



Algorithms of arousal: an investigation of Pornhub videos and advertisements recommended to browser sessions without prior viewing

Citation

Toth-Rohonyi, Ivan. 2025. Algorithms of arousal: an investigation of Pornhub videos and advertisements recommended to browser sessions without prior viewing. Bachelors Thesis, Harvard University Engineering and Applied Sciences.

Link

<https://dash.harvard.edu/handle/1/42731099>

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#)

**Algorithms of arousal: an investigation of Pornhub videos and advertisements
recommended to browser sessions without prior viewing history**

Ivan Toth-Rohonyi

I'm thankful for the Harvard College Research Program for its funding for this project

Abstract:

This paper explores the characteristics of videos and advertisements recommended to browser sessions without prior viewing on Pornhub, focusing on the relationship between aggression and the race of female performers. To do so, it relies on a manual review of ads and on data obtained through the use of scripted browser instances visiting porn sites to model what recommendations are present in actual visiting users' experience. The specific question it aims to answer is how the level and kind of aggression varies based on the female performer's race/ethnicity in videos recommended to fresh browser sessions. The paper is inspired by recent writings by Adler (2024) and Srinivasan (2021) that draw attention to the harms the way porn today is consumed causes. Both discuss the view that sexual tastes are not self-determined and biologically essential but influenced by porn. As such, porn's harm is conceptualized both in the direct and horrible toll its production can take but also in the way pornography that is widely consumed influences sexual norms, and as they are closely intertwined with sexuality, gender norms. The paper's main findings are that in 2025, on Pornhub the ads shown to browser profiles without prior interaction are exclusively sexuality related, targeted towards men and mainly feature white women. For the videos recommended, based on the tags and categories found on the site, the ones that feature female performers of color also have aggressive terms more often. The effect is more pronounced for search results, but overall presence of aggressive tags or categories are higher for homepage recommendations. This finding is broadly in line with previous literature, most notably Shor and Golriz's 2019 article *Gender, Race, and Aggression in Mainstream Pornography*, an important inspiration, with this paper having a significantly larger sample size but relying on external categorization for its operationalization of key categories.

Introduction

This paper explores the characteristics of videos and advertisements recommended to browser sessions without prior viewing on Pornhub, focusing on the relationship between aggression and the race of female performers. To do so, it relies on a manual review of ads and on data obtained through the use of scripted browser instances visiting porn sites to model what recommendations are present in actual visiting users' experience. The specific question it aims to answer is how the level and kind of aggression varies based on the female performer's race/ethnicity in videos recommended to fresh browser sessions. This is to investigate what types of content are prioritized by the platform's algorithms, and furthermore what types of recommended content are interconnected with racialized viewing and search history in a given browsing session.

The paper is inspired by recent writings by Adler (2024) and Srinivasan (2021) that draw attention to the harms the way porn today is consumed causes. Both discuss the view that sexual tastes are not self-determined and biologically essential but influenced by porn. As such, porn's harm is conceptualized both in the direct and horrible toll its production can take but also in the way pornography that is widely consumed influences sexual norms, and as they are closely intertwined with sexuality, gender norms. Both authors argue social construction of sexuality through porn is especially prevalent in the current pornography landscape where porn is easily available yet most access to it is mediated by a few giant platforms. (Adler 2024, Srinivasan 2021) This line of thought builds on similar arguments made earlier by many feminist scholars, most notably MacKinnon (1989). Importantly, MacKinnon sometimes goes further in arguing all porn is inherently harmful, while for this paper I am focusing specifically on describing characteristics

widely consumed porn has today and how these effects are influenced by platforms, without a detailed discussion of the possible effects of such depictions.

To understand the importance of Pornhub and similar sites it's important to consider the current pornography landscape. Around the world, by far the largest share of pornographic consumption happens online on so-called tube sites. (Adler 2024). Each of these large pornographic sites hosts hundreds of thousands to millions of videos and based on their browsing patterns which involve clicking on many videos during a single session. Morichetta et al. (2021) for example report a median of 5-6 web pages visited for web pornography sessions based on traffic measurement through a Point-of-Presence (PoP) of a European ISP. As such, it seems reasonable to assume a majority visitors don't arrive with a particular video in mind. Thus, what determines the exact video they watch is the site's recommendation algorithms, which determine what videos the homepage will include, what videos are displayed after a search, and what videos are displayed on a given video's page, making what videos are watched highly dependent on the way the recommendation algorithm operates. (Adler 2024)

As far as this author is aware, there have been no credible descriptions of the exact mechanisms of Pornhub's recommendation algorithms by the site itself or by peer-reviewed articles. On a site last updated in 2024, Pornhub states "To determine the content you may find most relevant, we consider the metadata of the videos you watched, how you interacted with these videos, and popular videos in your region based on views, ratings, comments, favorite markings, and total time watched". However, based on external testing of the system in 2021 by [pornhub.tracking.exposed](#) (poTREX), it seems reasonable to assume that it is some form of hybrid recommender system, with recommendations using both content- and collaborative filtering, relying on both the tags categories and characteristics available about videos and on how previous users have interacted

with them. Important to note is that the site's layout was somewhat different at the time of the study. . poTREX has also reported of the system favoring new videos, but this blog post : <https://cz.pornhub.com/blog/grow-your-audience-with-our-updated-algorithm> indicates that this characteristic of the algorithm has likely been since changed.

Thus, as a result of how widespread the consumption of online pornography is around the world, porn recommendation algorithms play an important role in shaping desires, sexual scripts and in turn interactions between sexuality and racial categorizations and stereotypes. While racial and ethnic categorizations are fluid social constructions, online pornography is an environment where clear-cut racial categories are frequently used in both the titles and categorization of videos. (Rama et al. 2022) In addition, the pornography industry is one where Latina and Asian women have relatively higher visibility (Shor & Golriz 2019) – and such disproportionate representation of women of certain races/ethnicities in porn might contribute especially significantly to racial and gender stereotypes in countries where non-white women are minorities.

The current era of pornography began around 2007, when technological disruption created a completely new setting in terms of what is being watched, how it is being distributed, and who is producing it. (Adler, 2024) In 2007 the pornography industry started its transition in earnest to streaming, algorithm-driven tech platforms, the most prominent of which is Pornhub. (Adler, 2024) Similar to the popular video-sharing platform YouTube, the porn “tube sites” stream user-uploaded content that is available to virtually anyone, is almost always free, and has a seemingly unlimited supply of most genres and situations conceivable. A majority of these platforms are controlled by one, near-monopoly company known until recently as “Mindgeek”, but since renamed Aylo. Self-professedly data driven and secretive, Aylo holds a unique place in the contemporary United States and even global pornography ecosystem. (Adler, 2024) To put into context just how widespread

the use of pornographic sites is, the top three porn sites get more monthly traffic and unique visitors than several household names such as Amazon, Netflix or Yahoo, or up-and-coming sites such as TikTok, OpenAI or Zoom. (Wright et al, 2023)

Hopefully the above gives a sense of the significance of what kind of videos Pornhub or some other large tube sites' algorithm prioritizes or the exact mechanisms through which it operates. Previous quantitative scholarly research with comparable magnitude of sample sizes on the topic is scarce, while the porn conglomerates or Aylo specifically release near to no information about how the algorithm operates. There have been past explorations of race and violence in online porn through manual review of videos (Shor & Golriz 2019), of the functioning of Pornhub's homepage and the interplays between category recommendations through the use of browser extensions (Pornhub Tracking Exposed) and of the role of platformization in online porn (Rama et al. 2022). However, this paper will be -as far as I'm aware – methodologically unprecedented in both the sheer sample size of videos retrieved and in the granularity of actions -clicks, searches, views - the script is capable of modelling.

Following previous research (Shor & Golriz, 2019, Rama et al 2022) I chose to focus on Pornhub specifically because of its dominant place in the market, global reach and prominence in mainstream discourse, Pornhub and digital porn sometimes being used in discourse interchangeably. I also chose to focus on Pornhub because it gives a glimpse into the operations of Aylo, whose position is entirely unique. For a sense of just how dominant Aylo's market position is: as of 2023, Aylo accounted for 81.6% of the pornography industry revenue in the United States (Adler, 2024).

Despite the stigma around pornography in mainstream academic and journalistic discourse and the obvious care that must be taken to engage with it in a sensitive way, online pornography is a topic

that merits researcher attention due to several factors. As I detailed above, watching pornography online is incredibly widespread around the world and thus what kinds of videos are watched heavily influences our social realities. Currently, online pornography is a key medium and site of mediation for both individual sexual preferences and for broader forms of gendering and racialization of desire. Importantly for the topic that this project focuses on, pornography thus provides a particularly sharp lens into processes of racialization and identity formation that have broad sociological importance. Yet relatively little academic research has been done on this topic, especially with a quantitative focus. This is despite the fact that the recommendation algorithms of porn sites are unique both in the way they function and in the way they can be studied – as unlike other recommender systems, porn sites usually recommend content to users about whom there is very little to no past information available to the site. The primary objective of this thesis has been to gain a deeper understanding of the functioning of Pornhub’s recommendation algorithm especially as it relates to race while also creating a tool that other researchers can use for further investigations into Pornhub’s recommendation algorithms.

In line with previous reports, this paper has found that both for homepage videos and videos displayed after searching for specific terms, videos that are categorized as having non-White female performers more often have terms and categories associated with aggression. Interestingly, levels of aggression for all ethnic categories are lower after searching for them specifically when compared to their occurrences on a homepage. Though without access to data only available to Pornhub itself it is challenging to ascertain what might lead to these effect, based on my review of the advertisements currently displayed on Pornhub, I theorize that the higher levels of aggression and particular stereotypes being represented on the homepage -Asian women being more likely to be restrained for example – are a result of trying to capture the attention of a small number of

potential paying customers whose tastes are more extreme than the average user. Some of my results however also point to Pornhub's algorithm being less sensitive than previously theorized, for example when it comes to navigating back to the homepage.

In the rest of this paper, I will first provide an overview of the relevant literature. Following this, I will present my findings on the advertisements currently displayed on Pornhub. After this, I will describe the methods used to scrape the original dataset that is the main contribution of this paper, and my choices about operationalization. After presenting the results, I will discuss both the possible effects the disparities found might have on those who watch pornography on Pornhub, and possible explanations for findings that are not in line with previous research.

Literature review

Currently, porn is a key site of mediation for individual sexual preferences and as such preferences about gender and race and for broader forms of gendering and racialization of desire. (Adler, 2024,

Shor & Golriz 2019) I hope this paper will help further understanding of how the sociotechnical system of recommendation algorithms shapes how individuals are exposed to racialized aggression and racialized violence. For this, it is important to take a wider look at the way previous scholarship, most notably feminist scholarship, has theorized pornography and specifically online pornography and on how interplays of gender and race have been theorized in other contexts.

A crucial concept for theorizing about gender and race (and for identities in general) is intersectionality. First articulated by Kimberlé Crenshaw, intersectionality is a theoretical framework that examines how multiple axes of identity—such as race, gender, and class—interact to produce unique modes of discrimination and social experience. Those belonging to multiple identities are often excluded from the protections afforded to those analyzed solely through one lens – in the case of pornography, a lens that focuses on women in general. In her work on violence against Black women Crenshaw notes that they are frequently either invisible in broader discourse or hypervisible in ways that reflect longstanding racist and sexist stereotypes, motivating this thesis’s focus on the interplay of race and aggression.

Another important canonical theorist of race and gender is Angela Davis, whose work traces the historical connection between racialized sexual violence, capitalist exploitation, and gendered labor. In *Women, Race & Class* (1981), Davis illustrates how Black women’s bodies were both sites of coerced labor and targets of sexual domination during slavery. She argues that this legacy laid the groundwork for long-lasting structures of racialized sexual commodification. The project hopes to investigate in what ways can this history of objectification be traced to the way Black women depicted in current pornography – and to the degree this depiction is generalizable for all people of color.

Another important body of thought closely connected to this topic is feminist scholarship, which has long grappled with the cultural and political implications of pornography, most notably during the 1980s and early 1990s. At the center of these debates was Catharine MacKinnon's argument that pornography functions as a powerful tool of social construction, shaping gender norms, sexuality, and reinforcing patriarchal power through representational violence. This view, which was closely connected to efforts to regulate pornography was contested by other feminist thinkers who defended sexual expression and emphasized the (possible) diversity and complexity of pornographic content. Over time, especially among younger feminists, MacKinnon's position has fallen out of favor, sometimes dismissed outright as outdated or overly moralistic, though it seems to be making a resurgence in recent years. For a detailed discussion of this debate and its reinterpretation in the age of Big Tech see Adler (2024).

Regardless of the exact position one takes in relation to the above debate, porn as constructing reality through its ubiquity, especially so for younger generations seems like a highly plausible theorem. This approach of not necessarily passing moral judgement but acknowledging porn's power to shape social relations is well-articulated for example by Amia Srinivasan in her book "The Right to Sex", where she also reports other studies that on a variety of gender-related topics, from sexual behavior to attitudes about as women's roles and likelihood to report sexual assault, heightened viewership of porn is correlated with outcomes adverse to women's equality.

While an evaluation of whether porn as a medium is necessarily harmful and/or a tool of patriarchal oppression is outside the scope of this thesis, my results are in line with previous reports that heterosexual porn on the large tube sites frequently depict aggression, and that this aggression is heightened for women of color. This viewpoint by itself is not particularly new but hopefully the paper will contribute to understanding to some of the mechanisms that made Internet porn –

originally hoped for to be liberating and diversifying of sexual tastes and norms – to be by now understood to be in many ways even more heteronormative and reinforcing of oppressive structures of power than previous modes of consumption. (Adler 2024)

The interaction between gender and race in pornography is also particularly important as even outside of sexual contexts conceptualizations of race and gender can be heavily influenced by media depictions, especially so as for some groups (notably Asian and Latina women and Black men) their representation in pornography can outweigh their representation in the media or even real-world presence in a given region. (Shor & Golriz, 2019) Also important are to consider are discussions of the role of platformization and categorization in the pornography industry: Rama, et al. (2022) argue that “the combination of platform affordances and algorithmic suggestions on Pornhub significantly contributes to a heteronormative perspective on sexual desire and sexuality typical of a heterosexual, white, and hegemonic masculinity”, especially through the categorization of content, but their research falls short of investigating the particular algorithmic recommendations that facilitate and reinforce this perspective.

Unfortunately, this heteronormative perspective also limits this paper’s discussion to some extent. While this project is informed by an intersectional perspective, the constraints of a large-scale quantitative study—particularly one focused on publicly available metadata and using platform categories and tags—necessitate a somewhat limited scope in what identities are discussed. This paper focuses primarily on “mainstream” heterosexual pornography, which tends to rely on rigid and reductive categorizations of gender and race and leaves little room for many non-conforming identities. Although gender identity, racialization, and sexuality are complex and fluid, pornographic platforms typically collapse this nuance into fixed categories. These categories are not self-identifications of performers but are imposed by the platforms, shaping how content is

recommended and consumed. As such, when I adopt terms used by porn sites in this paper it is not to endorse them or depict them as accurate or sufficient descriptions of reality, but rather to be able to investigate the functions and consequences of these labels within the algorithmic frameworks of pornographic recommendation algorithms. One of the main areas of further improvement I see right now for the project is further theoretical investigation of the tags and categories I rely on and what identities and practices they leave out.

Yet another important strain of thought is investigating the way Big Data is reorienting porn. Just like in other fields, the mythology of big data is invoked to justify how an algorithmized system works and even to make claims about what Pornhub's data, in particular its yearly Insights reveals about human sexuality – for example when mainstream news sites report Pornhub Insights as representative of broader sexual trends or when popular science books such as “Everybody Lies” make similar claims (Turner 2019). boyd and Crawford's emphasize how Big Data's capacity to aggregate and analyze vast datasets is often mythologized as offering objective and superior insights, coined originally for more “conventional” firms seems highly applicable here. In the pornographic context too claims to objectivity in Big Data often obscure inherent biases in data collection and algorithmic processes, which are shaped by economic incentives rather than neutral analysis. The implementation of analytics, specifically analytics that is often self-described as “Big Data”, in informing decisions about porn cannot be understated. As Turner (2019) points out, before the internet, porn producers relied on sales trends to guide content decisions, commissioning more of what seemed popular based on limited market feedback. In contrast, contemporary porn conglomerates, much like the data-driven platforms Boyd and Crawford describe, harness vast streams of granular user data. Every interaction—what users watch, pause, rewind, or comment on—leaves behind a traceable digital footprint. This allows companies to experiment with and

refine content at an unprecedented scale, testing everything from scene composition to performer positioning to even the furniture. The low production costs of adult content, relative to other video industries, facilitate constant iteration and optimization, resulting in a situation that is well aligned with Boyd and Crawford's argument that Big Data not only reflects but actively shapes user behavior and market trends. This is of course not to say that it is completely parallel to how data is being used on other platforms— for a striking example: it is speculated (though not confirmed) that in the adult video context engineers pay much more and different attention to the last video watched than on other platforms. (Turner 2019)

Pornhub.Tracking.Exposed (poTREX) is an empirical project that aims to investigate such practices employed by Pornhub— though it has been unfortunately discontinued, its insights are still crucial for substantiating the above theories with empirical data. Through data obtained by the help of volunteers using a browser extension, the project sheds light on the extensive data collection practices of Pornhub, illustrating how over-tracking is deeply embedded in its operational model, drawing attention to the platform's ability to monitor user behavior—down to minute details like search history, viewing duration, and geographic location. poTREX highlights how this over-tracking not only fuels hyper-personalized recommendations but also reinforces systemic biases in content suggestion, particularly in how it perpetuates stereotypes around race, gender, and sexuality, with some content categories being far over-recommended despite similar amounts of videos present. I

It is also important to consider the way the market-based logic of many entities in the online pornography space can directly influence the non-pornographic Internet. Safia Noble's description of search engines returning hypersexualized results for terms like "black girls" is a well-known early example of this. Noble points out that at some point searches for Black women and girls

consistently yielded pornographic content, reflecting a deeply entrenched association between Black female identity and sexual objectification – enabled and brought to the Internet by the commercial interests of porn sites using keyword optimization. However, looking at contemporary pornographic content, the position of black female sexuality is complicated further, as Shor and Gorliz (2019) report: "Contrary to our theoretical expectations and to the findings of previous research, we found that videos featuring Black women were less likely to depict aggression than those featuring White women, while videos featuring Asian and Latina women were more likely to depict aggression". Their discussion of aggression and race is central to this thesis's argument and will be expanded upon significantly in the results section.

All of the above points us to the argument that Adler puts forward in *Arousal by Algorithm*: that the current state of affairs in porn is similar to the way the entities we often refer to as "Big Tech" warp social realities by using algorithmic search engines and suggestions, rigid categorization of content, and artificial intelligence driven optimization of their profits while not having a clear understanding of what a given algorithm exactly does anymore. Adler's concept of "porn authoring itself" highlights the feedback loop between user engagement and content production in the digital pornography landscape. Similar to social media platforms, Pornhub's algorithmic engines categorize, optimize, and promote content based on user behavior, creating a system where pornographic categories emerge and evolve in response to granular data. This process, described by Adler, exemplifies "keywordization"—the relentless tagging and organizing of sexual content into hyper-specific categories designed to drive traffic and engagement, with logic dictated by the imperative to make profit.

The exact effects of such algorithmic recommendation are hard to tell, but there are signs that they can result in shifting viewing habits to strange places. For example, it has been theorized that the

data-driven algorithms exacerbated if not drove the ascendance of faux family (stepsister, stepmother etc) pornography, not particularly present at the start of the 2010s yet incredibly prevalent in produced and recommended content today. (Stoya, 2018). Choking is another trope that has been receiving media and legislative attention due to its frequent depiction in pornography and real-world habits among young people in fact shifting significantly in the past decade. Some of these trends are currently receiving public/legislative scrutiny due to the possibly harmful effect their popularity has on a wider societal scale, so it is in my opinion of theoretical interest how they interact with race. This paper unfortunately doesn't have the space to fully explore Adler's argument on the connection between porn platforms and big tech and the platformization of the social sphere in general, but hopefully just the above is a convincing argument for why the website design and algorithmic framework of these sites is of interest.

A final important framework for this paper is Nick Srnicek's conceptualization of platform capitalism. Srnicek argues that contemporary capitalism is increasingly dominated by platform-based business models, wherein data extraction, user intermediation, and monopolistic control over digital ecosystems become central to profit generation. This framework seems perfectly matched for understanding the operations of Aylo, Pornhub's parent company, which mirrors the tendencies of major tech firms of leveraging data as a core resource and reinforcing monopolistic dominance through control of the pornography ecosystem. Srnicek's -perhaps obvious yet in my opinion quite important - insight that tech firms are fundamentally capitalist entities - driven by profit, competitive pressures, and pursuit of data - also points us to how pornography platforms, despite their distinct content, operate within comparable economic logics as Google or Amazon. This conception of Aylo and the other porn companies as fundamentally capitalist entities is what

motivates the focus of my following chapter, advertisements recommended to browser sessions with no past activities.

Advertisements recommended to browsers without prior watch history

Especially for a company that claims to rely heavily on data to optimize its operations and has been running for a while it's crucial to have a sense of what the economic incentives at play are. While the production and recommendation of videos is obviously important, many decisions are based on maximizing revenue which comes almost exclusively through advertisement – either to subscription porn services or to other products, but through advertisement nonetheless, with all other sources of revenue secondary. As such, let us now focus on online pornographic advertisements specifically. The state of affairs in the porn industry currently, i.e. being mainly sustained through advertisement, is a relatively new one, with the rise of the tube sites facilitating both a move towards data analytics and towards advertisement. However, it has been around for long enough for us to be able to make some claims. The majority of internet porn users consume content passively, watching videos without making any financial transactions. Only a small subset of users actually generate revenue for porn. This typically happens when they click on advertisements featured on tube sites or proceed to dedicated content platforms (advertised on the tube sites) —many of which are owned by Aylo—where they make a payment to access premium content. (Turner, 2019, Adler, 2024)

The distinction between casual viewers and paying customers plays a critical role in shaping the pornography industry. Companies rely heavily on data from paying users to determine what types of content drive revenue. Since the financial success of professionally produced porn hinges on the preferences of this small group, their tastes and behaviors likely significantly influence the kind of content that is created, promoted, and normalized across the industry. This process

perpetuates certain categories and trends, reinforcing demand-driven production cycles that might seem niche or unexpected to broader audiences. (Turner, 2019, Adler, 2024) Showing the peculiarity of the space theorizing exists today, the above is based on a well-researched magazine article titled *My Stepdad's Huge Dataset*.

But indeed, what are these advertisements, clicks and impressions of which ultimately drive the profits for porn conglomerates? Mainstream brand advertising which sustains many other ad-driven platforms such as Youtube or Facebook is rare on adult sites. So, what are the specific ads, posters of whose provide the revenue source towards which at the end of the day the capitalist porn platform's data scientists and engineers are optimizing towards? How do they relate to race and aggression? To answer these questions, we first look at the advertiser's side of this issue by investigating how can an ad for a product or service indeed end up on a porn site.

Posting ads to porn sites

The service through which potential advertisers can post ads to Pornhub is called TrafficJunky. Its website is well-designed and professional, which seems fairly consistent with other major ad agencies providing access to the adult video space— TrafficFactory, which provides advertisements amongst other sites to Xvideos, is for example a bit less user-friendly but has a similar overall tone and structure. The rules laid out for advertisers on TrafficJunky are detailed and functional, clearly based on experience – no Facebook lookalikes, no explicit scams, proper notice next to any medical claims made, no auto-downloads, no attempts at subliminal messaging and so on. To get a sense of the website, here is the way they detail the different positions one can buy an ad in:

Display Banner

Top Right Square

Video A, Video B

Inplayer

Footer

Header

Underplayer

Interstitial

In-Stream Video

In-Stream Video

Banner CTA Image

Popunder

Most of these terms are fairly straightforward, yet are clearly meant for a specialist audience well-versed in digital advertising - for example popunder in plain language describes ads that open as new pages that don't become the active page.

TrafficJunky also specifies the types of ads that each of the above positions allows ,the union of which can presumably covers all the possible ad categories. These are:

Adult Games

Adult VOD Cams

Cannabis/CBD

Crypto

Dating

e-Commerce

Forex

Male Enhancement/Nutra

Financial Services

Apps/Games

Casino/Sports Betting

VPN

The categories are mainly self-explanatory. VPN stands for Virtual Private Network, a proxy server which amongst other uses can help users bypass Internet censorship such as geo-blocking and help users who want to protect their communications against data profiling – a somewhat ironic service to be offered on the porn sites that do deal with data profiling as we saw for example in the findings of Pornhub.tracking.exposed. Forex stands for foreign exchange markets and the services advertised under this category usually relate to trading currencies. We can see that there is a fair amount of variance in the type of content that can be offered – and that ads to other porn sites are not present, presumably because Aylo reserves that for in-network premium sites.

In line with our expectations about “big data” and the analytics-driven platform, TrafficJunky’s website also describes the targeting services it offers. While the exact details are not available to

the general public, the prospectus advertises the possibility to target your ads based on location, demographics, keyword, network and device type, IP and ISP, browser, browser language and operating system, a dizzying array of possibility – similar options can be found for TrafficFactory. In general, we can see here a subsidiary of Aylo behaving much like any other firm, providing guidance for potential customers, pricing information and so on. It also describes the bidding infrastructure and market for ads, a feature of online advertising that is unfortunately outside the scope of this paper yet probably also influences the exact ads being displayed.

However, when detailing payment methods, there is an option missing present on the rest of the Internet: the possibility to pay with credit cards. This is because after concerns about exploitative and abusive content appearing on Pornhub and the site not doing enough to take it down, Visa and Mastercard severely limited their business with Aylo– not only for payments on the porn sites themselves, but even for companies that try to post their ads on porn sites. (Adler, 2024)

This move was a heavy blow to Pornhub, which in the 2010s tried to become part of the more “mainstream economy” – it ran ads that normalized Pornhub with clean, corporate aesthetic, and for some time there seemed to be successes: for example the mainstream movie Don Juan had ads on the site and so did Dollar Shave Club, owned by Unilever. Pornhub even ran a program that gave out impressions on Pornhub to small businesses such as jewelers or barbershops during the Covid-19 pandemic. However, these efforts seem to have waned, with no recent reports of spending by household name companies, leading us to the next part of the article: findings about who still advertises on Pornhub and other tube sites.

For this section of the paper, I’ve conducted manual content analysis, in separate browsing sessions (n=129) with no prior interactions with Pornhub, varying attributes such as IP, browser language, browser. of advertisements offered (n=43) on mainstream porn sites today, with the sampling

taking place August of 2025. The focus was on content analysis rather than a full quantitative investigation involving scraping, as ads lack the kind of structured tagging that is available for videos. I focused mainly on ads on Pornhub itself, with a few alternatives investigated in less detail: Xvideos and XNXX, two of the main competitor sites and one of the other sites owned by Aylo: RedTube, were also investigated.

The focus was on the United States – within the US recommendation variety in ads was low, as tested by using a VPN with locations set to different US cities, but it is important to note that ads for example in France were markedly different so the findings of this paper should only be taken to be applicable to the United States setting in 2025. Rather than just recording what ads were present, I followed the links provided by the ads and to try to ascertain how and where money could be made. I employed saturation sampling: creating new browser sessions until I felt I have seen all of the general archetypes of ads in previous sessions.

Findings

The types of ads encountered on Pornhub right now (with the caveat of the research design as described above) can be described as almost exclusively sex related. Based on content analysis the following categories emerge:

Advertisement for other pornography sites- this category is dominated by ads for Brazzers (also owned by Aylo) and other in-network premium sites. The revenue stream is clear: it's from a subscription model for premium content for which free samples are already present at Pornhub. The aesthetic is usually quite clean and “consumerist” – clear lighting, good cameras, houses that seem to capture some idea of “ordinariness”. Position wise, they are usually banner ads in “prime”

locations, though can sometimes also appear in-video. Occasionally features women of color (2 instances)

Advertisements for camming sites – these ads will depict a camming experience, more often in-video but sometimes in some of the less well-positioned banner slots too. Clicking on them takes you directly to a specific live camera room (not the one in the ad), where payment in the form of tips can influence the behavior of a model. The models depicted are almost always white.

Advertisements for sex toys – sexual devices targeted towards both male and female users were advertised, yet as we will see in the recommendation section this seems to be largely disconnected from the information available to the site at any given point in time: female sex toys don't seem to disproportionately show up in browsing sessions where search and viewing activity is searching for porn targeted at women. Some of these devices are surprisingly expensive, often in the 200-300 dollar range. Notably, some of these sites redirect ultimately to Amazon or other vendors where conventional credit cards can be used.

Advertisements for sexual wellness products - these seem to have the most possibility of being scams or at the very least misleading – promising penis enlargement, increased sperm production, or a “fat compressor that helps reveal inches of your dick buried in fat”. Revenue is made off of those who purchase such products.

Advertisement for dating or hookup site – always targeted towards men, and in particular often towards older men, these ads are usually in the banner positions and often use the terms “lonely” and “immediately”. Promising companionship, payment is required to access to messaging

features of these sites – whether such paid messages reach actual people is unclear. The profiles depicted that I have encountered always feature white women.

Sexual advice ads - once again targeted at men, they present some sexual technique, often with a narrative that learning a given skill will result in women wanting to be with you despite your other potential shortcomings. Interestingly, once on a page outside the tube site, the videos are often unskippable and quite long (upwards of 15 minutes) with a significant amount of time spent on detailing the importance and uniqueness of the advice in question (for example through testimonies of medical professionals or adult film actors), with payment required to actually view the advice.

Pornographic AI ads – they don't yet seem to have their own specific category – probably would fall under adult games - but these ads are very much present, especially in the less privileged banner ads. They usually advertise a customizable female character that will respond to a user's requests. Payment is required after an initial free trial.

An overall impression is that many of the ads use terms that directly address the user, which might be specifically efficient in altering the behavior of someone who is browsing the website drunk or under the influence of substances.

Notably, many categories from TrafficJunky are missing from Pornhub's ads: gambling, finance and in fact anything non adult-themed. Interestingly, some smaller sites- even within the Aylo network- have more of the finance and gambling related ads, but they are still overwhelmingly full of sex-related advertisement. I never encountered anything that I would deem outside of the

categories that are described on TrafficJunky's sites – though they technically might be present as the “small businesses” anecdote shows.

Apart from just establishing what kind ads were present on the sites, a secondary goal of the project was to see the way user behavior influenced the type of ads being shown. For this, n=23 browsing sessions were kept open for longer periods of time and manual search and video clicks tried to emulate specific interests that ad targeters might pick up on. Pornhub's algorithm picks up easy on gay and trans preferences – ads immediately change, with the characters depicted as gay and trans, sometimes with other characteristics of the ads remaining constant. However, interestingly (and concerningly) the site almost never displays ads targeted towards women (regardless of their sexual identity), even in browsing sessions where search terms and watching history point towards the user being female – further research could actually create an account to see if long-term behavior would lead to different ads. Targeting based on other characteristics, such as keywords for race or porn genre, are harder to tell and would need an even more complex investigation, but based on preliminary trials, searching for a specific ethnic group or sex act doesn't significantly impact what ads are recommended, with the female performer appearing in ads appearing white for all but one each of the ads reviewed that featured live performers.

My findings point to Pornhub making money off of overwhelmingly if not exclusively sexuality related products and services, whether through third-party ads or ads for its own premium sites. These ads are overwhelmingly targeted towards a male audience. While these findings might seem trivial, they have profound consequences. It is important to note that not only mainstream products are missing –as noted previously it has been a long time since any well-known brand has advertised on Pornhub. It's that only explicitly sex-related products seem to be generating revenue for a

significant portion of the US market. Important to note that apart from reputational damage, this is also likely driven by the difficulty of payment services for corporate clients of porn companies and the relatively little information available about users (apart from the fact that they're watching porn) which makes the kind of targeting ads have elsewhere on the Internet not possible. There's a fairly diverse ecosystem of sex-related ads – it doesn't for example seem that only specific demographics or preferences are functioning as sources of revenue streams, though ads targeted at women are very rare. It seems reasonable to assume that whatever analytics Aylo and other porn conglomerates are conducting will be trying increase their revenues. If my findings about ads are representative of the US market as a whole, this would mean there is a stark difference in business incentives from the way other platforms like Spotify or Netflix who try to retain consumers, or even from the other ad-based platforms such as Facebook or Youtube, as porn conglomerates are only gaining revenue from those who are willing to spend money on sex-related products, a presumably tiny minority. This likely leads to the pornography shown to a wider public – which influences their sexual views and behaviors – to be based on the purchasing behavior of a slim minority. Thus the particular ways these ad categories influence recommendation and production are incredibly important to look at, especially the way revenue mainly coming from one gender influencing the viewing experience of all genders. Let us now turn to what videos are actually recommended.

Characteristics of videos recommended to blank browser contexts on Pornhub in 2025

This chapter details how certain characteristics, most importantly the level and form of aggression and violence in heterosexual videos involving female performers of different races differ. The reason for focusing on the race and gender of performers separated rather than race of performers

in general is two-fold. First is the theoretical importance of intersectionality as detailed in the previous section. The second consideration is methodological: in heterosexual porn most categories and tags on a given video that describe a performer's characteristics refer to the female performer or performers. This focus makes automated investigation using tags and categories only reasonable for female performers, while for men more elaborate methods – scraping information about the models involved and/or matching terms to the title of the video – would be required.

To keep the paper focused, only the most relevant findings are reported here, with further plots and results available at <https://github.com/Irgul/Thesis>.

Data gathering

The quantitative findings of this thesis are based on an original dataset that was scraped from Pornhub, from February 22, 2025, to October 11, 2025, available at <https://github.com/Irgul/Thesis>. Important to note that everything recorded in the dataset is publicly available information with no data about specific individuals and no videos or images stored, only video metadata such as views, duration, tags and categories.

To establish what kind of content is recommended to users, automated browser sessions were run periodically using a browser automation library called Playwright. Opening a new browser context for each session, the script cycles through forty distinct device profiles. These profiles are a combination from five device types, four United States time zones and two color schemes. The device types are Safari on iPhone 12, Desktop Chrome, Chrome on Pixel 5 Android, Desktop Firefox, and Safari on iPad Pro, which are mimicked by setting the user agent string, viewport dimensions and touch support flag. In addition, the computer running the script is connected to

ExpressVPN, a Virtual Private Network service whose IP rotation feature changes IP address for each new browser session, though importantly all the IP addresses it cycles through are United States ones. The purpose of this rotation is two-fold. It is meant to cover a variety of possible settings in case these influence what kind of videos are recommended – while this is unlikely for color scheme, it is a legitimate possibility for browser type, though for the metrics we’re interested in this did not seem to be the case. The secondary aim is to prevent browser profiling/fingerprinting, which there have been accusations of porn websites engaging in. While the above are far from covering all the possible variables Pornhub can collect about a visitor, hopefully varying just the above is enough to throw automated systems off.

After opening a new browser instance with the specifications detailed above, the “homepage” scraping process navigates to Pornhub.com, clears modals associated age requirements and cookies, extracts all visible video links present on the main page. The system then processed each identified video link by opening it in a new browser tab and extracting relevant metadata, with small uniform random waiting times (1-8 seconds) between each modal clearing click to make the browser’s behavior more organic. For the “search” scraping process, after the initial homepage loads, the script clicks into the search bar and searches for one of four terms, recording all the videos present after. Finally, the “journey” process does everything the search process does, clicks on one or more of the videos, watches them for some time, records the videos recommended here, then returns to the homepage and records what videos are present.

To motivate why this method of collecting data was used, it is crucial to understand how users use tube sites. For the majority of users, they don’t have an account (cite) and are using Pornhub in incognito mode (cite) which makes tracking based on cookies or similar methods harder – as

mentioned above, there have been allegations (denied by Pornhub) of Pornhub tracking users using so-called “fingerprinting”, but as far as modern content sites go, Pornhub has been theorized to have one of the most current session dependent recommendation systems. As such, unlike most media recommendation systems, actual visitors appearing as “blank” users without a search history and being recommended content largely similar to what the scripted browser instances are recommended aren’t unrealistic here. This makes of scripted visitors far more analytically sound than for example it would be for making claims about the content that someone using YouTube is shown, which makes using automated browser instances uniquely well-suited for our research.

From the videos collected this way, 3 smaller datasets corresponding to each mode were created, with entries deduplicated by URL, with many later analyses using these datasets. This decision was made after careful consideration and significantly impacts the interpretation of many of our findings – and also makes them be more analogous to previous studies. Though the scraping script can pick up on the same video several times – and in fact this does happen incredibly often – it doesn’t make sense to weigh videos by how many times they occur along scraping. This is because to establish likelihood, even over a period of time, we would need to know more about distributions of when people watch, on what browsers, etc, without which how many times certain videos appear could be completely unrepresentative of how many times they are being recommended. Thus, claims can be made about what the set of videos recommended to blank sessions consists rather than the likelihood of certain videos appearing. It is also important to note here the way that our data collection differs from sampling that is based on most watched videos in general or for a specific category that many other papers engage in. The videos in our sample are what are being recommended specifically to people without accounts or prior interaction during the months the

script was running, providing a snapshot of current recommendations, rather than an aggregate of viewing patterns that go back almost two decades.

Operationalization choices

Following Shor & Golriz (2019) I chose to focus on the four most prominent racial/ethnic categories present on porn sites: White, Black, Latino and Asian. For search results, the outcome we're interested in is what comes up when searching for ethnic categories and thus no further ethnic coding is required. An important choice for search is to use the terminology: "Ebony" and "Latina" are gendered for female performers specifically, are amongst the most used and searched categories (Pornhub Insights 2024) and as they are the subject of our research it made sense to use these terms. For homepage videos however, we need a way of categorizing of what race the female performer is depicted as being part of.

For this, I rely on the tags and categories scraped, specifically this mapping:

```
{  
  
  'white': ['white', 'caucasian', 'european', 'pawg'],  
  
  'ebony': ['ebony', 'black', 'african'],  
  
  'latina': ['latina', 'latin', 'hispanic', 'mexican', 'brazilian', 'colombian'],  
  
  'asian': ['asian', 'japanese', 'chinese', 'korean', 'thai', 'vietnamese', 'filipino', 'indian']  
}
```

Tags or categories immediately followed by “man” and “guy” are disregarded. If a video has one or more tag or category from a group and only that group present, it will be classified as belonging to that group. If a video has a term from more than one group, it will be classified as “several”.

Based on a manual review of $n=28$ videos categorized this way, I found that for these 28 videos the tags and categories that are present do accurately categorize the race of the female performer or performers present. This however does not mean that every video is categorized accurately. “White” not being present as a category and being assumed as the default (as for example Rama et al. discuss) contributes to many videos involving female white performers not being tagged at all with terms related to race/ethnicity, but this does happen for performers of other races too.

For aggression, the situation is somewhat more complicated. Though many canonical definitions stress the importance of consent, Bridges et al. put forward a reasoning for employing a broader and more inclusive definition, as the genre of pornography often requires performers to express enjoyment following just about any act or situation. Due to these powerful expectations, they argued, almost no act would be coded as aggressive under a definition that requires a target to clearly show displeasure or lack of consent, rendering sexual aggression invisible. Bridges et al. (2010) therefore defined aggression as “any purposeful action causing physical or psychological harm to oneself or another person, whereby psychological harm is understood as assaulting another verbally or nonverbally”. For this paper, we will consider aggression as purposeful action causing physical harm.

This is because symbolic or psychological harm – whether it be acts that are understood as humiliating in many contexts (ejaculating on face) or derogatory terms - is context-dependent in a way that makes classification nigh-impossible. I also considered levels of aggression as one possible instrument, but ultimately decided in favor of a binary classification, as how many

different tags describing potentially aggressive behavior were present seemed a poor sign of the length and seriousness of aggression present, and no seriously problematic (clear indication of lack of consent for example) terms are present anymore as tags or categories – whether this is representative of the actual acts being committed is unclear.

Due to the number of unique tags and categories, I utilized a script that made API calls to GPT-4 for establishing the possible aggression terms, working through the around 19 000 terms and categories present in decreasing order by frequency, with a prompt detailing the context in which the judgement has to be made, then manually reviewing those the LLM deemed potentially aggressive after every 20 terms flagged, with subsequent queries to the model also having the manual decisions so far made as context.

Looking at the resulting terms, they could be grouped into 3 broad categories: harm through physical pain, harm through restriction, and generalized harm, with the most frequent tags and categories associated with each listed below:

Physical/Pain: deepthroat, rough sex, rough-sex, rough sex, rough, rough-sex, hard rough sex, deep throat, deep throating, choking, deep-throating, hard-rough-sex, hard fast fuck, face fuck, gagging, deep-throat, screaming, hard-fuck, big dick small pussy, face-fuck, spanking, slapping, deepthroat, licking-pussy, face slapping, double vaginal

Restrictive: bondage, spread eagle, bound

Overall: gangbang, extreme

Gangbang was a term I spent some time debating whether to include or not, but ultimately how often it comes up and the fact that it describes a situation where aggression is implied – as opposed to “group sex” or “orgy”, it made sense to put it on the list.

Apart from aggression, I decided to also investigate affection. This one is even harder to operationalize, but using a similar technique, a binary affection variable was constructed, using these terms:

Affection: pussy licking, pussy-licking, romantic, pussy licking, romantic, pussy-licking, kissing, eating-pussy, sensual, passionate, spooning, passion, romantic sex, passionate real sex

These variables are not new in theoretical approach and are indeed more simplistic than some of what the previous literature uses. But they do enable comparison between the groups and are roughly consistent with previously reported levels of overall aggression. A manual review of 10 videos also found that the videos where they are present would be categorized as aggressive under the Bridgers definition – though some videos that are not classified as such from the sample with this coding would also be considered aggressive.

Results

The two most straightforward samples are the results of a deduplication of the videos scraped by the search and homepage scripts. The differences in the sample sizes don't necessarily mean that a kind of video is recommended less, just that there are less unique instances that the scraping has encountered. Number of views at time of scraping of the videos are roughly consistent between non-white groups, but significantly higher for White videos in both datasets.

Table 1: Homepage videos with inferred ethnic group and view statistics

Ethnic Group	N	Mean Views	Median Views
Asian	560	961,093	249,017
White	368	1,963,036	394,000
Ebony	434	1,365,317	264,000
Latina	954	1,236,958	298,785
Several	337	927,798	251,000
None	4,011	1,387,572	295,000

Table 2: Search videos with inferred ethnic group and view statistics

Search Term	N	Mean Views	Median Views
Asian	645	1,386,770	333,000
White	246	4,061,592	997,000
Ebony	617	1,322,735	236,000
Latina	723	1,725,121	385,000

Table 3: Aggression metrics in homepage videos

Ethnic Group	N	Any Aggression (Prop)	Physical Pain (Prop)	Restrictive (Prop)
Asian	1,244	0.305	0.268	0.053
White	614	0.321	0.303	0.018
Ebony	1,066	0.340	0.301	0.040
Latina	1,772	0.341	0.320	0.027
Several	818	0.290	0.269	0.021
None	6,815	0.310	0.286	0.031
Chi-squared p-value		0.0307	0.0150	0.0000

Aggression in homepage videos indicates a higher proportion of videos having at least one aggressive tag for videos with Latina and Black performers, and somewhat lower rates for those with Asian performers, when compared to either White or None. Restrictive aggression is especially notable in having higher rates for performers of color, especially Asian performers. I report chi-squared p values for each of my outcome categories, as the chi-test is an appropriate method when assessing whether the proportion of a binary outcome differs significantly across multiple independent categorical groups.

Table 4: Aggression metrics in search videos

Search Term	N	Any Aggression (Prop)	Physical Pain (Prop)	Restrictive (Prop)
Asian	775	0.235	0.205	0.035
White	319	0.176	0.166	0.022
Ebony	756	0.229	0.200	0.033
Latina	838	0.274	0.264	0.024
Chi-squared p-value		0.0039	0.0005	0.4434

For search videos, Latina and Ebony videos are yet again more often classified as aggressive, but here Asian videos also have higher percentages of being tagged as aggressive than White ones. Restrictive videos however are fairly evenly distributed here (with the p-value of the chi-test far from significant).

Apart from aggression, I was also interested in how affection and certain tropes or acts are distributed. Videos were tagged as “Step” if the term was found in their tags or categories and as “Choke” if choking or choke were present.

Table 5: Affection, step and choke metrics in homepage videos

Ethnic Group	N	Affection (Prop)	Step (Prop)	Choke (Prop)
Asian	1,244	0.162	0.100	0.026
White	614	0.256	0.300	0.007
Ebony	1,066	0.174	0.117	0.010
Latina	1,772	0.158	0.234	0.013
Several	818	0.130	0.219	0.004
None	6,815	0.164	0.208	0.014
Chi-squared p-value		0.0000	0.0000	0.0004

Affection and Step content are both much more common for videos classified as White. Rates of the choking category are low – though they are clearly most common for Asian videos.

Table 6: Affection, step and choke metrics in search videos

Search Term	N	Affection (Prop)	Step (Prop)	Choke (Prop)
Asian	775	0.133	0.098	0.015
White	319	0.135	0.169	0.000
Ebony	756	0.110	0.101	0.008
Latina	838	0.123	0.165	0.010
Chi-squared p-value		0.5078	0.0000	0.1074

For search results, Affection rates are a lot more similar between groups, while Step content is still more common for White videos but not the differences are smaller. Choke is once again rare, but still most common for Asian videos.

The journey results were counterintuitive. Ethnic category proportions didn't particularly change or changed in directions opposite to what I expected on the homepage for a variety of journeys that were meant to specifically model someone being interested in a given ethnicity – ie. journeys that search for and click on videos with a certain characteristics several times, then return to the homepage. This could be an indication of further data collection being necessary to isolate the effect, but even with these sample size the project currently has it seems to be an indication of a less responsive system than some of the literature reviewed has theorized. This points to Pornhub's recommendation algorithm not being as sophisticated as some fear or as they themselves have advertised in press releases.

Conclusion

The differing proportions of aggression in search results – and for Latina and Black women, in homepage results – are a strong indication that videos recommended to user sessions without prior interaction are significantly different for videos featuring female performers of ethnic categorizations other than White. Importantly, this claim shouldn't be taken to mean that this discrepancy is the result of Pornhub's algorithm or of the kind of videos that are present on Pornhub, just as a descriptive statement about an outcome observed. Depending on how much fingerprinting we assume on Pornhub's behalf, it does indicate that such differences could be representative of content that a huge portion of visitors to the website are recommended. Due to Pornhub's ubiquity in popular conceptions of porn, Pornhub's homepage and simple search function are also likely landing spot for teenagers interacting with porn for the first time, meaning the above differences will be present in the sexual content first encountered by some users– and possibly conditioning them to expect certain behavior in further pornography consumption.

For advertisements, the total lack of non-sexual advertisement while not surprising, is a significant finding that warrants further investigation. The lack of responsiveness in ads to racialized searches/viewing patterns should be validated with a larger sample size/potentially through an automated investigation similar to the one conducted for videos, but points to Pornhub's "big data" based business model being overstated.

Our findings on how rates of aggression differ between videos with female performers of different races are for the most part are in line with what Shor & Golriz (2019) report, though unlike them we find videos involving Black female performers to be more likely to contain aggression when

compared to videos with White categorization. The causality of this shift is hard to establish, but it is more in line with the expectations that Shor & Golriz report based on their review of previous literature and canonical conceptualizations of depictions of Black female sexuality.

A somewhat surprising findings is that aggression rates are smaller when searching for any of the ethnic categories investigated than on homepage videos of performers of the same ethnicity. One possible explanation for this is that those who are interested in aggressive videos involving women of a given ethnicity search for “rough [ethnicity]”, not just “[ethnicity]”, so when searching for “just” and ethnic term, updating based on this the algorithm recommends less aggressive content. Another possible explanation is that based on our findings on revenue streams and ads, it is plausible to theorize that the reason for higher aggression rates on the homepage might be in order to target potentially paying customers whose preferences are for more extreme stuff. While this is an elegant explanation and could also account for Internet pornography’s alleged shift towards more extreme content in general, we don’t have enough data on purchasing behavior to plausibly argue this – and this paper’s findings on lack of sophistication in Pornhub’s recommendation systems are also evidence against this line of argument.

For several outcome categories, the proportion of Asian videos belonging to them were clear outliers. This is also mentioned by Shor & Golriz (2019) in their study who also report that Asian women are likely to suffer from greater aggression despite their relative passivity, or perhaps even due to it. They also discuss previous accounts of feminist scholarship that linked the stereotypes of Asian women as passive doll-like “fucking machines”. (Shor & Golriz 2019) Based on our findings, Asian women are the ones most likely to be restrained and choked in videos commonly recommended.

For the non-aggression related tags investigated, affection rates being higher for videos involving White women is not particularly surprising, but does reestablish previously held assumptions with our present-focused methodology. Another category that had high proportions in videos involving White women is “step” content. Step-family dynamic seems a lot less likely to be present in videos with non-white women – one possible explanation for this is that videos need a limited number of “hooks”, with either the ethnicity of the performer, or step-family “script” providing one. Finally, despite significant legislative scrutiny, choking is near non-existent as a tag or category for the videos we investigated – this could however also point to choking being so normalized as a sexual act that it is not tagged anymore.

When it comes to insights to establish whether Pornhub’s algorithm is content- or collaborative filtering based, this paper was unfortunately able to establish little – the more complicated user journeys resulted in results that are hard to interpret. As such, it is hard to make any claims about whether the tendencies we see above are a reflection of previous user’s preferences manifesting, caused by what kind of videos are being produced, or if they are the result of some calculation based on tagging and are affecting preferences – the truth is likely a combination of these factors. Trying to overwhelm the filtering by having browsers watch two relatively obscure videos one after another an incredible number of times could be a way to test this, though guardrails for this sort of intervention are likely in place.

Overall, the project’s results points to meaningful differences between videos of female performers of different racial/ethnic categories. Further automated research into advertisements placed, and more detailed investigations of longer browser sessions would be quite valuable, as while the theoretical literature on the importance of recommendation algorithms and porn platforms is rich, there are very few large-sample empirical studies.

Bibliography

- Amia Srinivasan. *The Right to Sex: Feminism in the Twenty-First Century*. New York: Farrar, Straus and Giroux, 2021. 304 pp.
- Davis, A. Y. 1. (1983). *Women, race & class*. First Vintage books edition. Vintage Books.
- Williams, L.R. (1989). Hard Core: Power, Pleasure, and the "Frenzy of the Visible".
- Adler, Amy M. *Arousal by Algorithm*. NYU School of Law, Public Law Research Paper No. 24-41, NYU Law and Economics Research Paper No. 24-38, 2024. Available at SSRN: <https://ssrn.com/abstract=4937597> or <http://dx.doi.org/10.2139/ssrn.4937597>.
- Snitow, A. (2015). *The Feminism of Uncertainty: A Gender Diary*. Duke University Press. <https://doi.org/10.2307/j.ctv1198vn5>
- boyd, danah, & Crawford, K. (2012). "Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon." *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>.
- Rama, I., Bainotti, L., Gandini, A., Giorgi, G., Semenzin, S., Agosti, C., & Romano, S. (2022). "The Platformization of Gender and Sexual Identities: An Algorithmic Analysis of Pornhub." *Porn Studies*, 10(2), 154–173. <https://doi.org/10.1080/23268743.2022.2066566>.
- Shor, Eran & Golriz, G. (2019). "Gender, Race, and Aggression in Mainstream Pornography." *Archives of Sexual Behavior*, 48(3), 739-751. <https://doi.org/10.1007/s10508-018-1304-6>.
- Srnicek, Nick. *Platform Capitalism*. Polity Press, 2017.

Tracking Exposed. (n.d.). "Pornhub Tracking Exposed: Project Analysis." Retrieved December 6, 2024, from <https://pornhub.tracking.exposed/project-analysis/>.

Wright, P.J., Tokunaga, R., & Herbenick, D. (2023). "But Do Porn Sites Get More Traffic than TikTok, OpenAI, and Zoom?" *The Journal of Sex Research*, 60(6), 763-767. <https://doi.org/10.1080/00224499.2023.2220690>.

Urman, Aleksandra, Mykola Makhortykh and Anikó Hannák. "WEIRD Audits? Research Trends, Linguistic and Geographical Disparities in the Algorithm Audits of Online Platforms -- A Systematic Literature Review." (2024)

Sexuality, Pornography, and Method: "Pleasure under Patriarchy Author(s): Catherine A. MacKinnon Source: *Ethics*, Vol. 99, No. 2 (Jan., 1989), pp. 314-346 Published by: The University of Chicago Press Stable URL: <http://www.jstor.org/stable/2381437>

Andrea Morichetta, Martino Trevisan, Luca Vassio, Julia Krickl, Understanding web pornography usage from traffic analysis, *Computer Networks*, Volume 189, 2021, 107909, ISSN 1389-1286, <https://doi.org/10.1016/j.comnet.2021.107909>.

Additional Sources:

Stoya. "Every Story Is an Epstein Story." *Components One*, 2018. Available at: <https://components.one/posts/every-story-is-an-epstein-story-stoya-pornhub-step-porn>.

Turner, Gustavo. "My Stepdad's Huge Data Set." *Logic Magazine*, 2019. Available at: <https://logicmag.io/play/my-stepdad's-huge-data-set/>

