

REFLEXÕES SOBRE *BIG DATA*, SEXUALIDADES, DATIFICAÇÃO E MORALIDADES NO PORNÔ DIGITAL: A "NOVINHA" COMO CATEGORIA DE CLASSIFICAÇÃO E INDEXAÇÃO

Carolina Parreiras 

Universidade de São Paulo (USP),
Departamento de Antropologia. São Paulo, SP, Brasil.
E-mail: carolparreiras@gmail.com

Introdução

Este artigo, em certa medida, é uma aventura e um experimento teórico, empírico e metodológico que busca mapear possíveis conexões entre os amplos e intrincados conceitos de big data e pornografia. Aventura porque propõe abordar o estudo da big data a partir de uma perspectiva etnográfica, menos interessada nos números em si do que nos desdobramentos dessa quantificação e em que medida ela abre espaço para discussões sobre pornografia (em seus diversos gêneros, formas de produção e possibilidades) enquanto problema social que diz muito sobre moralidades e aquilo que, de acordo com Gregori (2016), podemos chamar de "limites da sexualidade".

É preciso ressaltar também que ambos os conceitos – pornografia e big data – estão envolvidos em disputas de sentido e são categorias socialmente modeladas, sendo que, muitas vezes, uma visão apressada dá significados equivocados ou muito genéricos para estes termos. Assim, neste artigo, busco trazer esta problematização, além de situar a partir de que perspectiva compreendo cada um deles. Saliento ainda que, nos últimos anos, houve um crescimento dos esforços de pensar representações pornográficas e o mercado erótico de modo mais amplo a partir

do digital e, especificamente, da ciência de dados. Cito como principais exemplos desta produção artigos recentes publicados pelo periódico *Porn Studies* (Caminhas 2022; Saunders 2020).

Do mesmo modo, há um importante movimento na antropologia de levar em consideração temas até então restritos a outras áreas do conhecimento, como algoritmos, big data e as várias infraestruturas do digital. Como afirma Seaver (2022), ao estudar "algorithm systems" (e o mesmo pode ser dito para big data), se ficarmos limitados aos aspectos técnicos, perdemos as pessoas e, conseqüentemente, os modos como elas influenciam e moldam o funcionamento dos artefatos tecnológicos (que são, como bem nos alertava Christine Hine, ainda no ano 2000, artefatos culturais). Assim, neste artigo, estou interessada em aspectos sociotécnicos envolvidos na circulação de formas de classificação de pornografia, a partir especificamente do termo "novinha", presente em produções heterossexuais (ou lésbicas para homens¹), mostrando os embates em torno de sua utilização e sua associação com outros termos, *hashtags* e categorias. Desse modo, me interessam tanto os dados como a relação das pessoas com esses dados, ou a "humanidade" e a "mecanicidade" (Haraway & Nakamura 2003) presentes nos arranjos sociotécnicos.

Entendo este artigo como um experimento, na medida em que, nele, trago minhas primeiras incursões etnográficas com uso de ferramentas de mapeamento e cartografia de plataformas digitais. Ainda que eu tenha uma longa trajetória nos estudos para o digital (Hine 2015; Parreiras 2008, 2012, 2013, 2015, 2022; Parreiras; Lins & Freitas 2020; Parreiras & Macedo 2020) e me considere uma *heavy user* de tecnologia, até então nunca tinha levado a cabo nenhuma pesquisa que lidasse com métodos computacionais ou ferramentas de métrica e mapeamento. Era, portanto, uma neófito neste campo, o que implicou um processo de aprendizado, como cursos de programação em R e data science ou assistir a inúmeros tutoriais de como usar programas como Nvivo ou Atlas.ti. Como ficará mais claro ao longo do artigo, foi central para a feitura deste texto o uso da extensão NCapture, do Nvivo, aplicada ao Twitter.²

Ainda que em caráter introdutório, creio ser necessário situar minha experiência de pesquisa com tecnologia e pornografia/erotismo, visto que ela é um dos pontos de partida para as reflexões aqui desenvolvidas. Desde 2006, venho buscando pensar as muitas interfaces de *digital studies* com os estudos de gênero e sexualidade. Entre 2009 e 2015, realizei uma pesquisa de doutorado que teve como foco a chamada pornografia alternativa³ ou *altporn*, buscando compreender como os gêneros alternativos apareciam no contexto brasileiro e em que medida isto nos ajudava a refletir sobre o que chamei de convenções — estéticas, de atos sexuais e relativas a desejos, fantasias e prazeres — do *altporn* e o modo como os sujeitos envolvidos com a produção deste tipo de pornografia utilizavam e davam sentido

aos aparatos tecnológicos, tanto para produção quanto no contato com o público consumidor, na apropriação do espaço urbano e para o estabelecimento de relações entre si e com o mercado erótico mais amplo.

Naquele momento, começávamos a falar sobre big data e a situar mais detidamente a discussão sobre algoritmos, *hashtags/tags* e redes sociais. Como pontuam danah boyd e Kate Crawford (2012), mesmo que possamos considerar big data um termo "pobre", é inegável que vivemos algo que poderia ser classificado como "era da big data", sendo que estudiosos/as e desenvolvedores/as das mais diversas áreas têm se dedicado a analisar estas quantidades massivas de dados e rastros digitais que todos nós produzimos — conscientemente ou não — em nossos usos cotidianos das tecnologias e redes de conexão. No meu caso específico, ainda que não tenha me dedicado a compreender aspectos relativos à big data em minhas pesquisas, já ensaiava uma discussão sobre *hashtags* e formas de indexação, especialmente do modo como era feita na pornografia alternativa.

Durante os anos seguintes, após a conclusão do doutorado, me afastei dos estudos sobre pornografia e, no intuito de pensar os já citados limites ou, para tomar as colocações de Carole Vance (1986), o modo como sexualidades transitam entre "prazer e perigo", passei a me dedicar ao estudo de violências sexuais contra adolescentes em contextos de favelas da cidade do Rio de Janeiro. Buscava compreender, nesta pesquisa, o que (e se) era classificado como violência e, especificamente, violência sexual, entendido como um termo amplo para uma miríade de atos sem consentimento e em que são centrais os marcadores de gênero e sexualidade. A partir da pesquisa de campo, ficou claro que o conceito de violência, quando aparecia nas narrativas, referia-se a formas pública de violência, como a entrada da polícia, sendo que havia uma invisibilização e uma quase naturalização de violências íntimas e ligadas a gênero e sexualidade.

Foi durante esta pesquisa que começou a surgir uma série de postagens em plataformas e redes sociais — especialmente Twitter e Facebook — que traziam tabelas extraídas do PornHub Insights com as palavras mais buscadas por brasileiros quando se falava de pornografia. O PornHub Insights — cujo descritor de página é "digging deep into data"⁴ — faz parte do PornHub e se dedica à pesquisa e à análise dos dados de acesso da plataforma, apresentando, a partir dos "bilhões de hits" das/os/es usuárias/os/es, um compilado de informações guiado por datas temáticas, assuntos do momento, os mais acessados do ano ou mesmo por país.

Uma destas palavras despertou minha atenção de imediato: "novinha". Em termos muito gerais, a novinha representa a jovem mulher que é alvo de atenção e desejo sobretudo de homens mais velhos. Meu interesse nesta categoria se justifica por sua constância nos contextos de favela em que eu fazia pesquisa de campo. A novinha era personagem do cotidiano das adolescentes/minhas interlocutoras, sendo que aparecia como modo de caracterização, como chamamento, nas letras

de músicas que ouviam e nas festas que frequentavam e nos imaginários — em termos de desejos e sexualidades — destes locais. A novinha evoca o jogo de fantasias e desejos que oscila entre proibição, interditos, risco e violência e as múltiplas possibilidades de prazer e de vivência positiva e consensual da sexualidade. Tudo isto, quando aparece no registro daquilo que é associado ao pornográfico e ao erótico, nos diz muito sobre ansiedades, pânicos e moralidades, o que corrobora a visão de que não se trata simplesmente e unicamente de ser anti ou pró-pornografia, mas de pensá-la como problema social altamente relevante para compreendermos diversas relações sociais e os processos de constituição dos sujeitos.

A partir deste disparador — as postagens em redes sociais sobre a palavra novinha enquanto categoria de indexação encontrada em sites pornográficos de conteúdo gratuito como o PornHub — meu objetivo neste artigo é refletir etnograficamente sobre uma faceta das relações entre big data e pornografia, com foco específico em *tags* e categorias, apontadas pelos levantamentos do PornHub Insights como as mais pesquisadas por usuários brasileiros. Como forma de ir além dos dados da plataforma — nada neutros e voltados a interesses mercadológicos — e de perceber como se dá na prática a associação entre novinha e pornografia/erotismo, também realizei um mapeamento exploratório desta categoria no Twitter.⁵ Meu intento, neste caso, não é dar centralidade aos dados quantitativos em si, mas recolocar o sujeito no meio desse conjunto de dados, percebendo seus fluxos (e os fluxos de formas de classificação e dos significados do pornográfico), a associação de categorias e os sentidos distintos que uma mesma categoria pode ganhar quando imersa em relações sociais diversas. Este objetivo está de acordo com pesquisas de colegas que compõem este dossiê, como Perin (2024) e Walford (2024), ao falarem especificamente dos dados georreferenciados do Projeto Embrapa-Moçambique e da extração de dados ambientais na Amazônia.

Com isto em mente, este artigo busca dar conta de algumas discussões, a saber, as possibilidades etnográficas para o trabalho com big data (localizado neste experimento de construir uma etnografia da *big porn data*), *tags* e categorias de classificação (sejam as *hashtags*, *tag clouds* ou outras formas de indexação) e seus possíveis significados e o modo como a categoria novinha permite refletir sobre moralidades e sobre convenções de gênero e sexualidade que passam a ser definidoras até mesmo do que seria o Brasil consumidor de pornografia, como se houvesse uma possibilidade de defini-lo de forma estanque e a partir dos dados de uma plataforma que está sob, ainda que pornográfica, uma lógica de "capitalismo de vigilância" (Zuboff 2019). Tudo isto reitera a pornografia como problema social que nos ajuda a compreender a realidade social de forma mais ampla. Além de trazer estes dados mais gerais e de larga escala, utilizo material etnográfico proveniente de minhas navegações nestes sites, como a simulação de busca das

mesmas palavras-chave. Neste caso, busco associar a estética e as representações encontradas em produções indexadas por estas categorias, propondo, mais uma vez, ir além dos números, da larga escala, das tabelas e dos recortes trazidos pela plataforma. Em suma, estou propondo mapear moralidades e os "limites da sexualidade" por meio da categoria "novinha" como indexadora de conteúdo pornográfico e que é apresentada como uma das formas de representação mais buscadas por brasileiros no Pornhub. Acredito que isto pode nos ajudar a entender os modos como pornografia e big data se intersectam e iluminar análises sociais mais amplas em termos dos usos do digital, segundo uma das várias relações entre tecnologia e vida cotidiana.

Em termos de sua estrutura, o artigo se divide em duas partes. Na primeira delas, descrevo brevemente o contexto brasileiro durante o período em que realizei o levantamento etnográfico e a análise dos dados. Isto é importante na medida em que as moralidades, as ansiedades sociais e os pânicos morais (Rubin 1986) existentes estão intimamente relacionados ao que a pornografia representa e aos modos como é vista enquanto ameaça. Faço ainda uma apresentação mais detida dos métodos utilizados, tanto a partir do uso dos dados fornecidos pelo Pornhub quanto do mapeamento exploratório realizado no Twitter. Na segunda parte, procuro refletir sobre a rentabilidade da etnografia para análise de big data, mostrando como podemos interpretar antropologicamente estes dados em larga escala. Na terceira seção, apresento os dados disponíveis no Pornhub Insights, buscando analisá-los etnograficamente e com foco na categoria "novinha", mostrando seus vários significados que não se restringem ao erótico/pornográfico.

Insights⁶ contextuais e metodológicos: etnografia, big data, moralidades e um Brasil que descobre a pornografia datificada

No carnaval de 2019, o então presidente brasileiro de extrema direita, Jair Bolsonaro, utilizou sua conta oficial no Twitter para divulgar um vídeo, de filmagem amadora e bastante rudimentar, claramente gravado em um celular, em que duas pessoas realizavam uma performance (que poderia ser considerada erótica por alguns) em cima do teto de um ponto de táxi. Ao redor desta cena, pessoas celebravam o carnaval, com música e dança. No que se via no vídeo, dois rapazes performavam e, em um dado momento, um deles urinava no outro, que mantinha o dedo no ânus.

Ao postar esse vídeo, a repercussão foi imediata. Nos comentários, pessoas indicavam que a cena representava um *golden shower*, conhecido ato sexual presente em produções pornográficas, que consiste em que uma das pessoas envolvidas nos atos urine em outra ou outras. Logo em seguida, o presidente tuíta

novamente: "o que é *golden shower*?". Ainda que este acontecimento tenha sido analisado em profundidade em outros trabalhos (ver Coelho 2020, por exemplo), sobretudo no que se refere ao anseio classificatório da cena — pornográfica ou não, erótica ou não, contestação política ou não, simples gozo carnavalesco ou não — ele aparece aqui em outro viés, diretamente relacionado à discussão sobre pornografia, big data e também às moralidades que dão o pano de fundo contextual para o Brasil de anos recentes.

Logo após a pergunta sobre o que seria o *golden shower*, houve um verdadeiro turbilhão de reações, seja dos defensores do presidente e da suposta moral e dos bons costumes que ele representaria, seja de seus opositores. Entra em cena aí um dos vieses da big data e daquilo que creio poderíamos chamar de uma das faces da datificação da pornografia. Para consumidores/as mais assíduos de pornografia ou mesmo praticantes de práticas sexuais dissidentes, o *golden shower* não representa uma novidade, na medida em que é uma das práticas do chamado pornô bizarro (Leite Jr 2006; Dias-Benitez 2012) que está na fronteira com outros nichos e categorias do mercado pornográfico, aparecendo como palavra-chave e classificatória em sites mais convencionais e conhecidos como Xvideos, Pornhub, RedTube, dentre outros. No entanto, para a maior parte dos seguidores do ex-presidente (especialmente de seus/suas apoiadores/as mais ferrenhos/as), parecia haver um desconhecimento sobre o significado do ato tuitado.

Algo que desde sempre foi estudado e falado em relação à internet é a capacidade de chegar a sites/redes/portais/plataformas por meio das ferramentas de busca, que estão na base da organização e da indexação dos conteúdos produzidos. Estudos desenvolvidos especialmente a partir dos anos 2000 (Granka 2010; Halavais 2017; Zuboff 2019; Noble 2018) têm apontado o caráter fundamental e cada vez mais preciso das ferramentas de busca, bem como o fato de que essas ferramentas não são neutras e, além de serem parte de todo um mecanismo de vigilância sobre nossos comportamentos, desejos e práticas, atuam de forma muitas vezes discriminatória (em termos de gênero, raça, sexualidade, classe social, geração) e movidas por interesses basicamente comerciais (como exemplo, os links financiados do Google). Além disso, sabe-se também que estas ferramentas atuam por algoritmos que criam personalização (ver Pariser 2012, por exemplo) e que chegam a prever ações e nossos possíveis interesses por meio de *machine learning*⁷ e absorção de nossa rotina de buscas e navegações mais gerais.

A situação não seria diferente para o mercado erótico, que depende também de ferramentas de busca para organizar os conteúdos e, sobretudo, para atrair seu público consumidor. É importante dizer, como analisei em minha tese de doutorado (Parreiras 2015), que a relação entre ferramentas de busca e mercado erótico não é das mais simples, na medida em que, a partir de uma série de moralidades que acompanham este tipo de conteúdo, são várias as estratégias

aplicadas por provedores⁸ de busca para cercear o que é lido como erótico e/ou pornográfico. No entanto, todas as plataformas nomeadamente pornográficas têm como parte fundamental as ferramentas de busca, que organizam o conteúdo (também de forma nada neutra) e permitem que as pessoas naveguem por uma base de dados gigantesca. No caso do Pornhub, para além de indexar conteúdo, a busca se transformou na base do Pornhub Insights, que utiliza dados de acesso para construção de análises breves sobre comportamento de usuários, conteúdos mais buscados por hora do dia, região, país ou mesmo em datas celebrativas (o último post no momento em que este artigo foi escrito tinha como mote o Super Bowl 2023, por exemplo).

Assim, após o tuíte do presidente, como consequência quase imediata, o Pornhub começou a receber uma quantidade muito maior do que a média de acessos localizados no Brasil, sendo que a palavra de busca mais utilizada era *golden shower*. Estes resultados, juntamente com uma breve explicação do ocorrido, foram publicados⁹ no dia 07 de março de 2019. Nos dias 06 (dia da postagem do ex-presidente) e 07 de março de 2019, as buscas por *golden shower* tiveram um aumento de 688% e 1896% respectivamente, apenas nos domínios do Pornhub. Os dados disponíveis também mostram que houve um aumento geral 98% (06/03) e 173% (07/03) nas buscas globais por *golden shower*, alavancadas pelo tráfego gerado por conexões advindas do Brasil.

Decidi iniciar esta contextualização por este fato porque ele indica pontos que estão no cerne do argumento deste artigo. O primeiro é como a discussão sobre pornografia e seus limites revela situações que extravasam o domínio da sexualidade e do gênero, pautando e refletindo moralidades, ansiedades e políticas. Como apresento e analiso no item seguinte, o disparador para este artigo foi uma postagem do Pornhub Insights em que se revelava que a palavra "novinha" figurava em anos sucessivos entre as mais buscadas por brasileiros/as na plataforma. Esta postagem circulou em redes sociais e diferentes sites conservadores e progressistas, sendo que havia em ambas as partes do espectro político uma associação quase imediata da busca por "novinha" em um site pornográfico com pedofilia ou a revelação do desejo secreto — e, por isso, reservado a um site pornográfico — do brasileiro por jovens meninas.

A postagem do presidente e sua evocação do *golden shower* têm muito a nos dizer sobre o que estamos qualificando como pornográfico e como isso se reflete, em termos tecnológicos, no aumento do acesso a sites declaradamente pornográficos, como o Pornhub. Em certo sentido, a cruzada moralizante resumida naquele tuíte também aguçou a curiosidade geral, o que nos faz questionar se não teve, em alguma medida, efeito contrário àquele pretendido por seu emissor. Sempre que retomo este caso fico imaginando quantas dessas pessoas, que a big data nos apresenta como números absolutos e porcentagens, de fato nunca tinham

ouvido falar no *golden shower* (ou já tinham visto, experimentado, desejado, mas sem dar uma nomenclatura — e lembremos que essa nomenclatura vem do mercado erótico). Assim, temos aqui um primeiro indício da complexidade desta discussão, sendo que são estes tensionamentos e contradições — sem pretensão alguma de resolvê-los —, evocados pelas categorias, nomenclaturas e sua forma tecnológica, simbolizada pelas *hashtags* e formas de marcação algorítmica, que gostaria de desenvolver no restante deste texto.

Podemos afirmar, desse modo, que há uma relação importante entre pornografia e big data, sendo que a escala, os grandes números e essa contabilização de dados quantitativos têm muito a nos dizer sobre a vida cotidiana e as práticas dos sujeitos. Neste sentido, estamos falando de um contexto de datificação da própria pornografia, bem como da existência — importante, sem dúvida, mas que deve ser problematizada — de uma empresa/plataforma que vem fazendo um uso estratégico de seus dados de acesso, criando inclusive uma parte específica de sua organização para lidar com estes dados.

É importante frisar que temos acesso apenas ao que a empresa decide tornar público e esta é uma informação fundamental em termos metodológicos. Isto significa que grande parte da análise desenvolvida neste artigo está baseada em uma quantificação à qual não tenho acesso direto. Trata-se de um levantamento realizado por uma empresa/plataforma, com fins mercadológicos e a partir de interesses que não são os de pesquisa acadêmica, por exemplo. Isto aponta para a necessidade de contextualizar esses dados, bem como a empresa/plataforma que nos permite acesso a eles, evitando desta forma análises apressadas como as que proliferaram em redes sociais quando os dados anuais de acesso a conteúdo pornográfico por brasileiros/as, através do Pornhub, foram divulgados.

Sempre é muito difícil para nós, pesquisadoras, ter acesso aos bastidores de plataformas digitais — o que lembra muito a "caixa preta" proposta por Latour, ou mesmo a "ansiedade terminológica" (:2) causada pelos algoritmos e por nossa tentativa de entender o que eles são, mencionada por Seaver (2017). No geral, temos acesso apenas ao que é tornado público pelas plataformas em termos de dados de acesso, políticas de uso e da própria estrutura de funcionamento, tratada, muitas vezes, como secreta para garantir a exclusividade de mercado e a competitividade da empresa.¹⁰ O mesmo ocorre com grandes plataformas de conteúdo pornográfico como o Pornhub, o que torna sempre desafiador qualquer análise que pretenda ir além do conteúdo, das representações pornográficas que são veiculadas, ou mesmo da parte visível e acessível de seu site.

Como mostra MacDonald (2019), os modos de operação do Pornhub são intrincados, visto que de fato há o oferecimento de uma grande quantidade de conteúdo pornográfico sem a cobrança de qualquer valor monetário. De acordo com ela, ao aderir a uma lógica de plataforma, o Pornhub demonstra pouca clareza

em relação ao seu funcionamento. Algo que chama a atenção é o fato de ser uma empresa que pertence a um grupo de tecnologia (e não do mercado erótico estritamente) mais amplo chamado MindGeek (atualmente Aylo), que também é a dona do Xvideos e do RedTube, com funcionamento similar ao Pornhub. Além disso, também é a proprietária de conhecidos sites pagos, como Digital Playground, Men.com e Mofos. A manutenção de sites sem cobrança, como o Pornhub, se dá então pela geração de fluxo e tráfego entre seus outros sites (apresentando, por exemplo, trechos curtos de vídeos que podem ser assistidos na íntegra e em melhor qualidade de som e imagem mediante pagamento), pela veiculação de anúncios (presentes em todos os vídeos gratuitos) e por meio dos dados gerados pelos milhões de acessos diários. Este último ponto é uma das características mais marcante do capitalismo plataformizado: a aparência de gratuidade quando, na verdade, os dados — invisíveis e, por vezes, desconhecidos pelo/a usuário/a — são a moeda de troca mais valiosa.

Ao consultar os termos de uso do Pornhub, escondidos em face da abundância de imagens, estáticas e em movimento, de corpos e atos sexuais, em letra minúscula, em um item chamado "Trust and Safety", a questão da coleta de dados é mencionada, mas de forma um tanto confusa. A plataforma se compromete a não coletar dados pessoais, como nome, e-mail e senha, a menos que seja judicialmente interpelada para tal. No mais, os dados de acesso em si não constam dos Termos de Uso. Em letras amarelas, indicando o acesso para um link, consta destes termos a Política de Privacidade,¹¹ na qual fica mais clara a política aplicada para o uso de dados. Neste item, lê-se:

Para usuários não cadastrados na plataforma:

Dados de atividade no website: coletamos informações de como você usa nossos sites, produtos e serviços e interage com nosso conteúdo e anúncios, incluindo as páginas que você visita em nossos sites, histórico de busca e a página de referência pela qual você chegou aos nossos sites. Coletamos informações de navegador e sistema operacional, dispositivos que você usa para acessar os sites e sua configuração de fuso horário. Nós também coletamos identificadores on-line. Especificamente, coletamos localização de IP (internet protocol) e habilitamos cookies, como explicado abaixo na seção Cookies and Automatic Data Collection Technologies.¹²

Para usuários cadastrados na plataforma (registro simples ou premium):

Dos usuários registrados coletamos as mesmas categorias de informação descritas acima para os usuários sem registro.

Informações pessoais submetidas pelo usuário: Coletamos informações que você submete a fim de personalizar sua conta ou para uma função específica, como, por exemplo, data de nascimento, idade, gênero, seus interesses, preferências, feedback, respostas a surveys, suas preferências de recebimento de nosso marketing ou de parceiros e suas preferências de comunicação, bem como qualquer outra informação que você nos fornece voluntariamente para uma função específica.¹³

A partir deste excerto dos termos relativos à privacidade e à coleta de dados transcritos acima, é possível começar a compreender a gestão de dados realizada pela plataforma. O primeiro ponto que chama a atenção é o uso de sites no plural, sendo que isto indica o caráter "opaco" (MacDonald 2020) das operações do Pornhub, bem como sua lógica plataformizada, em que os dados fornecidos – vale dizer que, na maior parte das vezes, com desconhecimento dos usuários – durante o acesso a um site específico acabam servindo de indicadores (e identificadores) para outros sites que fazem parte da MindGeek.

Como mostram van Dijck, Poell e de Waal (2018), ao buscarem compreender o que chamam de "platform societies", plataformas não devem ser entendidas unicamente como um fenômeno econômico e nem como um construto apenas tecnológico, mas sim a partir da "relação inextricável entre plataformas on-line e estruturas sociais"¹⁴ (:2). Outro ponto trabalhado pelos autores é a ideia de que se cria, a partir das *big techs*, um "ecossistema de plataformas" (:4), em que a principal característica é a conectividade entre elas, tornada possível sobretudo pelo compartilhamento de dados. Assim, no modelo proposto por esses autores, na base técnica de funcionamento de uma plataforma estão a coleta automatizada de dados, por meio de algoritmos e de interfaces específicas, e um modelo de negócios baseado na posse destes dados. Nesta lógica, os dados são continuamente compartilhados pelo ecossistema, proporcionando o acesso a detalhes dos usos e dos comportamentos dos/as/es usuários/as/es. A partir desses dados, é possível não só a personalização de conteúdos – também citada nos termos de uso e aparentemente disponível apenas para usuários/as/es pagantes do Pornhub –, mas também a geração de métricas e levantamentos com diferentes fins. É exatamente aí que entra o Pornhub Insights, fazendo uso destes dados para criar perfis de usuários/as/es para a plataforma e compartilhando partes selecionadas desses dados com um público mais amplo.

A partir dos itens "Purposes for which we use your personal information" e "Cookies and automatic data collection Technologies", podemos avançar nessa reflexão sobre o uso de dados pela plataforma, o que influi diretamente nos modos como a análise é feita neste artigo, na medida em que ela se baseia, em grande

medida, em dados em segunda mão e sem qualquer influência da pesquisadora. Ainda que sejam "opacas" as políticas e as estruturas de funcionamento da plataforma, o mais perto que conseguimos chegar de uma maior clareza sobre essas questões é por meio dos Termos de Uso e da Política de Privacidade. Isto porque é nestes documentos — também opacos e quase invisíveis — que encontramos indícios de como o Pornhub e a MindGeek coletam e fazem uso dos dados de acesso. É importante notar que uma série de acusações e batalhas legais¹⁵ mundo afora foram responsáveis por mudanças nestes termos, especialmente no que se refere à hospedagem de conteúdo considerado abusivo, tais como vídeos em que não é possível garantir a idade dos/as participantes ou mesmo que houve consentimento para a filmagem; e os que seriam enquadrados na categoria *revenge porn*.¹⁶ Curiosamente, em termos de política de dados e da plataformação, não há grandes contenciosos públicos e muito menos produções a respeito, com exceção do importante trabalho de Margaret MacDonald (2020). Para levar à frente estas reflexões, transcrevo abaixo outras partes da Política de Privacidade:

Propósitos para os quais usamos suas informações pessoais

Análise: Usamos identificadores e dados de atividade do site para determinar se os usuários de nossos sites são exclusivos ou se um mesmo usuário está usando os sites em múltiplas ocasiões e para monitorar métricas agregadas, tais como número total de visitantes, páginas vistas, padrões demográficos.

Cookies e tecnologias de coleta automática de dados.

Na medida em que você navega por nossos sites e interage com eles, usamos tecnologias de coleta automática de dados para coletar dados de atividade do site. Usamos cookies, pequenos arquivos de texto que são armazenados em seu navegador ou baixados em seu dispositivo quando você visita um site. [...]

[...] Uso do Google Analytics: usamos o Google como provedor do serviço de coletar e analisar informações sobre como os usuários utilizam os sites, incluindo a coleta de dados de atividade por meio de cookies primários definidos por nossos domínios e cookies de terceiros definidos pelo Google. Porque ativamos a anonimização de IPs para o Google Analytics, o Google irá anonimizar o último octeto de um endereço de IP específico e não vai armazenar seu IP inteiro. O Google só irá usar as informações para nos prover os serviços do Google Analytics e não vai utilizar estas informações para outros fins. As informações coletadas pelo Google Analytics

podem ser transmitidas e armazenadas pelo Google em servidores nos Estados Unidos nos termos das cláusulas contratuais padrão aprovadas pela União Europeia.¹⁷

Ainda que o PornHub Insights não seja mencionado em nenhum momento nestes termos transcritos, eles nos ajudam a depreender de que modo é feita a coleta de dados, bem como os propósitos da plataforma para tal. Em todo o documento relativo à privacidade, o PornHub Insights aparece apenas uma vez e de forma bastante curiosa:¹⁸ "podemos usar seus dados para produzir e compartilhar insights agregados que não o identificam direta ou indiretamente e não são associados com você. Essa informação não é informação pessoal". Por esta afirmação, podemos inferir o que a plataforma entende como dados pessoais, sendo que, a partir do momento em que são anonimizados e agregados em conjuntos — e geram, então, isto que leva o nome de big data — eles são despersonalizados.

De fato, se pensarmos nas grandes quantidades de dados geradas continuamente pelos inúmeros acessos, o primeiro pensamento é que são dados impossíveis de serem identificados. No entanto, como teóricos como boyd e Crawford (2012) vêm mostrando, essa impossibilidade não é real, na medida em que estes dados podem ser relacionados com outros dados e não há garantia de que a anonimização automática de fato acontece. Como os excertos transcritos mostram, a operação do PornHub se dá tanto por cookies próprios e específicos — sobre os quais sabemos pouco ou nada — e com o uso de cookies de terceiros, com a utilização do Google Analytics. Ainda que, ao final do texto da Política de Privacidade, seja indicado um link no qual é possível entender melhor o funcionamento dessa ferramenta do Google, continuamos sem muitos indícios do que é realmente coletado ao longo dos acessos e se há a remoção de identificadores pessoais quando estes dados são armazenados e analisados.

Toda essa discussão é relevante na medida em que aponta para a necessidade de sempre adotar uma postura crítica quando falamos de big data e, especialmente, quando vamos lidar, em nossas pesquisas, com dados provenientes diretamente das bases públicas das plataformas. Os dados do PornHub Insights, desse modo, são completamente enviesados e imbuídos de uma lógica de mercado, sendo que os principais interesses por detrás da coleta de dados é fornecer subsídios que permitam à plataforma atingir mais pessoas e gerar maiores dividendos. Há, então, como afirmam Nieborg e Poell (2018), uma constante assimetria de poder entre plataformas e usuários. É uma das bases dessa assimetria a já apontada opacidade, que trata, como mostra MacDonald (2020), de manter a "infraestrutura de um site pornográfico intencionalmente invisível" (:11).

Foi com tudo isto em mente que me lancei à empreitada de compreender como categorias e rankings são apresentados pelo PornHub Insights e, especificamente,

como o Brasil é retratado nestes enquadramentos gerados pela plataforma. Para tal, sigo o fluxo de uma categoria específica — a "novinha" — a partir dos contenciosos e das manifestações públicas que ocorreram após a divulgação de rankings anuais de palavras mais buscadas onde esta categoria figura entre aquilo que "os brasileiros" mais procuram na plataforma.

A ideia de fluxo é importante porque marca uma postura metodológica. Em minhas pesquisas anteriores (Parreiras 2008, 2012, 2015, 2018) insisti bastante em uma compreensão do digital que estivesse centrada nos usos feitos pelos sujeitos, mas que, sobretudo, seguisse fluxos — de sujeitos, identidades, subjetividades, espaços, categorias, plataformas, infraestruturas, dados. Foi o que busquei fazer, por exemplo, ao lidar com o Twitter e o modo como ele foi fundamental para a produção, a divulgação e o consumo de pornografia alternativa, sendo que os sujeitos que encontrei em campo estavam em constante fluxo entre plataformas (ou redes sociais, como chamei naquele momento) e entre as próprias possibilidades de vivência das sexualidades e dos desejos. Para este artigo, parto do mesmo princípio, seguindo de diferentes formas o modo como uma categoria específica do pornô, mas não apenas dele, circula on-line e demarca diferentes significados.

Foi no ensejo de me deixar levar pelos fluxos — algo quase impossível à primeira vista se pensarmos que a base da argumentação são os dados do PornHub Insights — que decidi ir além dos gráficos e das análises provenientes da plataforma. Afinal, tinha sido em meio a fluxos que eu mesma tinha chegado a este tema e formulado minhas inquietações, sobretudo a partir de inúmeras publicações do ranking anual de 2014 no Facebook, no Twitter e em blogs e sites de notícias de grandes veículos de comunicação. Assim, resolvi mapear de forma experimental e etnograficamente a circulação do termo "novinha" no Twitter, a fim de compreender seus diferentes significados, bem como seus usos pornográficos como categoria indexadora. Para tal, fiz uso da extensão NCapture do software Nvivo, disponível para instalação no navegador Google Chrome. Estabeleci como período de coleta o tempo demarcado desde o convite para compor este dossiê (outubro de 2022) até o final de janeiro de 2023. A partir dos dados gerados, construí uma tabela com as diferentes formas de aparição do termo, o que apresentarei no item seguinte. De todo modo, o que este experimento mostra é não apenas o caráter polissêmico da palavra novinha, mas também o quanto são fluidas as posições de sujeitos, práticas e comportamentos que ela demarca em termos bastante práticos. Vale ressaltar que este levantamento não tem finalidade estatística ou de geração de métricas, mas representa um modo de mapear sentidos adquiridos por uma categoria quando imersa em relações sociotécnicas.

A escolha pelo Twitter se justifica pela maior facilidade de realizar levantamentos nesta plataforma, especialmente devido à sua API aberta. Saliento, no entanto, que, após a compra da plataforma por Elon Musk, vem ocorrendo uma série de

mudanças nas operações, sendo que uma delas é a cobrança do acesso à API para a realização de pesquisas, métricas e levantamentos, sejam eles acadêmicos ou mercadológicos. Assim, é algo ainda a ser compreendido o que estas mudanças irão significar para as formas como utilizamos o Twitter para cartografias e estudos quantitativos.¹⁹

Entendo que este é um estudo experimental etnográfico. Acredito que a etnografia, entendida não como método canônico e manualesco, mas como um conjunto múltiplo de técnicas de pesquisa que são relacionais e abertas à experimentação, é uma forma rentável de lidar com dados que conformariam isto que chamamos de big data. Minha proposta é, assim, ir além do dado numérico, representado sobretudo pelas porcentagens que demarcam os muitos temas relacionados à pornografia e iluminados por ela. Em última instância, estou interessada nas práticas e nos significados que a categoria analisada demarca, buscando ultrapassar explicações rápidas e baseadas apenas em uma suposta realidade do pornográfico que deixa de fora sua base na representação e no jogo de fantasias e desejos.

Por fim, em termos éticos, o levantamento realizado utilizou como base apenas o conjunto de tuítes públicos no período selecionado. Não houve coleta de nenhum dado pessoal (nome de usuário, localização, foto de perfil ou bio), sendo que eles foram descartados da análise. Esta é uma prática bastante corrente entre os estudos para o digital e que eu mesma utilizei em minha pesquisa de doutorado (Parreiras 2015). No próximo item discuto alguns dos achados deste experimento, mostrando certos vieses da plataformização da pornografia e dos modos como isto ocorre em relação ao Brasil.

Insights sobre o Brasil em dados do pornô plataformizado: a categoria “novinha” e seus desdobramentos

Desde que se iniciou o debate sobre o que foi nomeado como big data, há uma preocupação em situar tanto o que se entende por este termo quanto os desafios metodológicos e éticos para seu uso. Neste artigo, o ponto de partida conceitual em relação à big data está de acordo com o proposto por boyd e Crawford (2012) e Manovich (2011), que visam pensá-la não a partir da característica que mais chama a atenção, a quantidade maciça de dados, mas sim a partir da relacionalidade destes dados com outros. Neste sentido, a big data permite mapear conexões e relações entre os sujeitos, grupos sociais, tecnologias e estruturas da informação. Uma pesquisa etnográfica sobre big data, assim, pode se servir dessa capacidade de traçar conexões entre os dados, ainda que eles tenham como fonte, como no caso do PornHub, uma plataforma.

É também de uma provocação feita por boyd e Crawford (2012) que deriva a inspiração para a discussão aqui proposta. De acordo com as autoras, lidar com big data exige colocar em questão nossos "pressupostos analíticos, quadros metodológicos e preconcepções". Sobretudo, em minha leitura, as autoras nos instigam a ir além da coleta automatizada de dados e do uso dos softwares de análise, direcionando "perguntas críticas" sobre o significado dos dados coletados e aplicando um olhar subjetivo para seu entendimento.

Como dito anteriormente, a análise empreendida se baseia no uso da palavra "novinha", que figura, de acordo com o Pornhub Insights, entre as mais buscadas pelos/as brasileiros/as no Pornhub em diferentes anos. Este termo funciona tanto como palavra-chave de busca, como um desejo/fantasia/fetiche a ser consumido e uma forma de indexação de conteúdo dentro da plataforma.

Em termos tecnológicos, o uso de palavras-chave, ou mesmo *hashtags*, como marcadores classificatórios e de indexação (sejam elas visíveis ou não, como é o caso das *tag clouds*), está de acordo com um dos princípios que regem a estruturação da web desde sua versão 2.0: a chamada folksonomia.²⁰ Em linhas gerais, a folksonomia consiste em um modo de indexar o conteúdo disponível on-line por meio de marcadores, que são colaborativos e coletivos. Isto significa que, apesar de as categorias serem criadas por programadores e desenvolvedores, as marcações são ressignificadas e se sustentam a partir do uso amplo e da criação de sentido para um maior número de usuários. Deste modo, há uma classificação coletiva de informações, que incide diretamente nas ferramentas de busca e nos modos como se dão as circulações e as conexões.

Esse processo de marcação é importante, porque permite realizar levantamentos de acesso e uso da web. Por meio de uma palavra-chave, por exemplo, é possível estabelecer um mapeamento de dados e criar métricas para análise dos mesmos. É por este princípio que podemos saber de onde partem os acessos, a popularidade, a relevância, ou como determinados temas circulam on-line. No caso de sites em que há oferecimento de produtos e atividades comerciais, é fundamental que seja criado um eficiente sistema de marcação, o qual faz com que os dados ali presentes sejam mais facilmente encontrados. Munhoz (2024, em artigo neste dossiê) apresenta uma reflexão semelhante ao apontar que as ferramentas de busca devem "antever expressões que os usuários mobilizarão", além de criar relações entre os conteúdos indexados.

Patterson (2004), ao analisar o consumo de pornografia na era digital, traz algumas considerações sobre o ato de classificar e categorizar os materiais encontrados on-line. De acordo com ele, o pornô on-line funciona por meio da rápida criação e propagação de categorias, as quais permitem ao sujeito consumidor/espectador assimilar e emular uma posição particular em relação aos materiais, além de criar uma "promessa alucinatória de fluidez". Eu iria além

neste argumento, na medida em que as categorias, ainda que sejam fluidas em muitos momentos, apontam também para certa fixidez, passando a demarcar de forma muito contundente determinados gêneros de produções, formas de produzir e até mesmo a associação de certas categorias com posições de sujeito e imaginários sociais mais amplos.

Assim, as categorias, enquanto termos de descrição, têm a função de guiar o usuário em relação ao tipo de conteúdo encontrado. Elas funcionam como um resumo do que será visto nos vídeos e nas imagens e ajudam a direcionar o olhar a partir de determinadas convenções. A escolha de trazê-las aqui se deve à minha crença de que, assim procedendo, é possível entender com mais profundidade a lógica de organização do site e as interpretações possíveis destas categorias. Classificar é um ato lógico e que mobiliza operações de pensamento, não sendo, de modo algum, um processo aleatório. O que encontramos no compilado de dados do PornInsights caminha em sentido semelhante e, como venho defendendo, nos ajuda a compreender questões sociais mais gerais ou mesmo os tensionamentos e as moralidades em jogo em um determinado momento.

Como dito anteriormente, o *insight* para este experimento partiu do incômodo gerado pela publicação, em 2018, sobretudo em plataformas de redes sociais (Facebook e Twitter), de um gráfico que indicava a palavra mais buscada por alguns países no Pornhub em 2014. É interessante notar que este gráfico foi originalmente publicado no site da revista *The Economist*,²¹ em 26 de setembro de 2015, seguido de um tuíte²² no perfil oficial da revista em 29 de setembro no mesmo ano. Com o título "The Porn Supremacy" (Figura 1), o gráfico era uma construção feita a partir dos dados do ano de 2014 publicados pelo Pornhub Insights. Indicavam-se não apenas os termos mais buscados por país, mas também a quantidade total de acessos à plataforma Pornhub naquele ano: 18 bilhões. Outro dado era a porcentagem de acessos de cada país dentro do total. Reproduzo abaixo este gráfico, porque ele ajuda a visualizar melhor o efeito pretendido pela revista, bem como justifica porque ele foi o disparador de minhas reflexões.

Se consultarmos o Pornhub Insights, fonte original dos dados, notaremos algumas inconsistências em relação à tabela. Os dados utilizados são provenientes da publicação, bastante popular e das mais completas, que leva o nome de "Year in Review" e traz informações que resumiriam o ano em "tendências de busca" e "como as interações dos usuários reagem a grandes eventos mundo afora".²³ O gráfico criado pela *The Economist* sumariza outros gráficos criados pelo Pornhub (Figura 2). Como é possível notar, a revista exclui termos de nacionalidade (como "brazilian" e "brazil") e, por isso, "novinha" aparece como termo mais buscado. Mas, como a publicação original no Pornhub Insights (Figura 2) deixa claro, este não foi o termo mais buscado em 2014 por brasileiros/as.



Economist.com

Figura 1. Termos mais buscados em alguns países no PornHub em 2014.



Figura 2. PornHub Insights, 2014.

Como podemos notar, os termos mais buscados foram, na verdade, "brazil" e "brazilian". Novinha aparece como terceiro, ao contrário do que o resumo da *The Economist* sugere. É importante lembrar que este gráfico foi posteriormente, em especial ao longo da campanha eleitoral de 2018, apropriado fora de contexto, simplesmente reproduzido em postagens em redes sociais com a associação rápida e sem questionamentos entre a busca pela palavra "novinha" e o caráter "pedófilo" do brasileiro ou, como li certa vez em uma destas publicações, "o brasileiro é o que bate punheta para criança".²⁴

Aqui creio que já podemos notar algumas inflexões interessantes em relação ao termo novinha. A primeira delas é sua associação imediata com infância e criança, o que, em termos de representação pornográfica que é encontrada nos sites, não se justifica. Novinha, para o pornô, aparece como quase sinônimo de "teen"/"teenager", que são categorias clássicas do pornô *mainstream* (e que, como mostram os dados do Pornhub, também aparecem como palavras muito buscadas em outros países). Não estaríamos aqui, para além da nomeação rápida de pedofilia, diante de um tensionamento dos "limites da sexualidade" (Gregori 2016), a partir do consumo/aceso a certos tipos de produções de pornografia? Antes de aventar possíveis respostas a esta pergunta, creio ser necessário dar alguns passos atrás, mostrando meus próprios caminhos e fluxos em campo.

Ao checar dados de outros compilados anuais, é notável que este padrão de buscas se mantém, com pequenas variações certamente causadas por acontecimentos específicos (como o já citado *golden shower*). Também é notável o interesse da própria plataforma por tudo que se refere ao Brasil (em termos de eventos e acontecimentos) ou aos padrões de busca e comportamentos sexuais dos/as brasileiros/as.²⁵

Ao acessar o Pornhub Insights e restringir a busca de artigos e levantamentos a partir da palavra-chave "Brazil", nos é apresentada uma quantidade considerável de dados. O mais recente deles diz respeito ao tráfego gerado pela busca de produções pornográficas durante os meses iniciais da pandemia de COVID-19, sendo que o Brasil é um dos países que entram no comparativo realizado que buscava mostrar o quanto, diante da necessidade de isolamento social e com um maior quantitativo de pessoas em casa, houve um aumento considerável na quantidade de buscas e acessos por pornografia.

Em termos dos relatórios anuais (The Year in Review), o último em que o Brasil aparece é o de 2022 e o primeiro é de 2014. Vale a pena mencionar que há flutuações e variações nos relatórios, sendo que nos anos mais recentes um foco interessante tem sido no tipo de dispositivo utilizado para acesso ao site e em comparativos em relação ao conteúdo buscado por usuários/as/es de diferentes faixas etárias. No entanto, novinha (e *teen* em alguns anos) e *brazilian*/brasileira aparecem em quase todos eles. Mesmo não sendo comuns a todos, não deixam

de ser menos relevantes para os objetivos aqui pretendidos. Para dar um exemplo dessas mudanças de padrões, no relatório de 2022, chamam a atenção para o crescimento acentuado de busca por "trans/transgender" advindo do Brasil, algo ausente em quase todos os anos anteriores. Além disso, pela primeira vez, hentai foi o tipo mais buscado de vídeo por brasileiros/as. A categoria novinha (ou *teen*) não aparece neste relatório.

Algo que chama a atenção é como a delimitação numérica de milhões de dados de acesso fixa comportamentos e a própria definição de qual seria a relação do/a brasileiro/a com a pornografia. Os números, apresentados em gráficos coloridos e esteticamente impecáveis, parecem sugerir que é possível sumarizar estes comportamentos e definir uma quantidade enorme de sujeitos por aquilo que consomem. Se retirado de contexto — como é o caso das publicações já citadas em relação à novinha —, a situação se torna ainda mais complexa, na medida em que estamos falando de algo que, pela legislação brasileira, é um crime, qual seja, a pedofilia.

De modo algum estou sugerindo que não encontramos no Pornhub práticas que poderiam ser enquadradas como "pornografia infantil"²⁶ ou outros atos violentos e abusivos. Como é permitido o *upload* de vídeos por usuários/as, o Pornhub tem enfrentado uma série de batalhas jurídicas em diferentes países quanto a estes conteúdos e realizado, em consequência disto, a exclusão de uma grande quantidade de vídeos.

Lowenkron (2015), a partir de sua análise da CPI da Pedofilia, realizada no Brasil entre os anos de 2008 e 2010, mostra o quanto a categoria pedofilia está imersa em disputas. Assim, para entender a construção da pedofilia como problema social, é necessário levar em consideração diferentes planos que envolvem "deslocamentos de sensibilidades, modelos de inteligibilidade, categorias e tecnologias de controle social" (:35). É na interseção de diferentes "planos" de análise" que podemos situar a figura do pedófilo, o "monstro contemporâneo", como sugere Lowenkron. Deste modo, pedofilia e pedófilo, que delimitam atos e marcas reais, são categorias cultural e socialmente modeladas (Hacking 2013), que respondem a interesses e anseios de uma época específica, com base em diferentes discursos (médico, jurídico, político) e que possuem diversos desdobramentos práticos na vida social. É possível, então, associar a "novinha" buscada no Pornhub com pedofilia?

Ao realizarmos uma busca simples da palavra "novinha" no Pornhub — feita como usuária sem cadastro, com IP associado ao Brasil (sem uso de VPN) e com configurações automáticas da própria plataforma (linguagem português e geolocalização no Brasil) — aparecem mais de 10 mil vídeos, 122 fotos (as mais relevantes apenas e todas associadas a perfis de atrizes e performers) e 598 *gifs*. Na página inicial de busca são mostrados os vídeos considerados "mais relevantes" pela própria plataforma, índice provavelmente dado pela popularidade e pela

quantidade de acesso aos mesmos. O vídeo mais assistido nesta categoria, com mais de 46 milhões de visualizações, é de uma conhecida produtora pornográfica ligada ao *mainstream* e que se anuncia como uma das poucas que empregam modelos que representam "uma nova era do erotismo". Esta produtora não está localizada no Brasil e, se modificamos a busca para inglês, o mesmo vídeo reaparece, desta vez dentro da categoria "teen". Disso se depreende que, para a plataforma, novinha e *teen/teenager* são sinônimas e agrupam um mesmo tipo de corpos e atos sexuais em exibição. Em diversos outros vídeos a mesma situação ocorre, com a tradução automática entre os termos.

O jogo pretendido pelas produções encontradas dentro destas categorias parece claro: são vídeos e/ou fotos que jogam com o critério etário, com relações em que há assimetria de idades entre as pessoas envolvidas, sendo sempre uma mulher muito jovem — e as atrizes escolhidas para encenar estas produções costumam ter rostos e corpos que aparentam a pouca idade — e um ou vários homens mais velhos. Os roteiros todos jogam com o interdito e com a proibição da relação, exatamente pela existência de uma diferença de idades acompanhada de uma relação de poder assimétrica. A novinha é, assim, esse interdito, o proibido a ser conquistado. Vale dizer que, pelas regras da plataforma, só são permitidos vídeos em que apareçam pessoas maiores de idade (e essa idade legal varia entre os países, o que é também um campo de batalha), que consentiram com as filmagens e os atos encenados e que, se ligadas a produtoras, receberam pagamento por seu trabalho.

Não compactuo com uma visão da pornografia e do mercado erótico que enxerga apenas opressão, abuso e violência. Do mesmo modo, reconheço que se trata, especialmente quando falamos da pornografia *mainstream*, de um mercado altamente misógino e em que acontecem muitos atos e formas de violência. Mas, desde que comecei a pesquisar esta temática, em 2009, algo que sempre reitero é que há um lado da representação que os julgamentos apressados da pornografia deixam de lado. Mas onde a novinha entra em tudo isso?

A palavra novinha é polissêmica e ganha diferentes sentidos a partir das relações em que seu emprego aparece. Nos já citados posts e comentários ao gráfico que associam o brasileiro que procura pornografia em um site gratuito a um pedófilo, é evocado um sentido que iguala a novinha a crianças, o que não aparece nos vídeos. Nem mesmo para o algoritmo que faz traduções e associações na plataforma há essa relação, mas sim com adolescentes.

Em artigo anterior (Parreiras, no prelo), mostrei que a novinha é uma das muitas formas de nomear formas de vivência e experiência do período do curso da vida que chamamos de adolescência e que, em sua origem, é delimitada por marcadores de classe social e local de origem/moradia. É usualmente utilizada para se referir a mulheres adolescentes de classes sociais mais baixas e de locais

periféricos e empobrecidos. A partir de cenas etnográficas de uma pesquisa de campo realizada em favelas da cidade do Rio de Janeiro e de revisão de estudos que mencionam ou analisam a novinha (Fernandes 2017), argumento que a figura da novinha tensiona os limites da sexualidade (Gregori 2016), na medida em que, por um lado, evoca tentativas de moralização, controle, vigilância e tutela e, por outro, aponta para relações de constituição de sujeito de jovens mulheres que têm direito à vivência positiva e saudável da sexualidade.

Neste sentido, a novinha transita entre prazer e perigo e nas armadilhas deste entremeio, sendo uma delas os discursos moralizantes e tutelares. É importante dizer que o Estatuto da Criança e do Adolescente (ECA) se assenta na premissa de que crianças e adolescentes são sujeitos e não objetos de tutela do Estado, das famílias e da comunidade. Não quero com isso sugerir que novinha está relacionada apenas a prazer e desejos, mas que pode sim aparecer em situações de violência e abuso dos mais diversos tipos. Minha proposta é descolar esta nomenclatura dos dois polos demarcados, mas pensar nos trânsitos entre eles.

Como apresentei anteriormente, em 2018 e 2019, os dados de que novinha era uma das categorias mais buscadas por brasileiros no Pornhub veio à tona nas redes sociais. Tentei deixar claro que há uma polarização bastante nítida no cenário político brasileiro, em que um dos lados é partidário de uma cruzada moralizante e conservadora que tem como projeto o controle das sexualidades e, especialmente, da sexualidade de crianças e adolescentes. Assim, os dados do Pornhub — que foram descolados da postagem original feita nos Insights — foram apropriados e veiculados de diferentes modos e por ambos os lados da polarização. No geral, os argumentos eram simplesmente condenatórios, tomando a novinha como categoria encapsulada de uma adolescente sexualizada e alvo de abusos e violências ou de uma criança a ser salva.

Em meu mapeamento experimental, outros sentidos da novinha apareceram, bem como houve a associação do termo com práticas sexuais e a sexualização de determinados corpos. No período selecionado — entre outubro de 2022 e janeiro de 2023 — a extensão NCapture encontrou mais de 13 mil menções à palavra novinha, sendo excluídas postagens duplicadas e retuítes e incluídos os *usernames* (as @s) que porventura utilizassem o termo. Cinco sentidos principais da palavra foram encontrados: a simples indicação da idade e da adolescência (inclusive com brincadeiras com categorias como *boomers*, *millenials* e geração z); os usos sexuais; a utilização como adjetivo e forma de se referir a outras @s; a postagem de alguma música (um funk quase sempre) em que a letra fala da novinha; e a aplicação no *username* para indicar se tratar de um perfil de uma mulher jovem.

Entre as postagens que traziam alguma conotação sexual, chamou a minha atenção o quanto, no geral, eram postagens de perfis que buscavam trocas de nudes, imagens eróticas enviadas por meio de comunicadores ou redes sociais. No geral,

os perfis apresentavam descrições que fazem concluir serem de adolescentes — ainda que anônimos — em busca de interações sexuais com outros/as adolescentes. Foi criado, inclusive, um conjunto de *hashtags* que comumente acompanham este tipo de postagem e que, de tempos em tempos, abrem votações para escolher os melhores nudes, no que eles chamam de *#awards*. Para evitar a identificação de qualquer perfil ou as redes por onde perambulei, não transcrevo aqui as *hashtags* específicas, mas todas elas mencionam alguma parte fetichizada do corpo ou um ato sexual. De todo modo, a criação destes indexadores facilita encontrar pessoas que estão em busca de um mesmo tipo de conteúdo ou interação e, quem sabe e como muitas postagens mencionam, estejam dispostas a continuar a relação e as trocas pelo *direct message* (a DM) da plataforma.

Diante de tudo o que foi apresentado, ainda que em caráter exploratório, me parece que essa discussão é mais complexa e que é mais rentável pensar na novinha enquanto categoria tensionadora e potencialmente subversiva quando se refere à demarcação de produções pornográficas. É relevante mencionar que em sites como o Pornhub há uma política de controle do que é veiculado a fim de evitar a exposição de crianças e adolescentes e de relações sem consentimento em sentido mais amplo. As atrizes descritas como novinhas — ainda que corporalmente possam evocar a imagem de adolescentes — são legalmente maiores de idade e com capacidade de consentir. Talvez aqui seja rentável pensar no que Judith Butler (1990) propõe ao questionar os argumentos das feministas antipornografia: o pornográfico e o erótico operam a partir de fantasias. A presença do descritor novinha não nos diz que há na cena ou no filme em questão uma adolescente ou criança que é sexualizada, mas sim que há um jogo erótico em que a simulação é ponto central. Minha argumentação é, portanto, que a novinha deve ser pensada em razão dos limites destes trânsitos entre prazer e perigo e da articulação entre marcadores de idade, gênero e sexualidade. O discurso moralizante por si só não nos permite avançar nesta discussão.

Insights finais

Ainda que de forma bastante experimental, busquei indicar neste artigo alguns dos usos que antropólogos podem fazer da big data. Minha proposta foi, a partir de dados macro do Pornhub Insights e do meu uso da extensão NCapture para mapeamento de postagens no Twitter, fornecer algumas linhas possíveis de análise, com base na realização de etnografia, no que se refere a categorias e formas de classificação — sociais e algorítmicas — de produções pornográficas segundo o contexto brasileiro.

Além disso, tentei mostrar os múltiplos significados que uma categoria — a novinha — pode ganhar, fazendo sentido pensar em sua modelagem social. Para além de posturas que buscam apenas demonizar, condenar ou pensar nos efeitos viciantes e abomináveis da pornografia, meu objetivo é apontar os tensionamentos envolvidos quando lidamos com as sexualidades e suas formas de experiência. Em um momento de pânico morais (Rubin 1986), encampados sobretudo por grupos conservadores, faz sentido buscar um alargamento do nosso campo de visão, pensando inclusive no papel da internet como garantidora da propagação e da escala destes discursos, sobretudo por meio de sua lógica cada vez mais plataformizada.

Tudo isto reitera algo que os estudos sobre pornografia (Attwood 2009, 2017; Paasonen 2007, 2011; Sullivan & McKee 2015; Mulholand 2013; McNair 2013, para citar apenas alguns) vêm apontando há bastante tempo: o quanto a pornografia é um modo valioso — e agora plataformizado e datificado — de se pensarem relações que estão para além dela, sendo fundamental tomá-la como problema social, de forma a escapar de polarizações e visões moralizadoras que buscam apenas combatê-la.

Notas

- 1 É importante observar que na pornografia gay (categoria da própria plataforma) aparecem também categorias como "novinho" e "twink" (homens de aparência jovem ou adolescente). Além disso, no mercado pornográfico, filmes lésbicos para homens são aqueles que apresentam mulheres em relações homossexuais, mas mantendo todo o estereótipo encontrado em produções heterossexuais e destinadas ao prazer de homens.
- 2 Desde julho de 2023, o Twitter passou a se chamar X. Esta foi uma das mudanças introduzidas após a compra da plataforma por Elon Musk. Optei por manter o nome original neste artigo porque esta era a conformação encontrada ao longo do levantamento realizado.
- 3 Não pretendo fornecer uma definição fechada e acabada de pornografia alternativa, visto que esta (assim como os próprios conceitos de pornografia e erotismo) é uma classificação em disputa e relacional. Há um conjunto de produções bastante consolidado sobre pornografia alternativa, em diferentes contextos. Ver, por exemplo, Suzana Paasonen (2011) e Feona Attwood (2009).
- 4 Esta frase pode ser traduzida como "escavando a fundo os dados".
- 5 A escolha pelo Twitter se deu pela maior facilidade de encontrar conteúdos ditos pornográficos, na medida em que esta plataforma apresenta uma política menos limitante em relação à postagem e à divulgação de materiais com nudez ou atos sexuais, por exemplo. Além disso, a API (interface de programação) do Twitter permitia o acesso facilitado aos dados de uso por meio de softwares específicos para esta mineração de dados. É importante salientar que, após a aquisição do Twitter por Elon Musk, houve o anúncio de que o acesso à API, inclusive para fins acadêmicos e de pesquisa, passaria a ser pago, o que aconteceu de fato no começo de 2023.
- 6 Em todos os subtítulos do artigo utilizo a palavra insights (no sentido de conhecimento, informação) como forma de jogar com o nome utilizado pelo Pornhub para se referir aos dados coletados.
- 7 Machine learning, traduzido como aprendizado de máquina, é uma área da inteligência artificial. Em termos bem gerais, refere-se a sistemas que aprendem a partir dos dados. Para mais análises sobre o tema, ver Christian (2021) e Alpaydin (2016).
- 8 Os casos mais conhecidos de censura são o do autocompletar do Google, que exclui palavras com alguma relação com sexo e sexualidade e, mais recentemente, todas as tentativas do Instagram de dificultar a busca de conteúdos de profissionais do sexo – especialmente durante a pandemia com o boom do Only Fans (OF) e o uso de outras plataformas para divulgação e publicidade de vídeos e fotos posteriormente comercializadas no OF.
- 9 O post e a análise realizada pela plataforma podem ser acessados em <https://www.pornhub.com/insights/brazil-golden-shower>. Acesso em 20/03/2023.

- 10 Uma das exceções a esta lógica são os vazamentos que ocorrem recorrentemente, tais como o Wikileaks ou o da Cambridge Analytica, em relação ao Facebook e sua política danosa de dados.
- 11 Disponível em <https://www.pornhub.com/information/privacy>. Acesso em 23/03/2023. A última atualização destes termos ocorreu em 14 de abril de 2022.
- 12 No original (tradução minha): "Website activity data: We collect information about how you use our Websites, products and services and interact with our content and advertisements, including the pages you visit in our Websites, search history, and the referring web page from which you arrived at our Websites. We collect browser and operating system information, devices you use to access the Websites and your time zone setting. We also collect online identifiers. Specifically, we collect internet protocol (IP) address information and we set cookies as explained below in the section on Cookies and Automatic Data Collection Technologies."
- 13 No original (tradução minha): "We collect from registered users the same categories of information described above for unregistered users. User Submitted Personal Information: We collect information you submit in order to personalize your account or for a specific function, for example date of birth, age, gender, your interests, preferences, feedback, survey responses, your preferences in receiving marketing from us and our third parties, and your communication preferences, as well as any other information which you voluntarily provide to us at your direction for a specific function."
- 14 No original (tradução minha): "[...] inextricable relation between online platforms and societal structures."
- 15 Para mais informações sobre as acusações contra o PornHub, ver: <https://cherwell.org/2020/07/24/the-problem-with-pornhub-and-how-they-get-away-with-it/>. Acesso em 25/03/2023; <https://edition.cnn.com/2021/06/17/tech/pornhub-lawsuit-filed/index.html>. Acesso em 25/03/2023; <https://endsexualexploitation.org/articles/the-class-action-lawsuit-against-pornhub-and-mindgeek-explained-2/>. Acesso em 25/03/2023; <https://www.protocol.com/policy/pornhub-lawsuit>. Acesso em 25/03/2023; <https://www.theguardian.com/technology/2020/dec/14/pornhub-purge-removes-unverified-videos-investigation-child-abuse>. Acesso em 25/03/2023.
- 16 Como mostra Lins (2021), o termo *revenge porn*, apesar de sua popularidade, não é o mais adequado. Os vídeos e as imagens enquadrados nesta categoria se referem a vazamentos, sem consentimento ou autorização e, por isso, se configuram como uma forma de violência. É importante lembrar que pornografia implica a representação de relações consentidas e que, se ligadas a empresas, normalmente implicam relações de trabalho e pagamento de cachê.
- 17 No original (tradução minha): "Purposes for Which We Use Your Personal Information Analytics: We use identifiers and website activity data to determine whether users of our Websites are unique, or whether the same user is using the Websites on multiple occasions, and to monitor aggregate metrics such as total number of visitors, pages viewed, demographic patterns. Cookies and Automatic Data Collection Technologies:

As you navigate through and interact with our Websites, we use automatic data collection technologies to collect website activity data. We use cookies, which are small text files that are stored in your web browser or downloaded to your device when you visit a website. Use of Google Analytics. We use Google as a service provider to collect and analyze information about how users use the Websites, including by collecting website activity data through first-party cookies set by our domains, and third-party cookies set by Google. Because we activated IP anonymization for Google Analytics, Google will anonymize the last octet of a particular IP address and will not store your full IP address. Google will use the information only for the purpose of providing Google Analytics services to us, and will not use this information for other purposes. The information collected by Google Analytics may be transmitted to and stored by Google on servers in the United States pursuant to standard contractual clauses approved by the EU."

- 18 "We may use your data to produce and share aggregated insights that do not directly or indirectly identify you and are not associated with you. Such aggregate information is not personal information."
- 19 Em face das mudanças implementadas nos meses iniciais de 2023, acadêmicos de diferentes países se uniram na Coalition for Independent Technology Research, somando esforços para se adequarem à cobrança pelo acesso à API. Pelas atuais regras, a versão gratuita não permite mais a coleta de tuítes. A assinatura mais barata é 100 dólares ao mês e dá acesso a apenas 10 mil tuítes mensais. A coalisão publicou recentemente uma carta aberta sobre esta situação, que pode ser acessada em <https://independenttechresearch.org/letter-twiters-new-api-plans-will-devastate-public-interest-research/>. Acesso em 03/04/2023.
- 20 O termo "folksonomy" foi criado pelo arquiteto da informação Thomas Vander Wal. O nome vem da junção dos termos folk (povo) e taxonomy (taxonomia). Para ele, há dois tipos de processo de classificação: ampla e estreita. A folksonomia ampla é quando uma grande quantidade de usuários usa uma grande quantidade de marcadores para um mesmo conteúdo. Já a estreita ocorre quando marcam um objeto ou informação com um número reduzido de termos. Para uma análise mais detalhada, ver Vander Wal (2005, 2007) e Peter, Isabella (2009).
- 21 Disponível em <https://www.economist.com/international/2015/09/26/naked-capitalism?fsrc=scn%2Ftw%2Fte%2Fpe%2Fed%2Fnakedcapitalism>. Acesso em 05/04/2023.
- 22 Disponível em <https://twitter.com/TheEconomist/status/648956186906263552?s=20>. Acesso em 05/04/2023.
- 23 Os dados do ano de 2014, fonte do gráfico da *The Economist*, estão disponíveis em <https://www.pornhub.com/insights/2014-year-in-review>. Acesso em 05/04/2023.
- 24 A frase lida em uma plataforma de rede social e registrada em meu diário de campo foi modificada, sem perda do sentido original, para evitar que a postagem e seu/sua autor/a sejam identificados/as.

- 25 Tenho sempre colocado no masculino e no feminino porque, de acordo com um dos levantamentos do Pornhub Insights, o Brasil é, em média, o sexto país no mundo (dados de 2022) em que mais mulheres acessam conteúdo pornográfico. Há uma predominância de homens, mas no Brasil cerca de 39% dos acessos são feitos por mulheres, o que de forma alguma é irrelevante.
- 26 Ainda que, em termos legais, o termo utilizado seja pornografia infantil, creio ser mais correto falarmos de abusos e violências. Pornografia, como tenho repetido, implica atos consensuais entre pessoas que podem consentir. Mantenho o termo pornografia infantil, no entanto, porque ele é o que aparece em todas os processos enfrentados pelo Pornhub.

Referências

- ATTWOOD, Feona (ed.). 2009. *Mainstreaming Sex*. London: IB Tauris.
- ATTWOOD, Feona (ed.). 2017. *Sex Media*. Cambridge: Polity Press.
- BUTLER, Judith. 1990. "The force of fantasy: feminism, Mapplethorpe, and discursive excess". *Differences*, v. 2, n. 2. Disponível em: <https://read.dukeupress.edu/differences/article-abstract/2/2/105/298383/The-Force-of-Fantasy-Feminism-Mapplethorpe-and?redirectedFrom=fulltext>. Acesso em 25/06/2024.
- BOYD, Danah; CRAWFORD, Kate. 2012. "Critical questions for big data". *Information, Communication & Society*, 15 (5). Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/1369118X.2012.678878>. Acesso em 25/06/2024
- BROWN, Wendy. 2019. *In the Ruins of Neoliberalism. The Rise of Antidemocratic Politics in the West*. New York: Columbia University Press.
- CAMINHAS, Lorena. 2022. "The politics of algorithmic rank systems in the Brazilian erotic webcam industry". *Porn Studies Journal*, v. 9:1-18.
- COELHO, Clara. 2020. *Pós-pornografia em foco: Um estudo sobre tensões políticas e usos do corpo*. Dissertação de Mestrado, IFCH, Unicamp, Campinas.
- CHRISTIAN, Brian. 2021. *The alignment problem. Machine learning and human values*. New York: Norton & Company.
- DIAZ-BENITEZ, Maria Elvira. 2010. *Nas redes do sexo*. Rio de Janeiro: Zahar.
- DIAZ-BENITEZ, Maria Elvira. 2012. Sexo com animais como prática extrema no pornô bizarro. *Cadernos Pagu*, n. 38. <https://periodicos.sbu.unicamp.br/ojs/index.php/cadpagu/article/view/8645039>. Acesso em 22/05/2024
- FERNANDES, Camila. 2017. *Figuras da Causação: Sexualidade Feminina, Reprodução e Acusações no Discurso Popular e nas Políticas de Estado*. Rio de Janeiro: Museu Nacional.
- GREGORI, Maria Filomena. 2016. *Prazeres Perigosos. Erotismo, gênero e limites da sexualidade*. São Paulo: Companhia das Letras.
- LEITE JR., Jorge. 2006. *Das maravilhas e prodígios sexuais: a pornografia bizarra como entretenimento*. São Paulo: Annablume.
- LOWENKRON, Laura. 2015. *O monstro contemporâneo: a construção social da pedofilia em múltiplos planos*. Rio de Janeiro: Eduerj.
- MACDONALD, Margaret. 2020. *PornHub and the Platformization of a culture industry*. Thesis in the Department of Communication Studies, Concordia University, Canadá.
- MANOVICH, Lev. 2011. Trending: the promises and the challenges of big social data. In: M. K. Gold (ed.), *Debates in the Digital Humanities*. Minneapolis, MN: The University of Minnesota Press. Disponível em http://www.manovich.net/DOCS/Manovich_trending_paper.pdf. Acesso em 12/03/2024
- MCNAIR, Brian. 2013. *Porno? Chic! How pornography changed the world and made it a better place*. New York: Routledge.

- MULHOLAND, Monique. 2013. *Young people and pornography. Negotiating pornification*. New York: Palgrave MacMillan.
- MUNHOZ, Sara. 2024. "Comparações, sínteses e extensões: a jurisprudência do Superior Tribunal de Justiça como metadados digitais". *Mana*, 30 (2).
- PAASONEN, Suzanna. 2011. *Carnal Resonance: affect and online pornography*. Cambridge: MIT Press.
- PARREIRAS, Carolina. 2008. *Sexualidades no pontocom: espaços e homossexualidades a partir de uma comunidade on-line*. Dissertação de Mestrado, PPGAS, Instituto de Filosofia e Ciências Humanas, Unicamp.
- PARREIRAS, Carolina. 2012. Altporn, corpos, categorias e cliques: notas etnográficas sobre pornografia on-line. *Cadernos Pagu*, v. 38:197-222. Disponível em <https://doi.org/10.1590/S0104-83332012000100007>. Acesso em 22/05/2024
- PARREIRAS, Carolina. 2013. "Mainstream, Alternative and Amateur: Brief Notes on Categories of Classification of Online Pornography". In: Jessica Pfeffer (org.), *Desire, Performance and Classification*. 1ed. Oxford: IDP.
- PARREIRAS, Carolina. 2015. *Altporn, corpos, categorias, espaços e redes: um estudo etnográfico sobre pornografia on-line*. Tese de Doutorado, PPGCS, Instituto de Filosofia e Ciências Humanas, Unicamp.
- PARREIRAS, Carolina. 2018. Pornografias. com: as convenções do altporn. *Antropolítica*, v. 42:16-40. Disponível em <https://doi.org/10.22409/antropolitica2017.1i42.a41883>. Acesso em 15/04/2024.
- PARREIRAS, Carolina. No prelo. "Menor and novinha: a reflection on the meanings, nominations and narratives of adolescence in Rio de Janeiro favelas". In: *Gender, Sexuality and Life Course: research and dialogues over contemporary social transformations*. New York: Vernon Press.
- PARREIRAS, C.; LINS, B. & FREITAS, E. 2020. "Estratégias para pensar o digital". *Cadernos de Campo (USP)*, São Paulo, v. 29:1-11. Disponível em <https://doi.org/10.11606/issn.2316-9133.v29i2pe181821>. Acesso em 15/05/2024.
- PARREIRAS, C. & MACEDO, R. 2020. "Digital inequalities in Brazil during the Covid-19 Pandemic: a brief reflection on the challenges of remote learning". *Digital Culture & Education*, v. 1. Disponível em <https://www.digitalcultureandeducation.com/reflections-on-covid19/digital-inequalities-and-education-in-brazil>. Acesso em 24/06/2024.
- PATTERSON, Zabet. 2004. "'Going online: consuming pornography in the digital era". In: Linda Williams (ed.), *Porn Studies*. Durham: Duke University Press.
- PERIN, Vanessa. 2024. "Coletando dados, produzindo ativos: sobre as similaridades criadas pela cooperação sul-sul brasileira em Moçambique e seus efeitos". *Mana*, 30 (2).
- RUBIN, Gayle. 1989. 'Thinking sex: notes for a radical theory of the politics of sexuality. In: Carole Vance (ed.), *Please and Danger*. London: Pandora.
- SAUNDERS, Rebecca. 2020. *Bodies of Work: The Labour of Sex in the Digital Age*. New York: Palgrave.

- SEEVER, Nick. 2022. *Computing Taste. Algorithms and the Makers of Music Recommendation*. Chicago: University of Chicago Press.
- SEEVER, Nick. 2017. "Algorithms as culture: Some tactics for the ethnography of algorithmic systems". *Big Data & Society*, v. 4, n. 2. Disponível em: <https://journals.sagepub.com/doi/10.1177/2053951717738104>. Acesso em 25/06/2024.
- SULLIVAN, Rebecca & MCKEE, Alan. 2015. *Pornography: Structures, Agency and Performance*. Cambridge: Polity Press.
- VANCE, Carole. 1989. *Pleasure and Danger*. London: Pandora.
- WALFORD, Tone. 2024. "Map, Territory, and Everything in Between: environmental data and territorialisation". *Mana*, 30 (2).
- WILLIAMS, Linda. 1999. *Hard Core. Power, pleasure, and the "Frenzy of the visible"*. Berkeley/Los Angeles/London: University of California Press.
- WILLIAMS, Linda. 2004. *Porn Studies*. Durham: Duke University Press.

REFLEXÕES SOBRE *BIG DATA*, SEXUALIDADES, DATIFICAÇÃO E MORALIDADES NO PORNÔ DIGITAL: A "NOVINHA" COMO CATEGORIA DE CLASSIFICAÇÃO E INDEXAÇÃO

Resumo

O objetivo deste artigo é refletir sobre a relação entre pornografia digital e big data, com foco especial nos modos como a categoria "novinha", utilizada como forma de classificação de conteúdos pornográficos, circula e é apropriada, transformando-se em dados e métricas. Esta categoria é parte do vocabulário do senso comum no Brasil, sendo bastante utilizada em músicas, em contextos periféricos e como forma de classificar um tipo específico de vivência da adolescência. A perspectiva aqui desenvolvida se baseia em observação etnográfica conduzida no Pornhub. Mais especificamente, interessam a esta reflexão as informações divulgadas pelo chamado Pornhub Insights, dedicado à publicação e à análise de dados de acesso. A partir disso, busco compreender como a divulgação destes dados — plataformizados e nada neutros — levanta questões relativas à interseção entre sexualidades, moralidades, violência e adolescência, além de funcionar como um índice do que seria parte do comportamento sexual brasileiro. Neste sentido, parto das ideias de "datificação sexual" (Saunders 2020) e de plataformização para entender os processos que transformam práticas sexuais, corpos e sexualidades em dados supostamente quantificáveis, estabelecendo formas de governança das sexualidades. Trago também observações de minha própria navegação pelo Pornhub, bem como recortes dos debates suscitados pelos dados de 2016 a partir de publicações no Facebook e no Twitter. Por fim, com o uso do Ncapture, trago um mapeamento experimental de como a categoria "novinha" circula pelo Twitter, estando associada a *hashtags* e postagens que corroboram sua apropriação para indicar conteúdos eróticos. Assim, busco mostrar como esta categoria e seu uso como demarcadora de conteúdos pornográficos coloca em questão os limites da sexualidade (Gregori 2016).

Palavras-chave: Pornografia; Big data; Moralidades; Datificação.

REFLECTIONS ON BIG DATA, SEXUALITIES, DATAFICATION AND MORALITIES IN DIGITAL PORN: THE "NOVINHA" AS A CLASSIFICATION AND INDEXING CATEGORY

Abstract

The purpose of this article is to reflect on the relationship between digital pornography and big data, with a special focus on the ways in which the "novinha" category, used as a way of classifying pornographic content, circulates and is appropriated, transforming itself into data and metrics. This category is part of the common sense vocabulary in Brazil, being widely used in songs, in peripheral contexts and as a way of classifying a specific type of adolescence experience. The perspective developed here is based on ethnographic observation conducted on Pornhub. More specifically, this reflection is interested in the information disclosed by Pornhub Insights, dedicated to the publication and analysis of access data. From this, I seek to understand how the dissemination of these data — platformized and far from neutral — raises questions related to the intersection between sexualities, moralities, violence and adolescence, in addition to functioning as an index of what would be part of Brazilian sexual behavior. In this sense, as proposed by Saunders (2020), I start from the idea of "sexual datafication" (Saunders 2020), which transforms sexual practices, bodies and sexualities into supposedly quantifiable data, establishing forms of governance of sexualities. In addition to the data from Pornhub Insights, I also bring observations from my own browsing through Pornhub, as well as excerpts from the debates raised by the 2016 data from publications on Facebook and Twitter. Finally, using Nvivo Capture, I bring an experimental mapping of how the "novinha" category circulates on Twitter, being associated with *hashtags* and posts that corroborate its appropriation to indicate erotic content. Thus, I seek to show how this category and its use as a demarcator of pornographic content calls into question the limits of sexuality (Gregori 2016).

Keywords: Pornography; Big data; Morality; Datafication.

REFLEXIONES SOBRE BIG DATA, SEXUALIDADES, DATIFICACIÓN Y MORALIDADES EN EL PORNO DIGITAL: LA "NOVINHA" COMO CATEGORÍA DE CLASIFICACIÓN E INDEXACIÓN

Resumen

El propósito de este artículo es reflexionar sobre la relación entre la pornografía digital y el big data, con especial atención a las formas en que la categoría "novinha", utilizada como forma de clasificar los contenidos pornográficos, circula y se apropia, transformándose en datos y métricas. Esta categoría forma parte del vocabulario de sentido común en Brasil, siendo ampliamente utilizada en canciones, en contextos periféricos y como forma de clasificar un tipo específico de experiencia adolescente. La perspectiva desarrollada aquí se basa en la observación etnográfica realizada en Pornhub. Más concretamente, esta reflexión se interesa por la información que divulgan los denominados Pornhub Insights, dedicados a la publicación y análisis de los datos de acceso. A partir de eso, busco comprender cómo la difusión de esos datos —plataformizados y lejos de ser neutrales— suscita cuestiones relacionadas con la intersección entre sexualidades, moralidades, violencia y adolescencia, además de funcionar como un índice de lo que sería parte de la sexualidad brasileña de comportamiento. En este sentido, como propone Saunders (2020), parto de la idea de "datificación sexual", que transforma las prácticas sexuales, los cuerpos y las sexualidades en datos supuestamente cuantificables, estableciendo formas de gobierno de las sexualidades. Además de los datos de Pornhub Insights, también traigo observaciones de mi propia navegación en Pornhub, así como extractos de los debates suscitados por los datos de 2016 de publicaciones en Facebook y Twitter. Finalmente, utilizando Nvivo Capture, realizo un mapeo experimental de cómo circula la categoría "novinha" en Twitter, asociándose a hashtags y publicaciones que corroboran su apropiación para indicar contenido erótico. Así, busco mostrar cómo esta categoría y su uso como delimitador de contenidos pornográficos cuestiona los límites de la sexualidad (Gregori 2016).

Palabras clave: Pornografía; Big data; Moralidade; Datificación.

Agradecimentos

Este artigo não seria possível sem o generoso apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp — processos n. 2021/06857-7 e 2022/13233-2). Agradeço ainda às pesquisadoras e aos pesquisadores do grupo de estudos e pesquisa que coordeno na USP com a professora Heloisa Buarque de Almeida pelos preciosos comentários na versão inicial deste texto.

Carolina Parreiras é antropóloga, doutora em Ciências Sociais (Unicamp). Atualmente, é pesquisadora do Departamento de Antropologia da Universidade de São Paulo (USP) e coordenadora do Laboratório Etnográfico de Estudos Tecnológicos e Digitais (LETEC), financiado pela Fapesp.

Editora-Chefe: María Elvira Díaz Benítez

Editor Adjunto: John Comeford

Editor Associado: Luiz Costa

Recebido em: 10 Abril 2023

Aceito em: 3 Junho 2024