

# This Is My Private Business!

## Privacy Risks on Adult Websites

Pelayo Vallina<sup>†◊</sup>, Julien Gamba<sup>†◊</sup>, Alvaro Feal<sup>†◊</sup>, Narseo Vallina-Rodriguez<sup>† ‡</sup>, Antonio Fernández Anta<sup>†</sup>  
<sup>†</sup> IMDEA Networks Institute, <sup>◊</sup>Universidad Carlos III de Madrid, <sup>‡</sup> ICSI

**Abstract**—Millions of users from all over the world access pornographic content on the web. However, as of today, there’s a dearth of knowledge about the porn web ecosystem and, in particular, about the privacy risks to which porn web users get exposed. As a matter of fact, porn websites can infer sensitive private information from users, including their sexual orientation and preferences, which can be later linked to their identity. As a result, many porn web consumers resort to ad blockers and private browsing modes to protect their identity when browsing adult websites. In this work in progress, we aim to expose the porn web ecosystem hoping to analyze aspects like their age verification mechanisms — which are required in certain regulatory frameworks as the British one —, their user tracking and data collection practices, and their privacy policies. By leveraging simple methodologies, we have already obtained relevant preliminary results: (1) we have identified a significant presence of third-party tracking services on porn websites; (2) we have discovered the use of both first-party and third-party cookies, most times installed without the user consent; and (3) we have found a concerning lack of age verification mechanism to prevent the access of minors to porn websites. In the future, we plan to study in depth the many stakeholders in this industry as well as the effectiveness of anti-tracking mechanisms such as ad blockers, VPN services, and safe browsing modes.

**Index Terms**—Privacy, Adult Content Websites, Porn Web

**Tipo de contribución:** Investigación en desarrollo

### I. INTRODUCTION

Porn sites are among the most visited websites globally. According to the Alexa Rank (as of March 6th, 2018) two porn sites are among the top 40 most popular websites. Despite their importance in terms of the number of on-line users, the porn web ecosystem remains mostly opaque and unexplored. Only a number of isolated efforts, mostly executed by cyber-security firms, have revealed concerning aspects of this industry. For instance, Kaspersky labs revealed that PornHub.com, one of the most popular porn websites, was responsible for distributing malware [1]. Moreover, porn websites are considered sensitive websites, due to their capacity to obtain and infer private information, such as a user’s sexual orientation, habits, and behavior.

Many web users want to remain anonymous when browsing pornographic content, and resort to anti-tracking solutions or safe browsing modes to protect their privacy. However, these mechanisms may not offer them total protection. That is the case of anti-tracking solutions based on incomplete blacklists [2], or on-line services using advanced user and IP fingerprinting. Despite these issues, the academic efforts have ignored so far the tangible privacy risks of porn websites for the end-user. Instead, they have focused on measuring

and characterizing how users browse and interact with those services [3], and their integration with social networks [4].

Policy makers have made important steps for regulating adult websites and protecting consumer’s privacy. According to the European regulation and the upcoming one (GDPR and ePrivacy directives), it is illegal to obtain and process sensitive data without explicit consent from the users. UK legislators are putting efforts to prevent minors from accessing porn websites [5]. In their new regulatory framework, online porn services will have to verify the user’s age, for instance requiring the users to provide their credit card or their national id card. Many users may consider such mechanisms intrusive, and a mean for de-anonymizing their activities. Consequently, it is unclear whether the pornographic industry will effectively implement age verification mechanisms that could negatively impact their number of daily visitors.

In this project, we will develop methodologies and techniques to uncover the online porn ecosystem. Specifically, we aim to answer the following questions:

- Can porn websites harm user’s privacy?
- Do they use third-party tracking services?
- Do they interact with each other via affiliate programs?
- Are safe browsing mode and anti-trackers effective to guarantee user’s anonymity?
- Do porn websites use age verification mechanisms?

### II. METHODOLOGY

The first step in our methodology is selecting a large and representative corpus of porn websites. In our case we have selected 86 relevant free porn websites obtained from four different sources, specialized in recommending and classifying pornographic content<sup>1</sup>. Then, we use a Selenium-based crawler using a Firefox browser equipped with two add-ons: a VPN [6] for accessing the content from vantage points located in four countries (Spain, Russia, USA, and UK), to analyze geographical differences and Lightbeam [7], a tool to analyze the interactions between first and third-party services present on a given website as Falahrastegar *et al.* implemented in their previous work [8]. The crawler fetches the HTML content of the landing page and gets a screenshot of the main page. We complement our headless browser with ZBrowse [9], a command line tool based on the headless Chromium browser that allows analyzing dependencies on websites such as network

<sup>1</sup><https://www.alexa.com/topsites/category/Top/Adult>, <http://toppornsites.com>, <http://mypornbible.com>, <http://www.only4adults.net>.

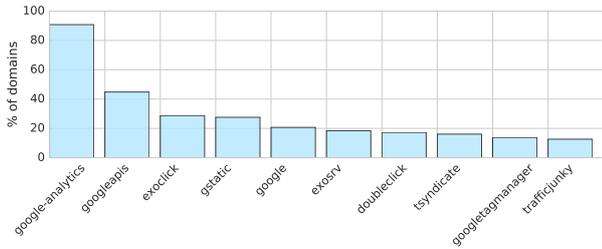


Fig. 1. Top 10 most common third-party services across porn websites.

TABLE I  
PRESENCE OF COOKIES IN OUR CORPUS OF PORN WEBSITES. WE DISTINGUISH BETWEEN FIRST- AND THIRD-PARTY COOKIES.

Feature	#Porn Sites ( $N = 86$ )	Cookie consent
First-party Cookies	42	4
Third-party Cookies	36	0
First and Third-party Cookies	20	0

request and object dependencies. We only used ZBrowse from a vantage point located in Spain. Both crawlers create fresh browsing sessions between each page crawl to minimize bias.

### III. PRELIMINARY RESULTS

**The online porn network:** The analysis of the HTML code of the porn websites revealed interesting links between various services. This suggests the presence of a large number of affiliated or subsidiary websites, all of them associated with the company MindGeek, which has a dominating position in the online porn industry.

**Third-party services:** Figure 1 presents the 10 most common third-party services over a total of 168 different ones identified in our web crawls. 19 of them are associated with advertising and tracking services, which we found on 79 porn websites. Just Google Analytics is present in 90% of the adult-content domains. We can also identify the presence of Exoclick in 30% of the porn websites. This is a Spanish ad network who, according to the BBC, specializes in distributing ads for sites and products for adults such as sexual content and gambling [10], a market intentionally ignored by Google and Facebook. This is a response to the restrictions set by countries like Germany, Russia, and China among many others who ban the distribution of adult and sensitive content through ad networks [11]. Surprisingly, Google is still present and capable of tracking porn web users through their analytics products.

**HTTP Cookies:** According to the European legislation, it is mandatory for all websites operating in the EU to inform users about the use of third-party cookies. In Table I we represent the presence of cookies in our set of porn sites. We found 36 porn webs installing third-party cookies without user’s consent. Also, we found 42 of these domains installing first-party cookies. Only 4 sites inform users but in these cases there were only of first-party cookies. However, according to the European legislation, it is not necessary to inform and request permission from users to install first party cookies as long as it is required to deliver the expected service [12].

**Age verification:** We leverage our screenshots to manually inspect the presence of age-verification mechanisms in our corpus of porn sites. For this specific study, we used the VPN service to identify and report differential behavior of those porn sites due to regulatory requirements. In the USA, Spain and the UK there is only one porn site that has implemented age verification mechanism. However, in Russia we have identified 6 of them. The mechanisms to prevent access to minors are far from effective: they are simple pop ups informing that the content is restricted to adults and a checkbox for verification. We didn’t find any variation in other EU countries. We can conclude that, as of today, the majority of the websites lack of effective age verification mechanisms to comply with current law requirements (as in Russia) and future ones (as in the UK).

### IV. CONCLUSIONS

The online porn ecosystem remains opaque due to the lack of systematic and comprehensive studies focused on this ecosystem. In this work, we have seen that by applying simple techniques, it is possible to audit the regulatory compliance of these websites and obtain striking results. We have found the use of cookies by those websites without the user’s consents and also the lack of age verification mechanism, which in some countries is already mandatory to prevent minor’s access to such sensitive content. This extended abstract only presents our preliminary results in our larger efforts towards illuminating comprehensively the online porn web ecosystem and its associated privacy risks for end users. We plan to extend this work by extending the scale of our porn site corpus, the geographic differential analysis, their privacy policies, the use of advanced user-profiling and user-tracking mechanisms and the effectiveness of anti-tracking software to protect user’s anonymity.

### REFERENCES

- [1] Kaspersky, “Malware spread through pornhub,” <https://www.kaspersky.com/blog/pornhub-malvertising/19698/>.
- [2] A. Razaghpanah, R. Nithyanand, N. Vallina-Rodriguez, S. Sundaresan, M. Allman, C. Kreibich, and P. Gill, “Apps, trackers, privacy, and regulators: A global study of the mobile tracking ecosystem,” *NDSS*, 2018.
- [3] G. Tyson, Y. Elkhatib, N. Sastry, and S. Uhlig, “Demystifying porn 2.0: A look into a major adult video streaming website,” in *Proc of ACM IMC*, 2013.
- [4] G. Tyson, Y. Elkhatib, Nishanth, and S. Uhlig, “Are people really social in porn 2.0?” in *Proc of ICWSM*, 2015.
- [5] The Guardian, “Pornography sites face uk block under enhanced age controls,” <https://www.theguardian.com/culture/2016/nov/19/pornography-sites-face-uk-block-under-enhanced-age-controls>.
- [6] Firefox, “Hoxx-vpn,” <https://addons.mozilla.org/es/firefox/addon/hoxx-vpn-proxy/>.
- [7] Firefox, “Lightbeam,” <https://addons.mozilla.org/en-US/firefox/addon/lightbeam/>.
- [8] M. Falahrastegar, H. Haddadi, S. Uhlig, and R. Mortier, “Tracking personal identifiers across the web,” in *PAM*, 2016.
- [9] ZMap, “ZBrowse,” <https://github.com/zmap/zbrowse/>.
- [10] BBC, “Exoclick, la empresa que gana mas de 100M USD con los clientes que Google no quiso,” <http://www.bbc.com/mundo/noticias-39723776>.
- [11] Google, “Advertising policies help. adult content,” <https://support.google.com/adwordspolicy/answer/6023699>.
- [12] EU, “Cookies,” <http://ec.europa.eu/ipg/basics/legal/cookies/>.