 1,602 sites. In contrast, in NY, we find age verification providers on 619 sites.

Differences in AVP detection are apparent when comparing the same websites across states with and without mandates. Of the sites that are detected to have AVPs in GA, 1,054 (nearly 64%) are detected to *not* have AVPs in NY. Similarly, of the sites with AVPs in TX, 1,024 (also ∼64%) do not have AVPs in NY. Meanwhile, among the 619 sites with AVPs in NY, we find 23 (∼4%) and 41 (∼6%) that *do not* have AVPs in GA and TX respectively. Finally, the sets of GA and TX sites with AVPs are highly similar, with a Jaccard index of about 0.87.

Figure 3 highlights state-wise trends in the presence of popular providers, such as Yoti, Incode, and VerifyMyAge. We detect only 45 sites that use Yoti in NY, which is less than 5% of that in GA and TX. For VerifyMyAge and Incode, this is more pronounced: both providers combined appear on fewer than ten websites in NY, whereas they are used on over 200 distinct websites in GA and TX.

## 4.3. Self-labeled Adult Websites

We split the set of websites we considered based on whether they included a Restricted to Adults (RTA) meta tag, which is a form of self-labeling for adult content (see Section 2.2). The crawls from all three states have approximately 5,500 websites that include this label: 5,788 in GA, 5,516 in TX, and 5,230 in NY.

Figure 4 shows the overall age verification provider detection counts across the three states, split by whether an RTA label was included. We detect AVPs on approximately

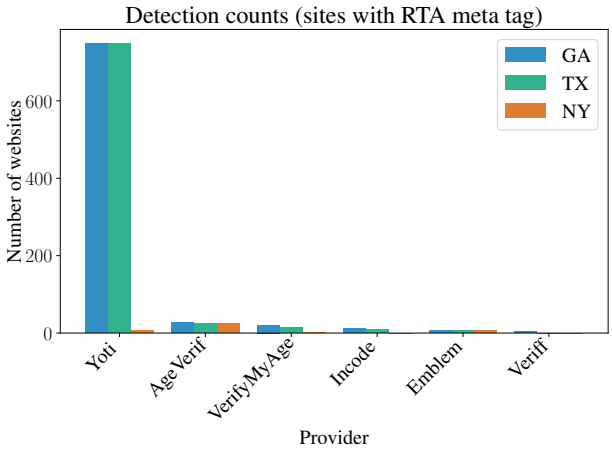


Figure 5. Yoti is by far the most commonly detected provider among sites that label themselves as "Restricted to Adults" through the RTA tag.

800 of the nearly 5,000 websites with the RTA label, and approximately 800 of the nearly 775,000 websites *without* the RTA label, in both GA and TX. In NY, we find fewer than 100 RTA sites with AVPs, and almost 600 non-RTA sites with AVPs.

Figure 5 shows the AVP detection counts among the websites that include the RTA meta tag, across GA, TX, and NY. We see that Yoti significantly dominates the AVP market among RTA-labeled websites, and we see it detected almost exclusively in GA and TX—the states with AV mandates. In GA and TX, Yoti appears on RTA sites 747 times and 778 times respectively, compared to NY's 11. The next most common provider, AgeVerif, appears fewer than 50 times in all three states. In NY, age verification providers are detected 39 times among RTA sites.

Figure 6, similarly, shows the AVP detection results among the nearly 775,000 sites that *do not* label themselves as adult content with the RTA meta tag. Here too, Yoti is

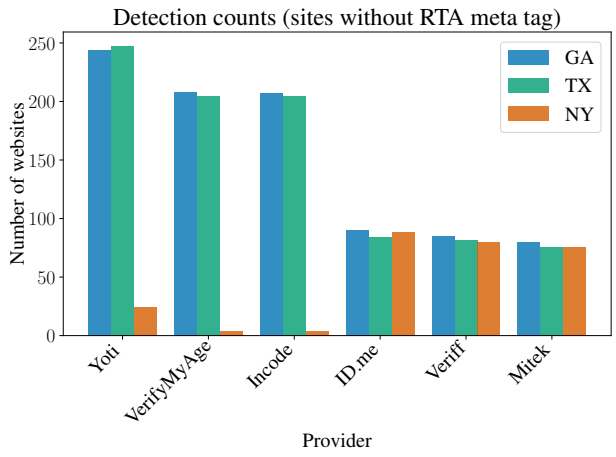


Figure 6. Among sites that do not label themselves as "Restricted to Adults" through the RTA `<meta>` tag, non-Yoti providers have relatively higher occurrence.

the most commonly used AVP in states with age verification mandates, with 244 and 247 sites using it in GA and TX respectively. However, its dominance is relatively less pronounced; the next two most commonly observed AVPs, VerifyMyAge and Incode, are observed on approximately 210 distinct sites in GA and TX. In NY, on the other hand, Yoti, Incode, and VerifyMyAge have a minimal footprint, covering just 31 non-RTA sites. Instead, ID.me, Veriff, and Mitek dominate.

### 4.4. Websites Employing Age Verification

Nearly half of the GA and TX sites found to use AVPs include the RTA label, compared to fewer than 10% of the ones in NY.

To understand which non-RTA sites contain age verification providers, we used a third-party API to retrieve category information for the non-RTA labeled sites (see Section 3.1.4). Based on the data fetched from the API, an additional 123 GA sites and 118 TX sites belong to the "Adult" category. Gambling websites make up another large group among non-RTA websites with providers, with 41 sites in GA and 39 sites in TX being classified as "Games/Gambling." The umbrella category "Sensitive Subjects" is applied to 24 GA sites and 19 TX sites, and seems to include mostly vape, cigarette, Cannabis, and knife stores. Unfortunately, the categorization API fails to return results for around 70 of the queried websites for both states, and many of the other categories are too vague to be informative. In NY, 27 websites are categorized as "Adult," and 14 as "Sensitive Subjects."

We find evidence that deployment of detectable age verification among likely mandated sites is low. We noted earlier that approximately 5,500 websites include the RTA label in each of the three states we considered. Since the RTA label is designed to let websites label themselves as "being inappropriate for viewing by minors" [52], it is likely that these websites would be subject to age verification mandates. However, we only detect 13.8% and 14.2% of RTA-labeled sites to contain age verification in GA and TX respectively.

We note that the number of websites that include RTA labels is a lower bound on the number of websites that primarily serve adult content, let alone the number of websites that fall within the scope of the various state AV mandates (see Section 2.1). The true number of websites where "material that is harmful to minors" is "more than 33.33 percent of total material" [53], or "sexual material harmful to minors" is "more than one-third" of the material distributed by a website [54] is likely to be higher [38].

We found that several websites that we detected to have age verification providers occupy high positions in the CrUX top 1M list. Among the top 1,000 most-visited websites, we detected AVPs on 37 websites in GA, 36 websites in TX, and 2 websites in NY. Of the ones in GA and TX, over 80% are adult entertainment websites likely operated by Hammy Media. Among the top 10K sites, we detect AVPs on 133 sites in GA and 124 sites in TX. Finally, considering the top 100K sites, there are close to 500 websites that we find to use age verification providers in GA and TX.

## 5. Yoti Privacy and Security

We report on our reverse engineering effort exploring Yoti's age verification suite. We separate our analysis into stages, representing the process by which a user would interact with Yoti's system: starting the verification flow (Section 5.1); the loading of the Yoti session configuration (Section 5.2); the age check process, which varies by method (Section 5.3); and eventually, redirection to the original website (Section 5.4). Finally, we describe the security and privacy risks uncovered by our analysis in Section 5.5.

### 5.1. Initialization

A website checking a user's age through Yoti starts by creating an ephemeral session for that user, and pointing them to Yoti's website through an iframe or a hyperlink. From the user's perspective, this is a page on `age.yoti.com` with a URL that includes a "session ID" and an "SDK ID." The SDK ID is a unique public identifier associated with an organization that signs up to use Yoti. Although it is possible to generate these sessions while rendering the HTML document on the server, we observed that many websites instead host internal APIs that generate and redirect to age verification sessions (e.g., `avsgate.com` from Section 4.1).

### 5.2. Session Configuration

When creating a new session, a website must specify various aspects of the age verification flow, including the set of supported age verification methods, how it would like to receive age verification results, and whether it wants the user's age in years or just a comparison against a threshold.

Most of the configuration details of a session are fetched in the user's initial interaction with Yoti. This includes the list of the configured verification methods and associated thresholds, parameters for integrity features (Section 5.3.1), third-party IP detection API endpoints (Section 5.2.1), the Stripe API 'publishable key' (Section 5.3.3), and more.

**5.2.1. IP Address and Geolocation Querying.** Soon after the user begins Yoti's age verification flow, Yoti makes a GET request to `api64.ipify.org` [55] to fetch the user's IP address. Next, it makes a GET request to `ipinfo.io` [56] to retrieve geolocation information associated with the request IP address. The response from this API includes the hostname, city, state, country, geographical coordinates, organization, and postal code associated with the IP address (see Listing 1). Notably, both of these requests include the `Referer` header with the value `https://age.yoti.com/`.

## 5.3. Age Confirmation and Estimation

After determining the user's region, Yoti allows the user to select between the configured age verification methods, such as facial age estimation, physical ID verification, and credit card verification. We examine each below.

**5.3.1. Selfie-based Age Estimation.** Yoti's age estimation method attempts to determine the user's age using ML applied on a photo of their face captured in real time from their device's camera. This photo is captured through a component Yoti calls the "Secure Capture Module" (SCM), which encapsulates the process of capturing a photo of the user, extracting it as an image, and preparing the payload uploaded to Yoti's servers for age estimation.

The SCM is a $\sim$500 KiB, minified, heavily obfuscated JavaScript module. Obfuscation methods include the mangling of strings, numeric literals, and control flow. Additionally, JavaScript property accesses, typically denoted as in `crypto.subtle.encrypt`, are mangled to instead use `crypto["subtle"]["encrypt"]` notation, but with obfuscated strings.

Once the user selects age estimation as their preferred method, the SCM is immediately downloaded from Yoti's servers. However, before the SCM is invoked, Yoti uses facial recognition on-device to detect and position the user's face. We observe requests to fetch weights and data for a "tiny face detector model" that is identical to the one in FaceAPI, an open-source facial recognition library [57].

Once the user's face is properly aligned, the SCM collects and processes a significant amount of data that is sent to Yoti's servers. In particular, it collects the photo captured from the user's camera and telemetry, including significant high-entropy browser and device metadata (see Table 2). It also includes data about the camera's properties, the frame rate of the camera stream, and metrics about download and processing times.

The SCM uses some cryptography, which we briefly describe here before returning to its implications in Section 5.5.3. If the image encryption setting is enabled (as it is by default), the SCM encrypts the captured image using AES-GCM with a key and initialization vector (IV) derived in the browser. Similarly, the telemetry and metadata collected are also encrypted under AES-GCM in the browser. The SCM also computes an HMAC using a similarly derived key over much of this data, presumably to achieve the integrity guarantees claimed in Yoti's developer documentation. The HMAC payload includes the (optionally encrypted) image, the encrypted metadata, the version of the SCM code, and the JSON Web Token.

TABLE 2. USER AGENT METADATA FIELDS ("CLIENT HINTS") SENT AS PART OF YOTI'S AGE ESTIMATION METHOD

| Metadata Field | Explanation |
|---|---|
| `Device-Memory` | Available RAM (approx.) |
| `Save-Data` | Whether reduced data usage is preferred |
| `Downlink` | Approximate Internet bandwidth |
| `ECT` | Effective connection type (e.g. 2G, etc.) |
| `RTT` | Approx. round trip time on the application layer |
| `Sec-CH-UA` | User agent brand & low-res version info. |
| `Sec-CH-UA-Arch` | CPU architecture |
| `Sec-CH-UA-Bitness` | Bitlength of a memory address |
| `Sec-CH-UA-Full-Version-List` | User agent brand & version info. |
| `Sec-CH-UA-Mobile` | Whether browser is on a mobile device |
| `Sec-CH-UA-Model` | Device model |
| `Sec-CH-UA-Platform` | Operating system |
| `Sec-CH-UA-Platform-Version` | Operating system version |

The data eventually sent to Yoti's servers in a POST request includes the image (either encrypted as described above or in plain); a "secure token," including the encrypted metadata, the aforementioned HMAC, the SCM version, and the JSON Web Token; and a boolean value signaling whether the user was prompted to consent to the collection of their biometric data. Yoti's API then responds to this request indicating whether the estimated age is above the configured threshold.

**5.3.2. Government-issued ID Verification.** Yoti supports verifying a user's age based on photos of their ID, and photos of their face to match with the one on their ID. Using a real-time photo to match the ID is optional, and the choice is left up to the first-party website. The first-party may also add validity checks, a comparison with the birthdate in the barcode on the document, a check against an unspecified "fraudulent document database," and more [58].

The first step in the ID verification process involves the browser loading an ID verification iframe with some session-specific identifiers included as query parameters. Once loaded, this page makes a GET request to fetch the configuration specific to ID verification. This includes the list of identity-related checks requested by the website and the set of supported documents. The configuration is fetched after each step in the verification process to check

the user's position in the ID verification flow. The verification page then makes a request to a third-party API, `ipv4.icanhazip.com` [59], to fetch the user's IP. This request includes the `Referer` header with a value of `https://api.yoti.com/`.

Many requests made to Yoti APIs during the ID verification flow include an `X-Yoti-Device-Meta` header, a Base64-encoded JSON object that includes device, OS, and browser metadata (see Listing 2).

Once photos for all document pages have been collected, Yoti makes PUT requests to upload the ID pages to Yoti's servers. These requests include the document photos as JPEGs and a description of the method used to collect the photo (i.e., capture or upload).

Once Yoti successfully performs text extraction on the ID pages on its server, the user can proceed to the next step. At this stage, if "liveness" or "ID document face match" is configured by the website, the Secure Capture Module is loaded and the face capture process proceeds much like in the selfie-based age estimation method (Section 5.3.1).

After all steps have been completed, Yoti confirms the status of the session by querying the server, and proceeds once it receives a "complete" status. Yoti notes in its developer documentation that photos uploaded in sessions that are left *incomplete* are stored for 25 hours [58].

**5.3.3. Credit Card-based Age Verification.** Yoti supports a credit-card-based age verification method, as explicitly permitted by South Dakota and Wyoming's age verification laws (see Table 1). This method collects a user's credit card details (card number, expiry date, CVV code or equivalent) and area code to authorize a credit card transaction for a small amount. Yoti relies on Stripe [60] to authorize this charge in order to verify its validity.

We used a US credit card issued by Visa to verify age for a website that configured Yoti to support this method. Upon successful completion, the issuing bank alerted us to an international transaction for 0.40 USD (the equivalent of 0.30 GBP at the time) under the merchant name "YOTI LTD." However, since Yoti configures Stripe to only authorize the transaction without capturing the funds, this transaction did not appear in our transaction list, and no money was transferred.

Over the course of this process, Stripe, the card network, and the issuing bank receive various pieces of information about the user's age verification request. The issuing bank is able to observe the charge authorization, including the amount, and the merchant name "Yoti Ltd." In addition to the user's card information, Stripe also receives information tied to Yoti, such as Yoti's Stripe API key, several `Referer` headers set to `https://age.yoti.com/`, and the merchant name "Yoti Ltd." Stripe also collects significant telemetry that could likely be used to uniquely identify a device. Most notably, we observed that Stripe also received the domain of *the first-party website that loaded Yoti*. Yoti typically receives the first-party website's domain through the `Referer` header,[1] which is accessible to Stripe's client-side JavaScript, and is eventually sent to Stripe as part of their "advanced fraud detection" signals.

## 5.4. Completion and Redirection

At the end of a successful age verification session or exhaustion of the permitted number of age verification attempts, the user is redirected to the first-party website. The age verification result can be communicated to the server either through a Session ID query parameter in the callback URL, or through a webhook notification. Depending on how the website configured Yoti, this result is either the user's age in years or a pass-fail result on the comparison against the configured threshold [58].

Websites that use Yoti can access records that include the Session ID uniquely identifying an age verification attempt, the age-checking method used, the status of the attempt, and a timestamp of when it was initiated [61].

## 5.5. Security and Privacy Analysis

Here, we elaborate on the security and privacy implications of Yoti's design and real-world implementation.

**5.5.1. High-entropy data collection.** As noted in Section 5.3.1, Yoti collects a significant amount of high-resolution data about the user's device. It is unclear what the use of this data is, and we note that little information collected here appears to be necessary in estimating the age of a user, assuming that one is doing so purely from the image captured or the user's ID. We further note that much of what is collected (OS version strings, available RAM, connection type, and CPU architecture) is also gathered by well-known fingerprinting libraries (e.g., [62]). Along with the user's IP address, it is likely that this data is uniquely identifiable, allowing for unpermissioned tracking of the user's device.

**5.5.2. Third-party IP APIs.** One of the earliest actions in the age verification process involves HTTP requests to third-party APIs to query the user's IP address and geolocation. The `ipify.org` API is used to retrieve the user's IP address and the `ipinfo.io` API is used to estimate the user's geolocation. These parties learn of the association between the user's IP address and a Yoti age verification attempt through the `Referer` header.

**5.5.3. Secure Image Capture.** Yoti's age estimation and ID verification methods support an optional, but enabled-by-default "secure mode," also referred to as *SICAP* (Secure Image Capture). Yoti's documentation claims to ensure that the image captured via SICAP cannot be manipulated before it is submitted to Yoti [63]. Therefore, we consider an

1. Unless websites customize the `Referrer-Policy` header to `no-referrer`, cross-origin requests, such as the one from the first-party site to Yoti, share the domain through the `Referer` header.

end user who is attempting to pass age estimation without presenting their face to Yoti's servers.

We found that it was possible to replace the captured image before it is sent to Yoti and achieve a successful verification result. The security features of SICAP involve cryptography performed *in the user's browser* through the Web Cryptography API. We built a Chrome extension that injects a script into Yoti's age verification pages and hooks into the `toDataURL` method on `HTMLCanvasElement`. It replaces the captured image with a carefully cropped selfie that passes Yoti's server-side checks without triggering the spoofing detection. We also confirmed our ability to observe and alter the behavior of the script during its cryptographic steps.

**5.5.4. Government-issued ID Verification.** We confirmed that Yoti's servers receive photos of the user's government ID, as well as device, OS, and browser metadata. As with the requests in Section 5.5.2, the requests to the `icanhazip.com` API signal the association between the user's IP address and their use of Yoti.

**5.5.5. Credit Card-based Age Verification.** Yoti's credit card-based age verification method associates a financial transaction (specifically, a payment authorization) with the age verification attempt. This provides Stripe and the issuing bank with a record associating the user's credit card details with a Yoti age verification attempt. We additionally observed a Stripe anti-fraud signals API collect the domain of the first-party website that led the user to Yoti. Stripe can thus associate a user's card details with a visit to an age-restricted website, despite being a *fourth party* to the user–website relationship.

**5.5.6. Additional Data Sharing.** Table 3 lays out Yoti's data sharing declarations per its privacy policies [64], [65]. During ID verification for a US driver's license or state ID, Yoti queries the *American Association of Motor Vehicle Administrators* (AAMVA) to check the validity of the data extracted from the user's ID. In states that do not allow this through the AAMVA, and for other checks of US documents, Yoti relies on *Aristotle*, a DC-based data broker [66]. It also relies on *Veratad*, a US-based provider of identity verification, age verification, and fraud prevention [67].

**5.5.7. Data Retention.** Yoti makes a number of statements about how long data is retained but these claims are somewhat difficult to parse and appear to differ by verification method. In its privacy policy, Yoti claims that "for most age-checking methods, the data is deleted as soon as the check is complete" [64], and biometrics policy applicable in some US states clarifies that this applies to the age estimation method [68]. They also specify that they store results of age checks (session ID, method, status, timestamp) for up to six months, and that websites may store these for longer. Yoti declares that they may share information in response to requests from "[l]aw enforcement agencies, regulatory bodies or other legal authorities [...] which [they] have a legal obligation to provide," and that they may be compelled to retain data for longer [64]. Regarding the ID verification method, Yoti notes that websites have the ability to configure manual reviews of ID verification-based age checks, in which case Yoti may process the data for up to 28 days, potentially involving a cross-border data transfer to Yoti India [64].

TABLE 3. YOTI DECLARES THE 'THIRD PARTIES' IT RELIES ON TO PERFORM AGE VERIFICATION FOR US USERS IN ITS PRIVACY POLICIES.

| Method | Data shared | External Service |
|---|---|---|
| ID verification | Name, Doc #, State, Gender, DoB, Exp. Date, Iss. Date | AAMVA, Aristotle |
| ID verification | Name, Address, DoB, SSN | Veratad |
| ID verification | "document information" | IDVerse |
| ID verification | [unspecified] | Amberhill, CIFAS |
| ID v. (manual) | ID, Face Photo | Yoti India |
| CC | CC Details, ZIP | Stripe |
| Mobile | Name, DoB, Mobile #, Address | Telesign, TMT Verify, Esendex, Twilio |
| SSN | SSN | "one of our data providers" |
| Email | Email | Versium |
| Database | Name, Address, DoB, SSN | Veratad |

**5.5.8. Data Location and Cross-border Transfers.** Yoti's privacy policy specifies that age checks and results of checks are processed and stored either on Yoti servers in their data centers within the UK, or in a US, EU, or UK AWS region [64]. We queried the authoritative DNS nameservers for the A and AAAA records for the two domains we observed in user-facing components of age verification: `age.yoti.com` and `api.yoti.com`. We found that they did not appear to be using round-robin DNS and resolved to IP addresses in AS204050 (Yoti Ltd) in London, UK.

## 6. Concluding Discussion & Recommendations

Our observations paint a concerning picture of the privacy and effectiveness of age verification. Compliance is low—only roughly 14% of sites self-labeling as adult content perform age verification in states with mandates. Worse, sites that do comply via the dominant provider subject users to significant privacy risks. These results have important implications for future technical and policy designs, which we outline below.

**Protocols and Tradeoffs.** There are ongoing efforts to standardize age verification as a cryptographic protocol. In particular, the Mobile Driver's License (mDL) standard [2] could be adapted to provide web-based age verification using anonymous credentials and zero-knowledge proofs, providing no more information to the site than that the user is above a certain age as guaranteed by a trusted state or government authority. Google's Longfellow is a preliminary implementation of this kind of protocol [69], and the W3C

has begun standardizing a method of exposing these kinds of credentials to the web [70].

From a privacy perspective, such standards *could* be a superior solution to what we examine in this work, but it is worth noting that their development is still ongoing and current designs pose substantial censorship and privacy risks. The mDL specification, as defined today, does not describe how to handle PKI failures, ID expiration, or ID revocation, so poor deployment and design could allow government entities to halt Internet access for arbitrary citizens by invalidating their digital IDs. For privacy, exposing age-related data provides yet another signal to adversaries attempting to track users via device fingerprinting [71].

**Censorship Risks.** The censorship risks inherent in age verification and existing cryptographic proposals extend beyond the sites we examine here: without a significant shift in policy direction, age verification suites may eventually control users' ability to participate in online speech. Age verification laws targeting social media platforms have already passed in the UK, Australia, and Mississippi [72]–[74]. As a direct result of Mississippi's law, the social media site Bluesky has chosen to not offer its service in the state [75]. Further, state-level mandates become unnecessary if sites adopt *universal* age verification to avoid regulatory risks. The absence of mandates does not mean the absence of verification—we discover sites enforcing age verification in New York, where no verification law exists.

**Balkanizing the US internet.** We previously had robust evidence for Internet balkanization *between countries* due to privacy regulations like GDPR [11]–[15] or US sanctions policy [23]. This balkanization typically shows up as differences in blocking and web accessibility across countries.

To the best of our knowledge, this is the first measured instance of web access balkanization between US states. Prior work has documented state-level differences in privacy compliance UIs, such as opt-out links and cookie banners [19], [20]. However, age verification mandates create differential *access* barriers—the first measured instance of state-level legislation producing content gatekeeping. This difference means that content flows freely in some states but is gated behind verification in others.

As of April 2026, over ten US states are considering new online age verification laws. Our findings point to a critical, but often overlooked, consideration for policymakers voting on these rules: significant privacy risks to their constituents; *where* users access the web from within the US now has a significant impact on their level of privacy online.

**Incomplete Compliance & Limited Effectiveness.** We find that, in GA and TX, fewer than 15% of sites that self-label as adult venues enforce age verification. Although this is likely a lower bound estimate, it still indicates that many sites will choose to ignore requirements. Further, our analysis of Yoti indicates that its suite may be circumventable.

Even if age verification rules were universally adopted, their ability to restrict minors' access to adult content is unclear. On-device protections do not prevent willing or unwilling adult cooperation in bypassing these checks, nor do they prevent providers outside the affected region from ignoring mandates. It may be that such behavior is incentivized: privacy concerns and friction may lead users to drive revenue to non-compliant sites.

**Privacy Regulation.** We encourage regulators to consider adopting straightforward requirements on age verification providers to ensure user privacy. Several third and fourth parties are used in Yoti's age verification suite, and it is likely that this list is incomplete given our limited access to Yoti's backend. Further, we found that Yoti's privacy policies made unclear and somewhat conflicting statements about data retention. Mandatory, machine-readable disclosure of all fourth parties receiving data during the verification process, strict minimum data collection and distribution requirements, and penalties for failures to meet these standards could help mitigate some of these concerns.

**Freedom of Speech on an Age-Gated Internet.** Regardless of how age verification is balkanizing the US web, our security and privacy analysis of the most common age verification provider has implications for future free speech debates. In 2025, the US Supreme Court ruled that online age verification was not a violation of the First Amendment. The majority decision applied intermediate scrutiny, arguing that online age verification would not significantly restrict adults' access to content. The court argued this based on key assumptions about online age verification.

First, the court assumed that online age verification is effective in restricting access of minors to adult content. We find that in the present day, this is false. Age verification is circumventable, either via technical means by subverting the provider's restrictions, or by *visiting the many sites that do not obey the mandate*. Such sites may be heavily incentivized to continue disobeying age verification laws as they likely enjoy the traffic given up by compliant sites.

Second, the court believed that online age verification is privacy-preserving, with risks similar to in-person age verification. The majority decision used this logic to dismiss any "chilling effect" caused by online age verification. The Yoti Ltd brief [76] to the Supreme Court echoes this, claiming: "Age Verification Technology Does Not Unreasonably Implicate or Intrude Upon the Privacy of Adults."

This paper demonstrates that the privacy impact of online age verification is nontrivial, and that there are critical differences compared to in-person age checks. Users commonly submit to having their photos, driver's licenses, and location sent, together with enough high-entropy browser data to track them, all to a centralized authority. We leave it up to the security and privacy community to determine whether these privacy concerns represent a "reasonable" or "unreasonable" intrusion on users. However, it is clear that a bartender need not gain universal knowledge of all patrons' PII in an easily copyable and indefinitely retainable format, and share much of this data with various third parties.

**Ethics Considerations.** None of our experiments involved the collection or submission of personally identifiable information such as IDs of any unaffiliated individuals. Where necessary to analyze the functioning of an age verification provider, the authors used their own laptop cameras and personal information. In the process of conducting our crawls, we sent the minimal amount of traffic necessary to the sites we crawled. During our reverse-engineering efforts, we sent no malicious or abnormal data to any external servers. We responsibly disclosed the Stripe information disclosure issue and the design issue we found in Yoti's Secure Capture Module to Yoti.

**Reproducibility.** Artifacts to replicate the main findings of our paper, including our crawling code, analysis code, and datasets, are available online [77]. This also includes code from our analysis of Yoti's Secure Capture Module.

## Acknowledgments

We thank Stephen Checkoway, Eric Rescorla, and Peter Swire for sharing valuable feedback, and Abhishek Bhaskar for assistance running RIPE Atlas queries. We also thank the anonymous reviewers and the shepherd for their valuable feedback. This work was supported in part by NSF CNS Award 2247514.

## References

[1] *Free Speech Coalition, Inc. v. Paxton, 606 U.S. 461 (2025)*, Supreme Court of the United States, Jun. 27, 2025. [Online]. Available: https://www.supremecourt.gov/opinions/24pdf/23-1122_3e04.pdf.

[2] NIST. "Digital Identities - Mobile Driver's License (mDL) — NCCoE Mobile Driver's License Project," National Institute of Standards and Technology. [Online]. Available: https://pages.nist.gov/nccoe-mdl-project-static-website/.

[3] *Reno v. ACLU, 521 U.S. 844 (1997)*, Supreme Court of the United States, Jun. 26, 1997. [Online]. Available: https://supreme.justia.com/cases/federal/us/521/844/case.pdf.

[4] *SP 800-63A: IAL2 Remote Identity Proofing*, Jul. 1, 2020. [Online]. Available: https://pages.nist.gov/800-63-3-Implementation-Resources/63A/ial2remote/.

[5] "Texas Secures Settlement with Operator of Major Pornography Website, Ensuring Compliance with Texas Law," Office of the Attorney General. [Online]. Available: https://www.texasattorneygeneral.gov/news/releases/texas-secures-settlement-operator-major-pornography-website-ensuring-compliance-texas-law.

[6] "Attorney General Ken Paxton Sues Major Pornography Distributor for Violating Texas Age Verification Laws," Office of the Attorney General. [Online]. Available: https://www.texasattorneygeneral.gov/news/releases/attorney-general-ken-paxton-sues-major-pornography-distributor-violating-texas-age-verification-laws.

[7] R. Louis. "xHamster Settles Texas AV Lawsuit, Pays $120,000," XBIZ. [Online]. Available: https://www.xbiz.com/news/293771/xhamster-settles-texas-av-lawsuit-pays-120000.

[8] T. Carpenter, "Anti-porn center files four Kansas lawsuits alleging violation of state's age-verification law," *The Lawrence Times*, May 14, 2025. [Online]. Available: https://lawrencekstimes.com/2025/05/14/anti-porn-center-files-lawsuits/.

[9] D. Fox. "Indiana AG Sends Cease and Desist Letters to Adult Sites Over AV," XBIZ. [Online]. Available: https://www.xbiz.com/news/286657/indiana-ag-sends-cease-and-desist-letters-to-adult-sites-over-av.

[10] S. Cole. "Florida Sues Hentai Site and High-Risk Payment Processor for Not Verifying Ages," 404 Media. [Online]. Available: https://web.archive.org/web/20250923135552/https://www.404media.co/florida-lawsuit-nutaku-spicevids-segpay/.

[11] M. Nouwens, I. Liccardi, M. Veale, D. Karger, and L. Kagal, "Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20, New York, NY, USA: Association for Computing Machinery, Apr. 23, 2020, pp. 1–13, ISBN: 978-1-4503-6708-0, DOI: 10.1145/3313831.3376321.

[12] I. Sanchez-Rola et al., "Can I Opt Out Yet? GDPR and the Global Illusion of Cookie Control," in *Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security*, ser. Asia CCS '19, New York, NY, USA: Association for Computing Machinery, Jul. 2, 2019, pp. 340–351, ISBN: 978-1-4503-6752-3, DOI: 10.1145/3321705.3329806.

[13] H. Ou, Y. Fang, Y. Guo, W. Guo, and C. Huang, "ViopolicyDetector: An Automated Approach to Detecting GDPR Suspected Compliance Violations in Websites," in *Proceedings of the 25th International Symposium on Research in Attacks, Intrusions and Defenses*, ser. RAID '22, New York, NY, USA: Association for Computing Machinery, Oct. 26, 2022, pp. 409–430, ISBN: 978-1-4503-9704-9, DOI: 10.1145/3545948.3545952.

[14] M. Kretschmer, J. Pennekamp, and K. Wehrle, "Cookie Banners and Privacy Policies: Measuring the Impact of the GDPR on the Web," *ACM Trans. Web*, vol. 15, no. 4, 20:1–20:42, Jul. 14, 2021, ISSN: 1559-1131, DOI: 10.1145/3466722.

[15] M. Hils, D. W. Woods, and R. Böhme, "Measuring the Emergence of Consent Management on the Web," in *Proceedings of the ACM Internet Measurement Conference*, ser. IMC '20, New York, NY, USA: Association for Computing Machinery, Oct. 27, 2020, pp. 317–332, ISBN: 978-1-4503-8138-3, DOI: 10.1145/3419394.3423647.

[16] D. Bollinger, K. Kubicek, C. Cotrini, and D. Basin, "Automating Cookie Consent and GDPR Violation Detection," presented at the 31st USENIX Security Symposium (USENIX Security 22), 2022, pp. 2893–2910, ISBN: 978-1-939133-31-1. [Online]. Available: https://www.usenix.org/conference/usenixsecurity22/presentation/bollinger.

[17] A. Bouhoula, K. Kubicek, A. Zac, C. Cotrini, and D. Basin, "Automated Large-Scale Analysis of Cookie Notice Compliance," presented at the 33rd USENIX Security Symposium (USENIX Security 24), 2024, pp. 1723–1739, ISBN: 978-1-939133-44-1. [Online]. Available: https://www.usenix.org/conference/usenixsecurity24/presentation/bouhoula.

[18] M. Nouwens, J. B. Kristensen, K. Maalt, and R. Bagge, "A Cross-Country Analysis of GDPR Cookie Banners and Flexible Methods For Scraping Them," in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, ser. CHI '25, New York, NY, USA: Association for Computing Machinery, Apr. 25, 2025, pp. 1–28, ISBN: 979-8-4007-1394-1, DOI: 10.1145/3706598.3713648.

[19] V. H. Tran, A. Mehrotra, M. Chetty, N. Feamster, J. Frankenreiter, and L. Strahilevitz, "Measuring Compliance with the California Consumer Privacy Act Over Space and Time," in *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, ser. CHI '24, New York, NY, USA: Association for Computing Machinery, May 11, 2024, pp. 1–19, ISBN: 979-8-4007-0330-0, DOI: 10.1145/3613904.3642597.

[20] M. V. Nortwick and C. Wilson, "Setting the Bar Low: Are Websites Complying With the Minimum Requirements of the CCPA?" *Proceedings on Privacy Enhancing Technologies*, 2022, ISSN: 2299-0984. [Online]. Available: https://petsymposium.org/popets/2022/popets-2022-0030.php.

[21] M. Zhang, W. Meng, Y. Zhou, and K. Ren, "CSChecker: Revisiting GDPR and CCPA Compliance of Cookie Banners on the Web," in *Proceedings of the IEEE/ACM 46th International Conference on Software Engineering*, ser. ICSE '24, New York, NY, USA: Association for Computing Machinery, Apr. 12, 2024, pp. 1–12, ISBN: 979-8-4007-0217-4, DOI: 10.1145/3597503.3639159.

[22] I. Reyes et al., ""Is Our Children's Apps Learning?" Automatically Detecting COPPA Violations," *ConPro*, 2017, DOI: 20.500.12761/350.

[23] A. Ablove et al., "Digital Discrimination of Users in Sanctioned States: The Case of the Cuba Embargo," presented at the 33rd USENIX Security Symposium (USENIX Security 24), 2024, pp. 3909–3926, ISBN: 978-1-939133-44-1. [Online]. Available: https://www.usenix.org/conference/usenixsecurity24/presentation/ablove.

[24] R. Kumar, A. Virkud, R. S. Raman, A. Prakash, and R. Ensafi, "A Large-scale Investigation into Geodifferences in Mobile Apps," presented at the 31st USENIX Security Symposium (USENIX Security 22), 2022, pp. 1203–1220, ISBN: 978-1-939133-31-1. [Online]. Available: https://www.usenix.org/conference/usenixsecurity22/presentation/kumar.

[25] A. McDonald et al., "403 Forbidden: A Global View of CDN Geoblocking," in *Proceedings of the Internet Measurement Conference 2018*, ser. IMC '18, New York, NY, USA: Association for Computing Machinery, Oct. 31, 2018, pp. 218–230, ISBN: 978-1-4503-5619-0, DOI: 10.1145/3278532.3278552.

[26] R. S. Raman, A. Stoll, J. Dalek, R. Ramesh, W. Scott, and R. Ensafi, "Measuring the Deployment of Network Censorship Filters at Global Scale," in *Proceedings 2020 Network and Distributed System Security Symposium*, San Diego, CA: Internet Society, 2020, ISBN: 978-1-891562-61-7, DOI: 10.14722/ndss.2020.23099.

[27] P. Pearce, R. Ensafi, F. Li, N. Feamster, and V. Paxson, "Augur: Internet-Wide Detection of Connectivity Disruptions," in *2017 IEEE Symposium on Security and Privacy (SP)*, May 2017, pp. 427–443, DOI: 10.1109/SP.2017.55.

[28] B. VanderSloot, A. McDonald, W. Scott, J. A. Halderman, and R. Ensafi, "Quack: Scalable Remote Measurement of Application-Layer Censorship," in *27th USENIX Security Symposium (USENIX Security 18)*, Baltimore, MD: USENIX Association, Aug. 2018, pp. 187–202. [Online]. Available: https://www.usenix.org/conference/usenixsecurity18/presentation/vandersloot.

[29] D. Lang, B. Listyg, B. V. Ross, A. V. Musquera, and Z. Sanderson. "Did Age Verification Bills Change Search Behavior? A Pre-Registered Synthetic Control Study. "[Online]. Available: https://osf.io/z83ev, pre-published.

[30] D. Xue, A. Ablove, R. Ramesh, G. K. Danciu, and R. Ensafi, "Bridging Barriers: A Survey of Challenges and Priorities in the Censorship Circumvention Landscape," presented at the 33rd USENIX Security Symposium (USENIX Security 24), 2024, pp. 2671–2688, ISBN: 978-1-939133-44-1. [Online]. Available: https://www.usenix.org/conference/usenixsecurity24/presentation/xue-bridging.

[31] L. Zhou et al., "POLICYCOMP: Counterpart Comparison of Privacy Policies Uncovers Overbroad Personal Data Collection Practices," presented at the 32nd USENIX Security Symposium (USENIX Security 23), 2023, pp. 1073–1090, ISBN: 978-1-939133-37-3. [Online]. Available: https://www.usenix.org/conference/usenixsecurity23/presentation/zhou-lu.

[32] S. Wilson et al., "Analyzing Privacy Policies at Scale," *ACM Transactions on the Web (TWEB)*, Dec. 4, 2018, DOI: 10.1145/3230665.

[33] R. Nokhbeh Zaeem and K. S. Barber, "A Large Publicly Available Corpus of Website Privacy Policies Based on DMOZ," in *Proceedings of the Eleventh ACM Conference on Data and Application Security and Privacy*, ser. CODASPY '21, New York, NY, USA: Association for Computing Machinery, Apr. 26, 2021, pp. 143–148, ISBN: 978-1-4503-8143-7, DOI: 10.1145/3422337.3447827.

[34] B. VanderSloot, "Device-based Age Verification," Aug. 8, 2025. [Online]. Available: https://www.ietf.org/slides/slides-agews-paper-device-based-age-verification-00.pdf.

[35] ASACP. "RTA - Parental Control Software. "[Online]. Available: https://www.rtalabel.org/?content=parents.

[36] Y. Xu, T. Price, J.-M. Frahm, and F. Monrose, "Virtual U: Defeating Face Liveness Detection by Building Virtual Models from Your Public Photos," presented at the 25th USENIX Security Symposium (USENIX Security 16), 2016, pp. 497–512, ISBN: 978-1-931971-32-4. [Online]. Available: https://www.usenix.org/conference/usenixsecurity16/technical-sessions/presentation/xu.

[37] B. Burgess, A. Ginsberg, E. W. Felten, and S. Cohney, "Watching the watchers: Bias and vulnerability in remote proctoring software," presented at the 31st USENIX Security Symposium (USENIX Security 22), 2022, pp. 571–588, ISBN: 978-1-939133-31-1. [Online]. Available: https://www.usenix.org/conference/usenixsecurity22/presentation/burgess.

[38] P. Vallina, Á. Feal, J. Gamba, N. Vallina-Rodriguez, and A. F. Anta, "Tales from the Porn: A Comprehensive Privacy Analysis of the Web Porn Ecosystem," in *Proceedings of the Internet Measurement Conference*, ser. IMC '19, New York, NY, USA: Association for Computing Machinery, Oct. 21, 2019, pp. 245–258, ISBN: 978-1-4503-6948-0, DOI: 10.1145/3355369.3355583.

[39] Y. V. Lin et al., "Carded by the Internet: Measuring User Responses to Online Age Assurance Mechanisms," presented at the Twenty-First Symposium on Usable Privacy and Security, Seattle, WA, USA, 2025. [Online]. Available: https://www.usenix.org/conference/soups2025/presentation/lin-poster.

[40] Y. Yao, S. McCollum, Z. Sun, and Y. Zhang, "Easy As Child's Play: An Empirical Study on Age Verification of Adult-Oriented Android Apps," in *Proceedings of the 34th USENIX Security Symposium (USENIX)*, 2025. [Online]. Available: https://www.usenix.org/conference/usenixsecurity25/presentation/yao-yifan.

[41] "Downloadable snapshots of the Chrome Top Million Websites pulled from public CrUX data in Google BigQuery," GitHub. [Online]. Available: https://github.com/zakird/crux-top-lists/blob/df19e35d556d8b9239b0cad9b600a1c46b889d54/data/global/202504.csv.gz.

[42] S. Mayhew. "New report forecasts the global face liveness detection market, highlights leading firms," EIN News. [Online]. Available: https://www.einnews.com/pr_news/793243715/new-report-forecasts-the-global-face-liveness-detection-market-highlights-leading-firms.

[43] Age Check Certification Scheme. "Registry Archive," Age Check Certification Scheme. [Online]. Available: https://web.archive.org/web/20250914233206/https://accscheme.com/registry/.

[44] "Downloadable snapshots of the Chrome Top Million Websites pulled from public CrUX data in Google BigQuery," GitHub. [Online]. Available: https://github.com/zakird/crux-top-lists/blob/df19e35d556d8b9239b0cad9b600a1c46b889d54/data/global/202506.csv.gz.

[45] K. Ruth, D. Kumar, B. Wang, L. Valenta, and Z. Durumeric, "Toppling top lists: Evaluating the accuracy of popular website lists," in *Proceedings of the 22nd ACM Internet Measurement Conference*, Nice, France and New York, NY, USA: ACM, Oct. 25, 2022, pp. 374–387, ISBN: 978-1-4503-9259-4, DOI: 10.1145/3517745.3561444.

[46] Leonard Richardson, *Beautiful Soup: We called him Tortoise because he taught us.* Version 3, n.d. [Online]. Available: https://www.crummy.com/software/BeautifulSoup/.

[47] "IAB Tech Lab Content Taxonomy," IAB Tech Lab. [Online]. Available: https://iabtechlab.com/standards/content-taxonomy/.

[48] "Klazify - Free Content Classification API. Turn any email or URL into a full company profile.," Klazify. [Online]. Available: https://www.klazify.com/.

[49] "Yoti Supermarket. "[Online]. Available: https://yoti.world/supermarket-avs/.

[50] "End of Life policy," Yoti developer documentation. [Online]. Available: https://perma.cc/2FTG-BUUT.

[51] "Overview," Yoti developer documentation. [Online]. Available: https://developers.yoti.com/age-verification-headless.

[52] "RTA - Parental Control Software - Website Label," RTA: Restricted To Adults. [Online]. Available: https://www.rtalabel.org/.

[53] *Protecting Georgia's Children on Social Media Act of 2024*, GA SB 351, 2024. [Online]. Available: https://www.legis.ga.gov/legislation/66023.

[54] *Liability for Allowing Minors to Access Pornographic Material*, TX HB 1181, 2023. [Online]. Available: https://capitol.texas.gov/BillLookup/History.aspx?LegSess=88R&Bill=HB1181.

[55] "Ipify - A Simple Public IP Address API," ipify. [Online]. Available: https://www.ipify.org/.

[56] "IPinfo — The Trusted IP Data Provider for Developers & Enterprises," IPinfo. [Online]. Available: https://ipinfo.io/.
[57] Vladimir Mandic. "@vladmandic/face-api," npm. [Online]. Available: https://www.npmjs.com/package/@vladmandic/face-api.
[58] "ID Verification," Yoti developer documentation. [Online]. Available: https://developers.yoti.com/age-verification/identity-verification.
[59] M. Hayden. "A new future for icanhazip," Major Hayden. [Online]. Available: https://major.io/p/a-new-future-for-icanhazip/.
[60] "Stripe Documentation. "[Online]. Available: https://docs.stripe.com/.
[61] "Portal View," Yoti developer documentation. [Online]. Available: https://developers.yoti.com/age-verification/portal-view.
[62] "Fingerprint — Identify Every Web Visitor & Mobile Device," FingerprintJS. [Online]. Available: https://fingerprint.com/.
[63] "Secure image capture," Yoti developer documentation. [Online]. Available: https://perma.cc/MDY7-CHSB.
[64] "Age Verification Privacy Policy [English] · Yoti," Yoti. [Online]. Available: https://perma.cc/4D74-5ZS9.
[65] "Yoti and EasyID App Privacy Policy," Yoti. [Online]. Available: https://www.yoti.com/privacy/app/.
[66] "Age Verification Systems – ID Verification Services," Aristotle. [Online]. Available: https://integrity.aristotle.com/.
[67] "Veratad: Age Verification — Identity Verification — Fraud Prevention," Veratad. [Online]. Available: https://veratad.com.
[68] "Biometrics Policy [United States]," Yoti. [Online]. Available: https://perma.cc/HPH9-9AQ2.
[69] M. Frigo and a. shelat. "Anonymous credentials from ECDSA. "[Online]. Available: https://eprint.iacr.org/2024/2010, pre-published.
[70] *Digital Credentials*, Apr. 17, 2026. [Online]. Available: https://www.w3.org/TR/digital-credentials/.
[71] P. Eckersley, "How Unique Is Your Web Browser?" In *Privacy Enhancing Technologies*, M. J. Atallah and N. J. Hopper, Eds., vol. 6205, Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 1–18, ISBN: 978-3-642-14527-8, DOI: 10.1007/978-3-642-14527-8_1.
[72] Department for Science, Innovation and Technology and The Rt Hon Peter Kyle MP. "Keeping children safe online: Changes to the Online Safety Act explained," GOV.UK. [Online]. Available: https://www.gov.uk/government/news/keeping-children-safe-online-changes-to-the-online-safety-act-explained.
[73] OAIC. "Social Media Minimum Age," Office of the Australian Information Commissioner. [Online]. Available: https://www.oaic.gov.au/privacy/your-privacy-rights/social-media-minimum-age.
[74] N. Totenberg and A. Ononye, "Supreme Court allows Mississippi social media law to go into effect," *NPR*, Aug. 14, 2025. [Online]. Available: https://www.npr.org/2025/08/14/nx-s1-5482925/scotus-netchoice.
[75] The Bluesky Team. "Our Approach to Age Assurance," Bluesky. [Online]. Available: https://bsky.social/about/blog/09-10-2025-age-assurance-approach.
[76] J. R. Marsh, "Brief Of Amicus Curiae Yoti Ltd. in Support of Respondent," Nov. 2024. [Online]. Available: https://www.supremecourt.gov/DocketPDF/23/23-1122/332630/20241122163435761_23-1122%20Amicus%20Brief.pdf.
[77] S. Minocha, *Artifact for 'Papers, Please: A First Look at Age Verification on the Web'*, version 1.0.0, Zenodo, Apr. 11, 2026, DOI: 10.5281/zenodo.19515393.

# Appendix A.
# Code Snippets and Tables


```
"ip": "[redacted]",
"hostname": "[redacted]",
"city": "Dallas",
"region": "Texas",
"country": "US",
"loc": "32.7831,-96.8067",
"org": "AS13213 UK-2 Limited",
"postal": "75201",
"timezone": "America/Chicago",
"readme": "https://ipinfo.io/missingauth"
```


Listing 1: Partially-redacted IPInfo response


```
"manufacture_name": "",
"model_name": "",
"os_name": "Linux",
"os_version": "6.12.51",
"client_version": "",
"product": 3,
"browser_name": "Chrome",
"browser_version": "141.0.7390.76",
"locale": "en-GB"
```


Listing 2: The device, OS, and browser metadata sent to Yoti as part of an ID verification attempt

TABLE 4. AGE VERIFICATION PROVIDERS CONSIDERED, THE BEST DATA SOURCES WE FOUND, AND HOW WE CONSTRUCTED PATTERNS

| Provider | Best Data | Domains | `<script>` | IDs | Classes | Iframes | `<a>` | `<link>` | Images | JS | Portal | Wildcard | Misc |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Acuant | Demo | ✓ | | ✓ | | | | | | | | | |
| AgeChecked | Demo | ✓ | ✓ | | | | | ✓ | | | | | |
| AgeChecker | Deployments | | ✓ | | ✓ | | | | | | | | |
| AgeGo | Deployments | ✓ | ✓ | | | | | | | | | | |
| AgeVerif | Deployments | ✓ | ✓ | | | | | | ✓ | | | | |
| Aristotle | Docs | ✓ | | | | | | | | | | ✓ | |
| Au10Tix | Deployments | ✓ | | | ✓ | | | | | ✓ | | | |
| Authenteq | Deployments | ✓ | | | | | | | | | | | ✓ |
| Aware | Docs | ✓ | | | | | | | | ✓ | | ✓ | |
| BioID | Docs | ✓ | | | | | | | | | | | ✓ |
| Blue Check | Docs | ✓ | | | | | | | | | | | |
| BorderAge | None | ✓ | | | | | | | | | | | |
| ComplyCube | Docs | ✓ | | | | | | | | | | | ✓ |
| Daon | None | ✓ | | | | | | | | | | | |
| EarthID | None | ✓ | | | | | | | | | | | |
| Emblem | Deployments | ✓ | ✓ | | ✓ | | ✓ | | | | | | |
| GBG | None | ✓ | | | | | | | | | | | |
| HyperVerge | Docs | ✓ | | | | | | | | ✓ | | | |
| ID.me | Deployments | ✓ | ✓ | ✓ | | | ✓ | | | | | | |
| Idemia | None | | | | | | | | | | | ✓ | |
| Idenfy | Docs | ✓ | | | | ✓ | | | | | | | |
| Identomat | None | | | | | | | | | | | ✓ | |
| IDNow | Demo | ✓ | | | | | | | | | | | |
| Idology | Docs | | | ✓ | ✓ | | | | | | | | |
| ID R&D | None | | | | | | | | | | | ✓ | |
| Ike | None | ✓ | | | | | | | | | | | |
| Incode | Deployments | ✓ | | ✓ | | | | | ✓ | ✓ | ✓ | | ✓ |
| IProov | Docs | ✓ | | ✓ | ✓ | | | | | ✓ | | | ✓ |
| Jumio | Docs | | | | ✓ | | | | | ✓ | | | |
| KWS | None | ✓ | | | | | | | | | | | |
| Luciditi | Demo | ✓ | ✓ | ✓ | ✓ | | | ✓ | | | | | ✓ |
| Mitek | None | | | | | | | | | | | ✓ | |
| Ondato | Deployments | ✓ | | | | | | | | | | | ✓ |
| 1Account | Deployments | ✓ | ✓ | ✓ | | ✓ | | ✓ | | ✓ | | | |
| OneID | Docs | ✓ | | | | | | | | | | | |
| OneSpan | None | | | | | | | | | | | ✓ | |
| Onfido | Deployments | ✓ | | | | | | | | | | | |
| Opale | Demo | ✓ | ✓ | ✓ | | | | | | ✓ | | | |
| Paravision | None | ✓ | | | | | | | | | | ✓ | |
| Persona | Deployments | ✓ | | ✓ | | ✓ | ✓ | | | ✓ | | | |
| Privado | None | | | | | | | | | | | ✓ | |
| Privately | Demo | ✓ | | | | ✓ | | | | | | | |
| Privo | Docs | ✓ | ✓ | | | | | | | ✓ | | | |
| Regula | Docs | ✓ | ✓ | | ✓ | | | | | | | | ✓ |
| RoC | None | ✓ | | | | | | | | | | | ✓ |
| Scytales | None | | | | | | | | | | | ✓ | |
| ShareRing | Demo | ✓ | | | | | | | ✓ | | | | ✓ |
| Shufti | None | ✓ | | | | | | | | | | ✓ | |
| Socure | Deployments | ✓ | ✓ | | | | | ✓ | | ✓ | | ✓ | |
| Sumsub | Deployments | ✓ | ✓ | | | | | ✓ | | | | | |
| Surepass | None | ✓ | | | | | | | | | | | |
| TokenOfTrust | Docs | ✓ | | | | | | | | | | | |
| Trulioo | Docs | ✓ | | | | | | | | ✓ | | | |
| Trustmatic | None | ✓ | | | | | | | | | | ✓ | |
| Veratad | Deployments | ✓ | | ✓ | | ✓ | ✓ | | | ✓ | | | |
| Veridas | Deployments | ✓ | | | | | | | | ✓ | | | |
| Veriff | Deployments | ✓ | ✓ | | ✓ | | | | | | | | |
| VerifyMyAge | Deployments | ✓ | ✓ | | ✓ | | ✓ | | ✓ | | ✓ | | |
| Yoti | Deployments | ✓ | | ✓ | | ✓ | ✓ | | | ✓ | ✓ | | |
| YouVerse | Docs | ✓ | | | | | | | | ✓ | | | |

## Appendix B. Meta-Review

The following meta-review was prepared by the program committee for the 2026 IEEE Symposium on Security and Privacy (S&P) as part of the review process as detailed in the call for papers.

### B.1. Summary

This paper presents a first look at the web age-verification ecosystem in the United States, combining a large-scale measurement of third-party age verification providers (AVPs) across three states (Texas, Georgia, and New York) with a detailed reverse-engineering and security/privacy analysis of the dominant provider, Yoti. The paper situates its findings within the context of recent US Supreme Court rulings on age verification mandates.

### B.2. Scientific Contributions

- Independent Confirmation of Important Results with Limited Prior Research.
- Provides a New Data Set For Public Use.
- Creates a New Tool to Enable Future Science.
- Provides a Valuable Step Forward in an Established Field.
- Establishes a New Research Direction.

### B.3. Reasons for Acceptance

1) The paper provides a valuable step forward in an established field and establishes a new research direction. Age verification on the web is a rapidly emerging policy phenomenon with direct constitutional relevance following the 2025 Supreme Court decision upholding Texas' age verification law. This paper is the first to characterize the AVP ecosystem empirically and connects "what laws require" to "what systems actually do," making it directly relevant to ongoing legislative and judicial debate.
2) The paper independently confirms important results with limited prior research and provides a new dataset for public use. The large-scale crawl across three states documents state-dependent deployment patterns, provider market concentration around Yoti, and fourth-party data exposure risks that typical site-to-provider framings miss. The authors commit to releasing their dataset and deobfuscated code openly.
3) The paper creates a new tool to enable future science. The authors develop and publish provider-detection heuristics, reversed client-side code from Yoti, and a proof-of-concept Chrome extension demonstrating end-to-end bypass of Yoti's Secure Image Capture (SICAP) mechanism. These artifacts provide a foundation for follow-on measurement and security research in this area.