



UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS HUMANAS, LETRAS E ARTES
COMUNICAÇÃO EM MÍDIAS DIGITAIS

NATALIA HAU FRANÇA PIRES

**ANÁLISE DAS FERRAMENTAS DE GERENCIAMENTO DE COMUNIDADE
DO BLUESKY E X NO COMBATE A CRIMES VIRTUAIS CONTRA AS
MULHERES**

João Pessoa

2024

NATÁLIA HAU FRANÇA PIRES

**ANÁLISE DAS FERRAMENTAS DE GERENCIAMENTO DE COMUNIDADE
DO BLUESKY E X NO COMBATE A CRIMES VIRTUAIS CONTRA AS
MULHERES**

Trabalho de conclusão de curso apresentado
ao Departamento de Mídias Digitais da
Universidade Federal da Paraíba, como parte
dos requisitos para a obtenção do título de
Bacharel em Comunicação em Mídias
Digitais.

Orientador: Prof. Dr. Cleber Matos de Moraes

João Pessoa

2024

Aos meus pais, Jorge e Raquel, que nunca mediram esforços para me apoiar em tudo o que eu me proponho a fazer.

AGRADECIMENTOS

Agradeço a Deus por sua graça concedida até aqui e minha família pelo amor que me deram, pelo consolo nos dias difíceis e pelo amparo em meio às nossas diferenças.

Ao meu pai, por cuidar de mim e me ensinar a nunca desistir de lutar por tudo o que mereço e por se orgulhar das minhas pequenas conquistas. Sem você, este trabalho não teria sido possível.

À minha mãe, por me mostrar que não existe problema que não possa ser superado; sua força ao longo dos anos fez de mim a mulher determinada que sou hoje. Sem a sua compreensão, eu não estaria aqui.

Aos meus irmãos, Geovane e Ester, que são a melhor parte de mim. Ver vocês sempre foi um sopro de alegria que ameniza a saudade de casa.

Ao meu orientador, Cleber Matos de Moraes, que embarcou nesta pesquisa comigo de forma inusitada e me guiou gentilmente até aqui. Você é o meu maior referencial de ensino.

Aos amigos que minha carreira me trouxe e que continuam a caminhar ao meu lado: Maria Isabelly, Lucas, Carlos, Maria Eduarda, Raquel e Lays. Sem vocês, meu dia a dia seria muito mais triste.

A Iury Alekxander, o amigo que a graduação me deu; sua presença foi a minha única constância durante os anos de trajetória acadêmica.

Aos amigos que me encontraram na caminhada: Shayenne, Davi Tito, Ana Luisa, Emmanuel, Dorivan e Bianca, Nicolý, Syanne, Aline, Gisele e Taisa. Caminhar com vocês é um privilégio que tenho a honra de desfrutar. E a todos aqueles que minha memória traiçoeira pode ter deixado de lembrar.

Tive a sorte de encontrar família e amparo em cada um de vocês. Sou quem sou hoje porque encontrei um lar e sei em quem posso confiar minhas alegrias e lágrimas.

“Toda vez que uma mulher se defende, sem
nem perceber que isso é possível, sem
qualquer pretensão, ela defende todas as
mulheres.” – Maya Angelou

RESUMO

Em um contexto onde as redes de relacionamento se estendem para o ambiente online de Redes Sociais, as interações antes propagadas fora da internet também sofrem a migração para o virtual. Essas interações podem ser comportamentos positivos (que reforçam laços de relacionamentos) e também comportamentos negativos, que podem gerar danos a outras pessoas. Tendo em vista a batalha pela igualdade de gênero que acontece em ambos os ambientes e as violências sofridas por mulheres, esta pesquisa visa averiguar a efetividade das ferramentas de gerenciamento de comunidades das redes sociais *Bluesky* e o *X*. O *Bluesky* é uma plataforma de microblogging de cunho federado e o *X* de funcionamento tradicional, centralizado no gestor. Cada uma dessas formas de gestão cria políticas para o combate a crimes contra mulheres no ambiente online. Esta pesquisa se propõe a estruturar uma análise comparativa dessas duas redes e suas ferramentas, avaliando seu potencial para impedir ou punir os crimes de *Doxxing*, *Revenge Porn*, Assédio e perseguição. A partir dos dados coletados, foram analisados os tipos de ferramentas de cada uma das Redes Sociais para verificar se as tais oferecem formas efetivas de combater ou punir tais crimes. Verificando, assim, quais das duas redes podem ser consideradas ambientes com mais recursos para que mulheres possam utilizar com segurança.

Palavras-chaves: Redes Sociais, Laços de relacionamento, *Bluesky*, *X*, Crimes contra mulheres

ABSTRACT

In a context where relationship networks extend into the online environment of social media platforms, interactions previously limited to offline settings now also migrate to virtual spaces. These interactions can manifest as positive behaviors, strengthening social bonds, or as negative behaviors that can harm others. Given the ongoing fight for gender equality across both environments and the forms of violence experienced by women, this study seeks to investigate the effectiveness of community management tools on the social networks Bluesky and X. Bluesky is a federated microblogging platform, while X operates with a traditional, centralized management structure. Each platform's management style implements distinct policies to combat online crimes against women. This research undertakes a comparative analysis of these two networks and their tools, evaluating their potential to prevent or penalize crimes such as doxxing, revenge porn, harassment, and stalking. Based on collected data, the study analyzes the types of tools available on each social network to determine whether they effectively address and penalize these crimes, thus identifying which of the two networks provides more resources to foster a safe environment for women.

Keywords: Social Networks, Relationship Bonds, Bluesky, X, Crimes Against Women

LISTA DE FIGURAS

Figura 1 - Tela inicial Bluesky	24
Figura 2 - Tela inicial Graysky	25
Figura 3 - Tela do filtro de palavras	32
Figura 4 - Quem pode responder?	33
Figura 5 - Filtro de notificações do X	34
Figura 6 - Mostrar com menos frequência do feed	34
Figura 7 - Formulário de denúncia de post	35
Figura 8 - Configuração de privacidade de conta em “Audiência e Marcação”	36
Figura 9 - Configuração de mensagens diretas no X	37
Figura 10 - Audiência e marcação	38
Figura 11 - Menu de quem pode responder?	38
Figura 12 - Menu de interações de post	39
Figura 13 - Menu de interações de post e usuários	40
Figura 14 - Feed inicial das comunidades	41
Figura 15 - Regras da comunidade do X	42
Figura 16 - Aviso de Spam.....	42
Figura 17 - Configuração de interação de post	47
Figura 18 - Menu de post - Silenciar Thread	48
Figura 19 - Adicionar palavras e tags silenciadas	48
Figura 20 - Removido pelo autor	49
Figura 21 - Denúncia de post	50
Figura 22 - Filtros de conteúdo	50
Figura 23 - O serviço de moderação do Bluesky	51

Figura 24 - Etiquetas de moderação	52
Figura 25 - Preferências de Thread	53
Figura 26 - Preferências de Thread	54
Figura 27 - Conta Silenciada	54
Figura 28 - Denunciar este usuário	55
Figura 29 - Lista de usuário	56
Figura 30 - Lista de moderação	57
Figura 31 - Pacote inicial	58
Figura 32 - Compartilhar Pacote inicial	58
Figura 33 - Compartilhar Pacote inicial	59
Figura 34 - Lista de rótulos	60

LISTA DE QUADROS

Quadro 1 - Análise das ferramentas de interação com conteúdo do X.....	65
Quadro 2 - Análise das ferramentas de interação com usuários do X	67
Quadro 3 - Análise das ferramentas nas comunidades do X	68
Quadro 4 - Análise das ferramentas de interação com conteúdo do Bluesky	70
Quadro 5 - Análise das ferramentas de interação com usuários do Bluesky	71
Quadro 6 - Análise das ferramentas dos feeds do Bluesky	73

LISTA DE GRÁFICOS

Gráfico 1 - Porcentagem de eficácia das ferramentas de interação conteúdo do X.....	75
Gráfico 2 - Porcentagem de eficácia das ferramentas de interação usuários do X	76
Gráfico 3 - Porcentagem de eficácia das ferramentas de comunidade do X	77
Gráfico 4 - Porcentagem de eficácia das ferramentas de interação de conteúdo do Bluesky	78
Gráfico 5 - Ferramentas de interação com usuários no Bluesky	79

LISTA DE TABELAS

Tabela 1 - Levantamento quantitativo de ferramentas e crimes X e Bluesky	74
---	----

SUMÁRIO

1. INTRODUÇÃO.....	14
2. JUSTIFICATIVA.....	18
3. FUNDAMENTAÇÃO TEÓRICA.....	21
3.1. Pesquisa em redes sociais.....	21
3.2 Microblogging.....	22
3.3 Redes sociais descentralizadas/federadas?.....	22
3.4 Crimes digitais contra mulheres.....	24
4. METODOLOGIA.....	26
4.1 Definição de recorte cronológico.....	27
4.2 Levantamento de ferramentas de gerenciamento de comunidades.....	27
4.3 Análise das ferramentas para notificação e combate a crimes.....	28
4.4 Análise comparativa de segurança.....	29
5. DESENVOLVIMENTO.....	29
5.1 As ferramentas de gerenciamento de comunidade do X.....	30
5.1.1 Ferramentas de interação com conteúdos.....	31
5.1.3 Comunidades do X.....	39
5.2 Termo de uso e diretrizes da comunidade do X.....	42
5.2.1 Termos de Uso do X.....	42
5.2.2 Diretrizes de comunidade no X.....	44
5.3 As ferramentas de gerenciamento de comunidade do Bluesky.....	45
5.3.1 Ferramentas de interação com conteúdos no Bluesky.....	45
5.3.2 Ferramentas de interação com o usuário.....	52
5.3.3 Feeds do Bluesky.....	59
5.4 Termos de uso e diretrizes da comunidade do Bluesky.....	60
5.4.1 Termos de uso do Bluesky.....	60
5.4.2 Diretrizes da comunidade do BlueSky.....	61
6. ANÁLISE.....	63

7. RESULTADOS E DISCUSSÕES.....	73
8. CONSIDERAÇÕES FINAIS.....	79
9. REFERÊNCIAS.....	81

1. INTRODUÇÃO

A partir do final do século XX, com o uso das tecnologias da comunicação, a sociabilidade intensificou uma nova esfera para atuação (CASTELLS,1999). Com o advento do virtual, uma parte da comunicação tradicional foi transferida para as mídias digitais em rede. Na década de 90 do século XX, essa estrutura de comunicação ainda era muito experimental, mas no início do século XXI houve uma grande aceleração do uso e relevância dessas mídias digitais.

Sherry Turkle (2011) afirma que as redes sociais são uma nova forma de comunicação que está transformando a maneira como nos conectamos uns com os outros. No entanto, também podem ter um impacto negativo na sociedade. Elas podem ser usadas para espalhar desinformação, promover o ódio e o assédio.

Não por acaso, fenômenos que acontecem no mundo real também foram transferidos para as redes sociais. Um deles é a violência contra mulheres. A violência virtual contra a mulher é um problema crescente, que pode ter consequências devastadoras para as vítimas. Conforme o estudo da ONU Mulheres:

[...] Em todo o mundo, 73% das mulheres que estão conectadas já foram expostas a algum tipo de violência online. As jovens entre 18 e 24 anos são as maiores vítimas de perseguições nas redes e de assédio sexual e muitas sofrem ameaças físicas (NewsUN, 2015).

De acordo com a pesquisa feita pelo SaferNet (2017), as mulheres são os principais alvos dos crimes virtuais. As redes sociais são mais um lugar onde violência virtual contra a mulher pode acontecer.

Essa violência pode acontecer em diversos tipos de interfaces comunicacionais, independentemente de suas estruturas. Redes sociais online, como Facebook, são ambientes em que esse evento de transferência da violência acontece de forma comum. A lógica de uma rede social como Facebook é de um laço mútuo, que presume afinidade. Diferentemente, outras redes sociais, como a rede de microblogging do X (antigo Twitter), não são uma rede de laços mútuos. Uma pessoa pode seguir a outra sem ser obrigado a seguir de volta. Apesar disso, pela dinâmica de articulação fácil,

nessas redes de microblogging também acontecem fenômenos de violência contra mulheres.

Lee Rainie e Barry Wellman (2012) defendem que as redes sociais online são um reflexo das antigas redes sociais formadas por diversas pessoas, agora, em grande escala. Segundo Rainie e Wellman (2012), as redes sociais se tornaram um conceito popular, que pode ser uma descrição apropriada das interações sociais, antes mesmo da revolução da internet. Como reflexo dessa mudança de ambiente, muitos hábitos que aconteciam fora das redes, passaram a ser reproduzidos no online.

Muitas das redes sociais mais recentes adotam algumas soluções e interações de outras mais conhecidas como uma forma de fazer com que a plataforma seja mais intuitiva, utilizando as bagagens trazidas pelas redes anteriores e seus acertos (D'Andrea, 2020). Temos como um bom exemplo o botão de coração, que é a forma de demonstrar interesse em uma publicação. Podemos ver esse botão até mesmo nas redes sociais mais emergentes. A primeira versão deste botão surgiu primeiramente no Facebook - plataforma da empresa Meta, no ano de 2007. Inicialmente um “curtir/joinha” era utilizado como símbolo de interesse rápido em uma postagem. Essa ação qualificava a interação em rede social, atribuindo relevância e pertencimento ao comportamento digital. Com o tempo, esse conceito de qualificação de conteúdo digital foi replicado em outras plataformas com diversas simbologias diferentes. O próprio Facebook atualizou para um conjunto de símbolos diversificados. Já outras redes, como Instagram, utilizaram a simbologia de “coração/amei”. Esse modo de qualificação social expandiu-se para diversas redes sociais, incluindo microblogs.

Cada rede social possui ferramentas que representam suas políticas internas. Como uma interface virtual para usuários, essas capacidades implementadas em cada ferramenta representam, num mundo virtual, a capacidade política dos atores nesse ambiente, reforçando a ideia de McLuhan (1964) de meio e mensagem.

Assim, ao permitir ou não certas ações, essas plataformas condicionam o agir nessa comunidade, uma teoria abordada por Dijck, Poell e Waal(2018).

[...]technology shapes society as well as the other way around. The ideal platform society, which governance builds on homogeneous generic public values, obviously does not exist. Public values need to be articulated in each sector and in each context because they are both stakes and outcomes in political and ideological¹ (Van Dijck, Poell e de Waal; 2018, p. 139)

¹ Tradução: a tecnologia molda a sociedade assim como a sociedade molda a tecnologia. A sociedade ideal de plataforma, na qual a governança se baseia em valores públicos genéricos homogêneos,

Seus usuários passam a fazer parte de sua sociedade virtual com uma regra que não necessariamente concordam ou podem mudar. Esse processo de mudança é uma relação legal, social e de relações públicas entre as plataformas e sua comunidade de usuários.

O X, antigo Twitter, foi uma rede social lançada em 2006 com a proposta de ser um microblog pessoal de mensagens curtas. Inaugurada e idealizada por Jack Dorsey, Biz Stone e Evan Williams, a plataforma se popularizou e continua ativa até hoje, sendo dirigida por Linda Yaccarino, atual CEO. A rede social tem mais de 10 anos de história, passou por diversas mudanças em suas ferramentas internas, algoritmo e a forma de gerenciamento, entendendo que as mudanças facilitariam o dia a dia de seus usuários.

O Bluesky é uma rede social que foi lançada em 2023, fundada por Jack Dorsey, cofundador e ex-CEO do X. Carregando a mesma proposta do X, a rede social visa ser um microblog pessoal de poucos caracteres, apresentando mudanças significativas em comparação com sua rede irmã. Uma delas é a ideia de que a empresa fundadora não precise gerenciar a sua comunidade, deixando esta missão para os usuários. Utilizando um novo algoritmo, o Bluesky acredita no autogerenciamento de comunidade através de uma rede social federada. Nas redes sociais federadas em vez de uma única empresa controlar a rede inteira, a federação permite a interoperabilidade entre diferentes serviços, criando um ecossistema interconectado. Esse sistema, busca romper com o controle centralizado e oferecer maior liberdade aos usuários.

Ao decorrer dos meses de funcionamento do Bluesky, uma comunidade de usuários brasileiros se formou. Com o princípio das federações, era esperado que as comunidades no Bluesky se autogerissem e conseguissem reduzir o problema da violência contra as mulheres na plataforma. Porém, com a adesão de novos membros, a ferramenta aparentou demonstrar problemas ao lidar com assédios nas suas comunidades. Essa insegurança cria questionamentos dos próprios usuários ao time de desenvolvedores: como uma comunidade que utiliza do autogerenciamento lidaria com esses casos? Quais são as ferramentas oferecidas pelo Bluesky para impedir que mais casos aconteçam?

obviamente não existe. Os valores públicos precisam ser articulados em cada setor e em cada contexto, pois são tanto apostas quanto resultados em disputas políticas e ideológicas.

Estabelecer regras, identificar condutas e avaliar usuários são formas de governança em redes sociais para garantir o bom funcionamento. Em geral, existem dois documentos de governança para as redes sociais, sendo eles o “Termos de serviço” e as “diretrizes para a comunidade”. Sobre o termo de serviço, Gillespie (2018, p.46) afirma:

[...] um contrato que explicita os termos sob os quais o usuário e a plataforma interagem, as obrigações que os usuários devem aceitar como condição para sua participação, e os meios adequados para resolver uma disputa que venha surgir.

Sobre as diretrizes de comunidade, consideradas escritas de forma informal, Gillespie (2018) destaca que o documento serve para especificar o que é apropriado ou não para aquela comunidade.

Ao navegar e utilizar as duas redes sociais, e presenciar a insatisfação dos usuários, é importante descobrir como as ferramentas oferecidas pelas plataformas citadas podem servir como ajuda para combater a violência contra a mulher no ambiente virtual. Por isso, ao olhar para o ocorrido na plataforma emergente, surge uma pergunta: em comparação com a rede social X, quais são as ferramentas de gerenciamento de comunidade do Bluesky que podem ser utilizadas para combater que os casos de crimes contra a mulher na plataforma se tornem comuns?

As plataformas lançam atualizações como forma de mitigar esses fenômenos através de ferramentas e soluções para trazer um ambiente mais seguro. Todavia, com a velocidade de mudança do digital, não parece que essas ferramentas atendam a escala da demanda social de proteção das mulheres em rede.

Este trabalho se propõe a investigar as ferramentas de gerenciamento de comunidade das redes sociais X e BlueSky e suas estratégias para reportar e impedir que crimes cibernéticos contra a mulher aconteçam ou sejam amplificados. Ambas as plataformas foram co-fundadas pela mesma pessoa, mas possuem modelos de negócio e sociabilidade diferentes. As políticas de cada uma de como lidar com essa questão mudam e isso se reflete nas suas ferramentas.

Este estudo se propõe a fazer uma análise comparativa das ferramentas de notificação e controle comunitário em duas redes sociais, identificando e classificando ferramentas de moderação e suas utilidades diante situações de assédio online, revenge porn, perseguição e Doxxing. O objetivo da pesquisa é coletar, analisar e catalogar as

ferramentas oferecidas pelas duas redes sociais, identificando quais são suas capacidades e seus papéis mediadores na comunicação digital e na segurança de comunidades. O resultado obtido é uma análise qualitativa comparativa das ferramentas e das políticas das plataformas analisadas em relação à violência contra as mulheres para melhor compreensão da sociabilidade online que destaca os pontos fortes e fracos de cada rede e trás luz sobre o assunto.

2. JUSTIFICATIVA

A violência virtual contra a mulher é um problema crescente, que pode ter consequências devastadoras para as vítimas. Os crimes cibernéticos, acontecem de diversas maneiras, podendo afetar ou não a vida das pessoas fora da internet, o JusBrasil, afirma que perante a lei, tudo o que é considerado crime fora das redes, também é considerado crime no ambiente cibernético.

Como usuária de diversas redes sociais durante muitos anos, através de um olhar detalhado pude identificar semelhanças entre ferramentas e soluções oferecidas por cada plataforma, algumas das quais são importantes para o gerenciamento de comunidades, ferramentas que dão opção de bloqueios, restringir e até mesmo publicar anonimamente. Ferramentas de moderação que muitas vezes são ignoradas pelo público pontual - público que não passa muitas horas online, se tornam grandes ativos no dia a dia de “heavy users” ou usuários fieis, sendo pontos geradores de confiança entre usuários ativos e plataformas.

Esse trabalho busca enriquecer o portfólio de pesquisa de Science and Technology Studies (Carlos d’Andrea, 2020), traduzida como Estudos de Ciência e Tecnologia, trazendo a perspectiva das ferramentas de gerenciamento de comunidade como ferramentas de combate a crimes contra as mulheres, tendo em vista denúncias que as redes X e Bluesky recebem sobre o assunto.

Como uma rede social federada, o Bluesky carrega o objetivo de que a plataforma se torne um ecossistema de mídias sociais descentralizado e que opere em código aberto, ou seja, não sendo propriedade de uma única empresa. Apesar de apresentar ferramentas que parecem promissoras como a aba “moderation” em português “moderação”, um local onde você supostamente tem o controle da micro comunidade que está ao seu redor, muitos usuários apresentaram denúncias sobre

acontecimentos indesejados na plataforma e a falta de ferramentas para combater. Tendo em mente que umas das redes sociais escolhida é descentralizada, a pesquisa levará em conta o funcionamento destas ferramentas em sua plataforma principal e sua federação, o Graysky.

O Graysky foi a primeira federação lançada através do Bluesky, trazendo a proposta do microblogging, mas com diferentes layout e ferramentas dentro de sua interface.

As plataformas X e Bluesky foram escolhidas por terem propostas semelhantes apesar de funcionamentos distintos. Assim, a comparação se torna importante na discussão da construção de gerenciamento de comunidades eficazes combatendo a violência contra mulher em ambientes virtuais.

O resultado da pesquisa tem o potencial de mostrar qual a rede social desempenha um papel melhor em suas ferramentas de gestão de comunidade para o combate da violência virtual contra a mulher, reforçando um dos Objetivos de Desenvolvimento Sustentável da Organização das Nações Unidas (ONU).

A Agenda 2030 para o desenvolvimento sustentável, estabelecida em 2015, representa um marco global para a promoção de um futuro mais equitativo e sustentável. Assinada por 193 líderes mundiais, ela é composta por 17 Objetivos de Desenvolvimento Sustentável (ODS), com o compromisso de alcançar três grandes metas até 2030: 1) erradicar a pobreza extrema; 2) combater a desigualdade e a injustiça; e 3) conter as mudanças climáticas.

No contexto desta pesquisa, o foco recai sobre o Objetivo de Desenvolvimento Sustentável número 5 (ODS 5), que visa promover a igualdade de gênero e empoderar todas as mulheres e meninas. Este objetivo reconhece que a equidade de gênero é fundamental não apenas para assegurar direitos humanos básicos, mas também para garantir o progresso social e econômico em escala global. A violência de gênero, a discriminação e as práticas prejudiciais contra mulheres e meninas continuam sendo barreiras significativas para o desenvolvimento sustentável.

O ODS 5 estabelece nove metas principais que abordam diversas dimensões da desigualdade de gênero, consideradas cruciais para a melhoria da qualidade de vida das mulheres e para a promoção de um desenvolvimento sustentável. As metas são as seguintes:

- Meta 5.1: "Acabar com todas as formas de discriminação contra todas as mulheres e meninas em todo o mundo" (Nações Unidas, 2015);
- Meta 5.2: "Eliminar todas as formas de violência contra mulheres e meninas, nas esferas públicas e privadas, incluindo o tráfico humano e a exploração sexual" (Nações Unidas, 2015);
- Meta 5.3: "Eliminar práticas nocivas, como casamentos infantis, precoces e forçados, e mutilações genitais femininas" (Nações Unidas, 2015);
- Meta 5.4: "Reconhecer e valorizar o trabalho de assistência e doméstico não remunerado, por meio de serviços públicos, infraestrutura adequada e políticas de proteção social, promovendo a responsabilidade compartilhada no lar" (Nações Unidas, 2015);
- Meta 5.5: "Garantir a participação plena e efetiva das mulheres em todos os níveis de tomada de decisão, seja na vida política, econômica ou pública" (Nações Unidas, 2015);
- Meta 5.6: "Assegurar o acesso universal à saúde sexual e reprodutiva e garantir os direitos reprodutivos, conforme acordos internacionais" (Nações Unidas, 2015);
- Meta 5.a: "Implementar reformas para assegurar às mulheres direitos iguais aos recursos econômicos, incluindo a propriedade de terras e outros bens, além do acesso a serviços financeiros e herança" (Nações Unidas, 2015);
- Meta 5.b: "Aumentar o uso de tecnologias de informação e comunicação para promover o empoderamento das mulheres" (Nações Unidas, 2015);
- Meta 5.c: "Adotar e fortalecer políticas públicas e legislações que promovam a igualdade de gênero em todos os âmbitos" (Nações Unidas, 2015).

Através da implementação dessas metas, busca-se não apenas promover uma sociedade mais justa, mas também gerar benefícios múltiplos para o desenvolvimento sustentável a um nível global. Este estudo, portanto, advoga pela importância da **ODS 5**, destacando a necessidade de estratégias robustas para combater a violência contra a mulher e as diversas formas de desigualdade de gênero que ainda persistem na sociedade através das redes sociais citadas acima.

Ao fazer uma análise comparativa entre as ferramentas de gerenciamento de comunidade para impedir que casos de crimes cibernéticos contra a mulher aconteçam em redes sociais de microblogging federadas. Esta pesquisa tem como questão principal

a investigação da possibilidade de que por sua característica federada, as redes sociais no Bluesky têm teoricamente maior capacidade de regular e instrumentalizar ferramentas para proteção da integridade e dos direitos das mulheres em redes sociais. Tendo como potencial, criar novas ferramentas para coibir crimes virtuais contra mulheres.

Sendo assim, os objetivos específicos que contribuíram para alcançar o objetivo geral desta pesquisa foram assim especificados: a) Entender como acontecem os crimes contra a mulher em redes sociais digitais. b) Fazer um levantamento das ferramentas de moderação de cada uma das plataformas escolhidas. c) Comparar a capacidade de cada ferramenta como uma forma de impedir que crimes aconteçam, avaliando como as ferramentas se aplicam nas diferentes comunidades federadas que estão localizadas dentro da plataforma Bluesky.

Esse projeto surgiu de acordo com a necessidade de avaliar as ferramentas de reporte em comunidades em duas das redes sociais de microblog, averiguando a confiança e usabilidade de suas ferramentas como forma de proteger os usuários de crimes cometidos contra mulheres.

3. FUNDAMENTAÇÃO TEÓRICA

Para que essa pesquisa seja bem fundamentada, precisaremos olhar para pesquisas que se aproximam do tema desejado, além de metodologias de pesquisas em redes sociais. Usaremos conceitos importantes de pesquisadores renomados nesta área, que por sua vez, é bastante recente, sendo eles: Tarleton Gillespie, Tarcízio Silva, Carlos d'Andrea e Manuel Castells, Barry Wellman e José Luis Orihuela.

3.1. Pesquisa em redes sociais

Em sua obra, Carlos d'Andrea (2020), nos apresenta os conceitos iniciais de redes sociais, plataformas, infraestrutura e governança. Tais assuntos são importantes para que entendamos como as plataformas de mídias sociais funcionam à medida que a internet 2.0 vai evoluindo. Ademais, d'Andrea (2020) afirma que “as plataformas são objetos empíricos em constante transformação, deixando uma sensação - injusta, mas compreensível de que a análise está realizada antes mesmo de ser finalizada”,

oferecendo grandes colaborações para lidar com esta realidade de redes sociais que estão sempre se modificando.

Manuel Castells (1999), argumenta que a sociedade contemporânea é organizada em torno da rede, uma estrutura social baseada na conexão e no fluxo de informação.

O autor defende que a rede é uma estrutura social baseada na conexão e no fluxo de informação. As redes são formadas por nós (indivíduos, organizações, etc.) conectados por laços (relações). A informação é o recurso mais valioso na sociedade contemporânea. Aqueles que controlam a informação têm o poder de moldar a sociedade.

De acordo com Lee Rainie e Barry Wellman (2012), as redes sociais já existiam antes mesmo do advento da internet. Elas poderiam ser caracterizadas por redes pequenas que poderiam incluir membros da família, amigos íntimos, vizinhos e pessoas conhecidas. Porém, Wellman e Rainie atribuíram à conexão com a tecnologia e internet, o papel de uma mudança revolucionária.

“Such networks had already formed before the coming of the internet. Still, the revolutionary social change from small groups to broader personal networks has been powerfully advanced by the widespread use of the internet and mobile phones” (Rainie e Wellman, 2012, p.8).²

3.2 Microblogging

Uma ferramenta de microblogging consiste numa fusão de blog, rede social e mensagens instantâneas (ORIHUELA, 2007). Os microblogs fazem parte do ecossistema de redes sociais e tem como objetivo serem redes sociais de poucos caracteres, trazendo mais objetividade e facilidade de informação e conexão. Atualmente, existem diversas plataformas de microblog em funcionamento, sendo algumas das mais famosas o X (antigo Twitter), Threads e Bluesky.

3.3 Redes sociais descentralizadas/federadas?

² Tradução: “Tais redes já haviam se formado antes da chegada da internet. Ainda assim, a mudança social revolucionária de grupos pequenos para redes pessoais mais amplas foi poderosamente impulsionada pelo uso generalizado da internet e dos telefones celulares.

Tarleton Gilliespie (2018, p.6) defende que “ plataformas devem ser, de alguma forma, moderadas: para proteger um usuário do outro, ou um grupo do outro e remover conteúdos ofensivos, vis ou ilegais”, enfrentando assim as propostas de plataformas descentralizadas, que ele considera ser uma utopia (p.28).

Além disso, o autor possui em sua publicação (2018), boas análises voltadas às moderações de plataformas e gerenciamento de comunidade, usando como exemplos redes sociais famosas pelo autogerenciamento de suas comunidades, sendo elas: Reddit, Flickr e Tumblr. O autor também aborda algumas das temáticas que serão trabalhadas durante a pesquisa, sendo elas Revenge Porn e assédio.

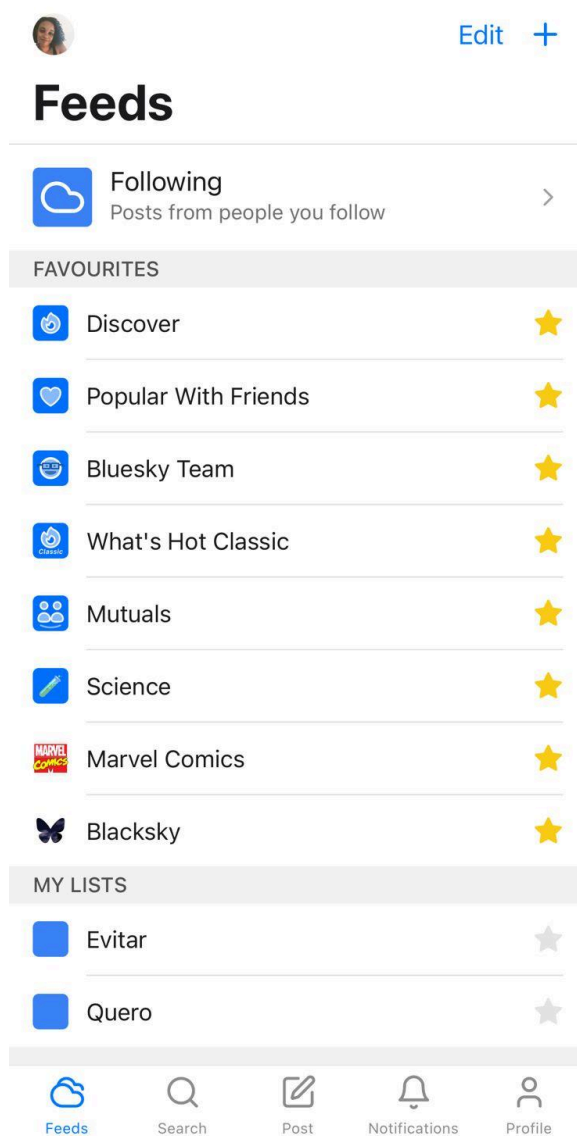
O Bluesky é uma rede federada recente, possuindo somente um federação lançada em 2024. O Graysky e o Bluesky, apesar de serem em sua essência a mesma rede social, as duas plataformas apresentam mudanças significativas em sua aparência principalmente em sua home page, sendo o Bluesky formado por uma tela de navegação em formato de linha do tempo (Figura 1) e o Graysky navegação através de um menu(Figura 2).

Figura 1 - Tela inicial Bluesky



Fonte: Bluesky

Figura 2 - Tela inicial Graysky



Fonte: Graysky

3.4 Crimes digitais contra mulheres

Sabendo que os fenômenos que acontecem fora das redes podem acontecer também nos ambientes virtuais, os crimes também podem mudar de ambiente. Por isso, é importante entender que crimes e tipos de violência contra mulher devem ser avaliados, investigados e adequados em suas leis no ambiente virtual.

O Direito deve-se adequar à nova realidade, sob pena de perder seu verdadeiro papel, qual seja disciplinar as relações sociais e impor normas de conduta. Assim, o binômio Direito e Internet não constitui fenômeno passageiro. Trata-se de uma realidade ainda pouco explorada, mas que deve ser analisada sob todos os campos das ciências jurídicas, a fim de garantir novos direitos fundamentais, bem como a efetivação dos já existentes (FIORILLO; CONTE, 2016, p.17).

Apesar disso, não é de conhecimento comum que algumas das coisas que são caracterizadas como crime na internet, também sejam crimes em ambientes virtuais. Durante esta pesquisa, quatro tipos de violência contra mulher serão abordados e caracterizados como crime seguindo os princípios de Fiorillo e Conte (2016): *Doxxing*, *Revenge porn*, assédio online e perseguição online.

Doxxing é o ato de expor informações pessoais de alguém na internet sem o devido consentimento. O termo deriva da frase "dropping dox"³ (JusBrasil, 2020). Nesses casos, reitera-se que é possível que a prática seja resultado de uma aquisição ilícita de informações, por violação a dispositivo informático previsto no referido art. 154-A do Código Penal (Brasil, 1940).

Revenge porn, considerado por lei um crime, consiste na divulgação de fotos e vídeos íntimos de ex-parceiros como forma de vingança. O ato pode ser enquadrado como difamação (art. 139) ou injúria (art. 140), considerados crimes contra a honra pelo Código Penal Brasileiro. O Artigo indica "Divulgação de cena de estupro ou de cena de estupro de vulnerável, de cena de sexo ou de pornografia" (Brasil, 2018).

Assédio online é o uso de tecnologias de informação e comunicação para apoiar comportamentos hostis, repetidos e deliberados. O assédio virtual pode incluir: Comentários sexuais, Divulgação de informações pessoais, Disseminação de discursos de ódio e/ou comportamentos hostis repetidos (JusBrasil, 2016).

Perseguição online, também conhecida como "*stalking*"⁴, é um crime que consiste em perseguir alguém repetidamente, inclusive de forma virtual. A perseguição virtual pode ocorrer, por exemplo, com a perturbação constante de acordo com a Lei 14.132/21 (Brasil, 2021).

Com base nos autores e publicações selecionados para fundamentar esta pesquisa, serão analisados também os documentos oficiais de cada plataforma. No que

³ Tradução: liberar documentação.

⁴ Tradução: perseguição.

diz respeito aos crimes, utilizaremos as definições previstas pela legislação brasileira para tratar comportamentos problemáticos como crimes.

4. METODOLOGIA

Este trabalho se propõe a analisar de forma comparativa as ferramentas de comunidade que ajudam a coibir e eliminar crimes contra a mulher, sendo eles: assédio online, revenge porn, perseguição e Doxxing.

A metodologia de análise comparativa qualitativa surgiu em 2001 Tendo como base os desafios das pesquisas em redes sociais apontadas por Carlos d'Andrea (2020). Para tal, se faz necessário realizar um recorte temporal para que a pesquisa não se torne obsoleta à medida que aconteça.

A pesquisa foi realizada com o acesso às redes sociais X, Bluesky em dias e horários alternados, com foco no *feed* geral, em comunidades e em *feeds* específicos, visando à devida comparação. Além disso, foi realizada a leitura detalhada dos manuais e termos de uso de cada plataforma, verificando as instruções oficiais de cada empresa no que se refere a crimes, assédios e comportamentos abusivos. A pesquisa foi estruturada em quatro partes para facilitar a coleta de dados:

1. Definição do recorte cronológico: realização das avaliações das ferramentas disponíveis durante um intervalo de tempo específico.
2. Levantamento de ferramentas de gerenciamento de comunidades: Identificação das ferramentas de gerenciamento de comunidades disponíveis em cada rede, considerando sua utilidade específica para prevenir ou combater crimes contra a mulher.
3. Análise das ferramentas para notificação e combate a crimes: Investigação de como as ferramentas disponíveis em cada rede social permitem que os crimes listados sejam notificados, combatidos, prevenidos ou punidos.
4. Análise comparativa: Comparação dos resultados para identificar qual das redes sociais oferece um ambiente mais seguro para o combate a ataques contra mulheres no espaço virtual.

4.1 Definição de recorte cronológico

Para garantir a atualidade da pesquisa e em conformidade com os princípios metodológicos de análise em redes sociais propostos por Carlos D’Andrea (2020), definiu-se um recorte temporal com início em 1º de agosto de 2024 e término em 3 de setembro do mesmo ano. Dados coletados fora desse período foram desconsiderados, visando a confiabilidade da pesquisa. As informações foram obtidas exclusivamente durante o período delimitado, por meio de capturas de tela (prints) e anotações feitas pela pesquisadora ao final da coleta, sendo posteriormente avaliadas e comparadas.

A análise da plataforma X apresentou uma particularidade devido aos bloqueios recentes realizados pelo Ministério da Justiça, em função do descumprimento de leis e ordens judiciais brasileiras. Esse fato tornou a plataforma indisponível temporariamente em todo o território nacional. Para contornar essa limitação, utilizou-se o repositório online Internet Archive, também conhecido como Wayback Machine, um site que arquiva páginas da internet desde 1996. A coleta foi realizada com base nos registros salvos até 3 de setembro, permitindo a análise dos dados pertinentes.

4.2 Levantamento de ferramentas de gerenciamento de comunidades

O segundo passo da pesquisa consistiu na análise das ferramentas de gerenciamento de comunidade das redes sociais Bluesky, X e Graysky. Foram utilizados dois métodos para listar os recursos disponíveis que poderiam ser utilizados no combate a crimes virtuais contra a mulher:

1. Uso Prático da Plataforma: Inicialmente, foi realizada a exploração da plataforma sem a leitura prévia dos Termos de Serviço, com o intuito de evitar que as diretrizes formais influenciassem a percepção inicial do funcionamento das ferramentas.;
2. Análise dos Termos de Serviço e Central de Ajuda: Em um segundo momento, os Termos de Serviço e os documentos auxiliares foram analisados para comparar a experiência prática com as normas e políticas formais estabelecidas.

O primeiro método envolveu a navegação e usabilidade das plataformas de forma individual e direcionada às comunidades, destacando as ferramentas disponíveis tanto no *feed* aberto e na página inicial, quanto nos *feeds* fechados das comunidades.

Esse processo incluiu a análise das diretrizes das comunidades e das ferramentas oferecidas internamente por cada uma delas.

Para a parte da pesquisa relacionada às comunidades, foram selecionadas duas comunidades de um mesmo nicho em cada plataforma. O objetivo é verificar como as ferramentas de gerenciamento se comportam conforme o nicho de atuação. As comunidades escolhidas foram de temas científicos e cultura pop voltada à marca Marvel Comics. A comunidade científica foi escolhida porque, no contexto dos últimos anos de pós-verdade, esses agrupamentos tentam se proteger em redes sociais com curadoria e crítica às interações de forma mais rigorosa. Dessa forma, essas comunidades são mais controladas e atentas a comportamentos não condizente com suas políticas. Já as comunidades de cultura pop incorporam o comportamento médio dos usuários de internet. Misturando o interesse comercial de marcas com cultura e consumo, essas comunidades dão apoio a comportamentos não tão supervisionados. Nessas comunidades serão avaliados os termos internos das comunidades, comportamentos, procedimentos e ferramentas disponíveis em cada uma delas, tendo em mente suas particularidades.

A segunda investigação é focada na leitura dos documentos de governança dessas redes. No caso do X, a investigação se concentra no documento de Termo de Uso e as Diretrizes de comunidade. Já no Bluesky, a documentação surge através do seu blog, visto que eles ainda não apresentam arquivos estruturados para a governança de suas plataformas. Serão utilizadas as publicações e comunicação oficial da equipe da Rede Social com o público para traçarmos as regras de governança adotadas por eles.

4.3 Análise das ferramentas para notificação e combate a crimes

Após identificar as ferramentas existentes, a pesquisa prosseguirá com a análise dos documentos de governança de cada rede social, como os termos de uso e as diretrizes de comunidades fornecidas por cada plataforma. Essa etapa visa verificar as instruções e orientações oferecidas pelas distintas redes sociais no que se refere ao tratamento de crimes virtuais contra mulheres e comportamentos abusivos cometidos por usuários.

Além disso, as ferramentas previamente identificadas serão examinadas quanto às suas formas de utilização. O objetivo é avaliar quais dessas ferramentas são eficazes

para coibir ou inibir crimes como revenge porn, doxxing, assédio online e perseguição virtual. Será realizada uma descrição detalhada de como cada ferramenta funciona e sua aplicabilidade para prevenir, proteger e denunciar tais condutas nocivas nas redes sociais. Essa análise crítica busca investigar a adequação de cada recurso como mecanismo de proteção e denúncia, considerando a eficácia de sua implementação no combate a esses tipos de crimes virtuais.

4.4 Análise comparativa de segurança

Neste estágio da pesquisa, será realizada a comparação dos resultados das ferramentas disponíveis em cada rede social e seu potencial para combater os crimes mencionados anteriormente. Para tanto, cada ferramenta identificada será categorizada em três níveis: denunciar, coibir e impedir que o crime ocorra novamente.

Ao final desse processo, espera-se expor nos resultados uma direção para pergunta central da pesquisa, indicando qual rede social, de acordo com os critérios utilizados, apresenta a maior capacidade de oferecer um ambiente seguro em seus diversos aspectos para as mulheres. Dessa forma, espera-se poder avaliar qual plataforma disponibiliza ferramentas de gerenciamento de comunidades que sejam, ao mesmo tempo, eficazes e efetivas no combate a crimes virtuais, garantindo a segurança de suas usuárias.

5. DESENVOLVIMENTO

Tomando como base os conceitos explicitados na fundamentação teórica e seguindo os procedimentos metodológicos descritos no capítulo anterior, a etapa de execução deste trabalho foi iniciada, se baseando na janela cronológicas de 1 de agosto até 3 de setembro, para a coleta de dados de usabilidade das ferramentas de gerenciamento de todas as plataformas citadas anteriormente nesta monografia. O objetivo central desta pesquisa é realizar uma análise comparativa das ferramentas de gerenciamento de comunidades em redes sociais de microblogging federadas ou não, com foco na prevenção de crimes cibernéticos contra mulheres. Através dessa avaliação poderá ser elucidado o potencial ou não dessas plataformas para desenvolver novas soluções voltadas ao combate de crimes virtuais contra mulheres.

Este capítulo apresentará em dois tópicos para cada rede social: (1) ferramentas de gerenciamento de comunidade e; (2) Termos de uso e diretrizes. A investigação exploratória das ferramentas que cada rede oferece será registrada no item 1. E os dados de uso estabelecidos juridicamente ou pela regulação comunitária serão extraídos das documentações. Para fins de imparcialidade, as documentações disponíveis por cada plataforma só foram acessadas após o fim do período determinado para uso na primeira etapa de exploração.

De acordo com Gillespie (2018), grande parte das empresas e redes sociais disponibilizam em seus ecossistemas documentos de governança, sendo os dois principais o "Termo de Uso" e as "Diretrizes de Comunidade". Esses documentos têm a função de especificar o que é considerado apropriado ou não para a comunidade da plataforma. Durante a presente pesquisa, após a coleta de dados por meio da análise de usabilidade, consultaram-se os documentos de governança das redes sociais estudadas, com o objetivo de compreender as diretrizes oficiais de cada plataforma no que diz respeito ao tratamento de denúncias de crimes contra a mulher e à veiculação de conteúdo abusivo.

5.1 As ferramentas de gerenciamento de comunidade do X

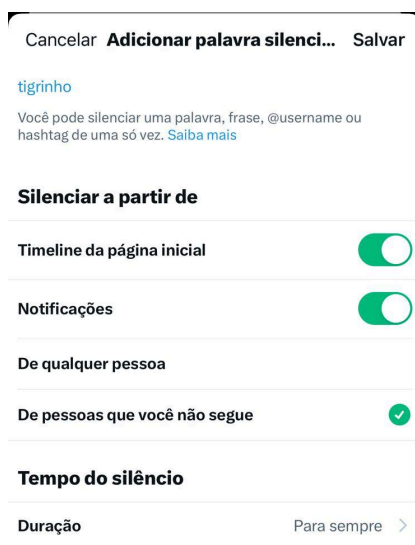
O X, antigo Twitter, é uma rede social de microblogging fundada em 2009 por Jack Dorsey, Biz Stone, Noah Glass e Evan Williams. A rede que inicialmente permitia apenas 145 caracteres, passou por mudanças significativas ao longo dos anos. Tais atualizações tiveram impacto na usabilidade de suas redes e do gerenciamento de comunidades. Em 2024 suas ferramentas continuam se atualizando. No entanto, com a mudança de administração algumas ferramentas foram retiradas, apesar da tela principal permanecer a mesma. As ferramentas atuais incluem ações de engajamento, contato com usuários e ações dentro de comunidades específicas. As ferramentas foram organizadas primeiramente a fim de explicar sua funcionalidade geral e após isso, ao elencarmos os resultados, será traçado um paralelo com os crimes contra mulheres citados anteriormente nesta monografia.

5.1.1 Ferramentas de interação com conteúdos

Ao utilizar o X, é possível verificar as ferramentas e listá-las de acordo com alguns critérios. Neste primeiro momento foram documentadas as ferramentas de gerenciamento de comunidade que dizem respeito ao conteúdo produzido e consumido dentro do *feed* e não necessariamente ao seu autor.

A primeira ferramenta analisada é o Filtro de palavras (Figura 3). O filtro de palavras é uma ferramenta de gerenciamento de conteúdo no *feed* que, através do filtro de palavras específicas, um usuário pode silenciar publicações e imagens que você não deseja receber. As palavras silenciadas são sensíveis a alteração de letras maiúsculas e minúsculas. Ou seja, cada palavra que se deseja que não seja mostrada na timeline devem ser silenciadas das duas formas. Além disso, a ferramenta aceita que você silencie palavras em todos os idiomas. Esta definição é encontrada dentro da aba definições de privacidade. Esta funcionalidade só permite que cada conta bloqueie até 200 palavras.

Figura 3 - Tela do filtro de palavras

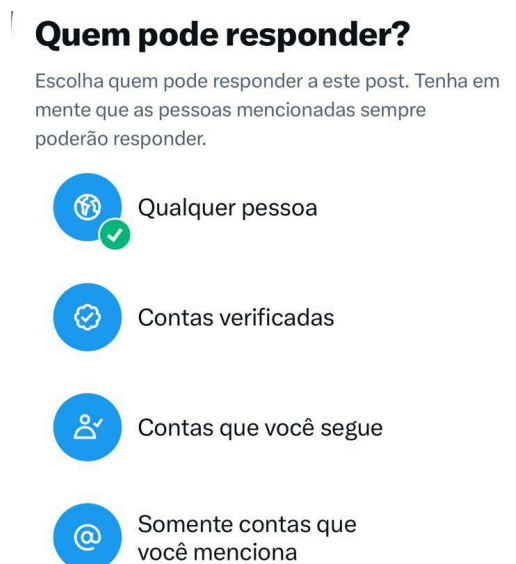


Fonte: *X silenciador de palavras* (setembro 2024)

A ferramenta “Quem pode responder?” , disponibilizada pelo X, permite que através dela o usuário possa decidir quem vai interagir com o seu conteúdo. Ao fazer uma publicação, o usuário clica no botão no canto inferior esquerdo e decide quem pode

responder, se “qualquer pessoa”, “contas verificadas”, “contas que você segue”, e “somente contas que você menciona” (Figura 4).

Figura 4 - Quem pode responder?



Fonte: X “Quem pode responder”

Outra ferramenta elencada é o Filtro de Notificações (Figura 5). Através dele o usuário tem a opção ativar um filtro onde você pode escolher que tipo de notificação você deseja receber em sua aplicação. As opções disponíveis são: comentários, curtidas, retweets, mensagens e novos posts de pessoas seguidas pelo usuário. Para ativar o filtro é necessário clicar no botão de configurações e selecionar a opção “notificação”. Neste setor da plataforma você terá a opção de selecionar a opção “filtro de qualidade”. Esse filtro, como explicado no próprio botão de selecionar, filtra conteúdos que o próprio X considera conteúdos de baixa qualidade. Logo abaixo existe a opção de filtros avançados, onde você pode encontrar mais opções para personalizar a sua experiência com as notificações nas redes.

Figura 5 - Filtro de notificações do X



Fonte: X

A quarta ferramenta é chamada de Mostrar com menos frequência (Figura 6). Através dessa ferramenta o usuário tem a possibilidade de marcar um post para que ele seja mostrado com menos frequência, assim personalizando e ensinando o algoritmo quais são as publicações do seu agrado dentro da plataforma. Para acessar essa funcionalidade essas informações, no post que o usuário não deseja mais receber, basta clicar nos três botões ao lado, o usuário pode clicar na primeira opção “Esse post não me interessa”.

Figura 6 - Mostrar com menos frequência do feed



Fonte: X

A quinta ferramenta é o Denunciar post (Figura 7). ao clicar no post que o usuário deseja denunciar, o menu oferece a opção de denunciar uma publicação, após isso, um formulário de denúncias se abrirá, oferecendo as opções com motivos para que

o post seja denunciado, sendo eles “ódio”, “Abuso e assédio”, “Discurso Violento”, “segurança infantil” “ Privacidade”, “Spam”, “Suicídio ou automutilação”, “Falsa Identidade”.

Figura 7 - Formulário de denúncia de post

Reunindo informações

Que tipo de problema você está denunciando?

[Por que estamos perguntando isso?](#)

Ódio ☐

Difamação, estereótipos racistas ou sexistas, desumanização, incitação de medo ou discriminação, referências a propagação de ódio, símbolos de ódio e logotipos

Abuso e assédio ☐

Insultos, conteúdo sexual indesejado e objetificação gráfica, conteúdo indesejado NSFW e gráfico, negação de eventos violentos, assédio direcionado e incitação ao assédio

Discurso violento ☐

Ameaças violentas, desejo de danos, apologia à violência, incitação da violência, incitação codificada à violência

Próximo

Fonte: X

As ferramentas de interação direta com os usuários não dizem respeito ao conteúdo postado, mas sim ao autor do conteúdo e a forma que as contas se relacionam uma com as outras. Nesta divisão, serão apresentadas ferramentas que limitam o acesso livre de usuários indesejados ao perfil pessoal de outras pessoas que utilizam a plataforma.

Já nas configurações de perfil o usuário tem a opção de manter a conta dele privada. Assim é possível o controle dos seguidores e das mensagens que os usuários vão receber. Para além da rede próxima, existe uma limitação para pessoas não seguidoras de não verem a interação desses usuários com postagens. Ou seja, ao comentar em um conteúdo, somente seus seguidores poderão visualizar esta interação.

Figura 8 - Configuração de privacidade de conta em “Audiência e Marcação”



Fonte: X

Já em Mensagens Diretas (Figura 9), também conhecidas como DMs, existem personalizações e formas de bloquear o acesso de outros usuários às mensagens diretas. Na aba de mensagens ao clicar no ícone da engrenagem, é possível acessar a área de ajustes. As primeiras possibilidades para os usuários são as opções para permitir ou não solicitações de mensagens. As opções que aparecem na tela incluem: “ninguém”, “usuários verificados” ou “todos”.

Figura 9 - Configuração de mensagens diretas no X

← **Configurações das mensag...** Concluído

Permitir solicitações de mensagem de:

As pessoas que você segue sempre poderão enviar mensagem para você. [Saiba mais](#)

- Ninguém** ☒
- Usuários verificados** ☐
- Todos** ☐

Habilitar chamadas de áudio e vídeo



Leve as mensagens a outro patamar com chamadas de áudio e vídeo. Quando ativado, você pode selecionar com quem você se sente confortável em usar. [Saiba mais](#)

Permitir chamadas de áudio e vídeo de:

Para reduzir chamadas indesejadas, você precisará ter recebido uma mensagem da conta pelo menos uma vez antes de a pessoa poder chamar você.

- Pessoas no seu catálogo de endereços** ☒
- Pessoas que você segue** ☒
- Usuários verificados** ☐

Fonte: X

Sobre as publicações, existem formas de definir se os post públicos são posts que todas as pessoas podem visualizar (independente se a conta do autor não é uma conta privada), então não existem restrições sobre o alcance da postagem ou para quem ela vai chegar.

Na utilidade Posts protegidos somente os seguidores do usuário poderão visualizar e interagir com os posts que têm essa marcação.

Figura 10 - Audiência e marcação



Fonte: X

Mais uma ferramenta é Configurações de interação de posts. Através desta ferramenta ao fazer uma publicação, o usuário pode clicar nos três pontinhos de sua própria publicação e consegue configurar de forma mais personalizada as regras de interação das pessoas com aquele post através do menu “Quem pode responder?” (Figura 11).

Figura 11 - Menu quem pode responder?



Fonte: X

Outra ferramenta é o Silenciar (Figura 12), através dela é possível escolher silenciar contas específicas, evitando que nenhum dos posts do usuário limitado apareça. Mas vale a pena ressaltar que ainda existe a possibilidade desta conta interagir com o usuário que a restringiu.

O botão Bloquear é uma ferramenta que permite bloquear uma conta, (Figura 12) é possível restringir a capacidade desta conta interagir com a sua conta, seja através de mensagens, curtidas e retweets. Silenciar é diferente de bloquear ou deixar de seguir: contas silenciadas não têm como saber que você as silenciou.

Figura 12- Menu de interações de post e usuários bloquear e silenciar



Fonte: X

O x também oferece a ferramenta Denúncia. Durante o uso e acesso a publicações é possível denunciar caso aconteça alguma interação indesejada com outro usuário. Seja através de posts ofensivos direcionados a uma única pessoa ou comunidade, mensagens privadas ofensivas ou com conteúdo inadequado e perfis que propagam conteúdos abusivos. É possível realizar uma denúncia abrindo um processo através de um formulário que é disponibilizado para a transformar, anexando prints e explicando os acontecimentos.

Figura 13 - Menu de interações de post Denunciar Post



Fonte: X

5.1.3 Comunidades do X

As comunidades no X constituem-se em grupos temáticos criados com o objetivo de permitir que usuários com interesses comuns interajam em um espaço mais direcionado e moderado dentro da plataforma. Tais comunidades oferecem um ambiente propício para a discussão de temas específicos, onde os membros podem compartilhar conteúdos relacionados e participar de conversas de maneira mais organizada e segmentada em comparação com o *feed* geral do X. Dessa forma, as comunidades proporcionam um espaço voltado à concentração de interesses específicos, favorecendo a interação entre os usuários que compartilham afinidades temáticas, ao mesmo tempo em que mantêm um controle mais rigoroso sobre a moderação e as regras internas de cada grupo.

Ao analisar a forma como as comunidades se estruturam para proporcionar uma experiência personalizada aos usuários, a presente pesquisadora inseriu-se em duas comunidades que foram selecionadas como base para esta porção do estudo: uma voltada à temática de cultura pop, intitulada "Marvel Comics", e outra relacionada à área científica chamada "Ciência". Essas comunidades foram escolhidas para permitir uma análise comparativa da interação dos participantes em diferentes contextos temáticos.

Figura 14 - Feed inicial das comunidades



Fonte: X

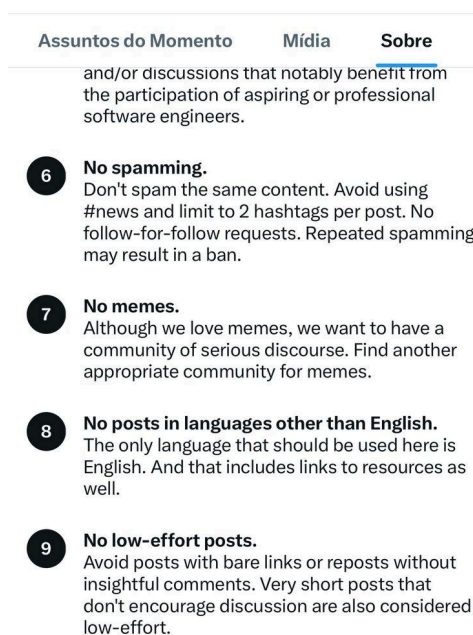
A Comunidade voltada à cultura pop foi escolhida pela diversidade de pessoas participando, tanto quando falamos em idade quanto localidades e contexto mais popular, Já a voltada a ciência inclui pesquisadores e pessoas que fazem parte de um contexto acadêmico, tendo regras bem rígidas quanto a divulgação de informações não verdadeiras e interação desrespeitosa com usuários.

Os dados coletados estão ligados com a percepção do uso dessas comunidades, suas regras internas e as ferramentas oferecidas para os moderadores, para manter a comunidade organizada sem a disseminação de conteúdos que violem suas regras.

Nas primeiras observações, foi possível verificar que todas as comunidades estão visíveis ao público, ou seja, o usuário consegue pesquisar pela comunidade de seu interesse e encontrá-la facilmente. Porém a participação na comunidade acontece via aprovação dos moderadores após a solicitação de entrada. Somente membros das comunidades podem postar e interagir com a comunidade.

A primeira ferramenta disponibilizada pelo X para as comunidades são as Regras de comunidade (Figura 15), onde cada comunidade pode se organizar de forma diferente, por isso, ao solicitar a entrada dentro de uma comunidade, uma aba de diretrizes se abrem. Nesta aba existem variações de regras e cada uma delas pode ser personalizada de acordo com o desejo dos moderadores.

Figura 15 -Regras da comunidade do X



Fonte: X

A segunda ferramenta é o Filtro de Spam para moderadores para manter o controle da comunidade. Nela os moderadores podem colocar filtros pontuando publicações que podem ser consideradas spam.

Figura 16 - Aviso de Spam



Fonte: X

Outra ferramenta é a Palavras chaves. Através das palavras chaves, os moderadores podem marcar palavras indesejadas para serem discutidas na comunidade. Assim, quando essas palavras forem utilizadas, os moderadores serão informados automaticamente.

Dentro de uma comunidade, os membros que desrespeitarem as normas da comunidade podem ser expulsos pelos administradores das comunidades utilizando a ferramenta “remover membros”. Após a remoção, o membro removido receberá uma notificação com o aviso da remoção.

Os moderadores têm a capacidade de aplicar as regras da comunidade, revisar conteúdos sinalizados, e tomar medidas contra comportamentos inadequados, como emitir avisos ou banir usuários problemáticos os próprios administradores ou pessoas da comunidade.

5.2 Termo de uso e diretrizes da comunidade do X

De acordo com Gillespie (2018), grande parte das empresas e redes sociais disponibilizam em seus ecossistemas documentos de governança, sendo os dois principais o "Termo de Uso" e as "Diretrizes de Comunidade". Esses documentos têm a função de especificar o que é considerado apropriado ou não para a comunidade da plataforma. Durante a presente pesquisa, após a coleta de dados por meio da análise de usabilidade, consultaram-se os documentos de governança das redes sociais estudadas, com o objetivo de compreender as diretrizes oficiais de cada plataforma no que diz respeito ao tratamento de denúncias de crimes contra a mulher e à veiculação de conteúdo abusivo.

5.2.1 Termos de Uso do X

Os Termos de uso regulam o acesso e uso dos serviços oferecidos pela X, os quais incluem uma variedade de sites, mensagens via SMS, APIs, notificações por e-mail, aplicativos, botões, widgets, anúncios, serviços comerciais e outros que estejam vinculados aos presentes Termos (denominados coletivamente como "Serviços"). Esses Termos também abrangem todo e qualquer conteúdo, como informações, textos, links, gráficos, fotos, áudios, vídeos ou outros materiais, que sejam enviados, baixados ou exibidos dentro dos Serviços, sendo referidos de forma abrangente como conteúdo.

Esta documentação de governança pode ser revisada periodicamente pela plataforma, sendo que as modificações realizadas não terão efeito retroativo. No X, ele se apresenta como “termos de serviço” e traz uma série de normativas para os usuários. Essas normas abrangem autorização de uso, propriedade intelectual, dados, limitações e responsabilidade deles como empresa para com os usuários. A empresa afirma que o documento funciona como parte do contrato com o usuário, documento que é juridicamente vinculativo, regendo o uso do X.

Durante as primeiras páginas do documento, temos contato com pontos de atenção ao uso da plataforma, falando sobre publicidade na rede social, contrato de usuário, licenças de propriedade intelectual, um dos pontos destacados pela empresa é o conteúdo consumido e produzido ser de total responsabilidade dos seus usuários, afirmando que é possível que o usuário seja exposto a conteúdos ofensivo e prejudicial postado por outros usuários.

Ao abordar os conteúdos produzidos e consumidos, os termos de serviço da plataforma afirmam que o usuário só deve compartilhar os conteúdos que deseja e esteja confortável em compartilhar. Na documentação a plataforma afirma que o X não apoia, declara ou garante a completude, veracidade, precisão ou confiabilidade de qualquer conteúdo ou comunicação postada por meio dos Serviços, nem endossar quaisquer opiniões expressadas pelos Serviços. O X deixa claro aos usuários que ao aceitar o termo de serviço, o usuário afirma estar ciente que estará exposto a conteúdos que podem ser classificados como conteúdo ofensivo, prejudicial ou inadequado. A Rede Social afirma que não pode monitorar ou controlar os conteúdos postados através do serviço e não pode ser responsabilizada pelos conteúdos publicados por usuários.

Porém, apesar dos avisos citados acima, a Rede Social diz ter o direito de remover conteúdos que violem o acordo do usuário incluindo violações que possam ferir autoria e propriedade intelectual, apropriação indébita de propriedade intelectual, falsa identidade, conduta ilegal ou assédio. Existe um processo de denúncia para violações de direitos, como infração de direitos autorais, e os usuários podem recorrer de decisões de moderação tomadas contra suas contas.

Os termos podem ser alterados de tempos em tempos, e devem ser lidos juntamente com o contrato de usuário, uma espécie de regras e políticas da plataforma.

5.2.2 Diretrizes de comunidade no X

As diretrizes de comunidade do X aparecem de forma menos formal, se apresentando como Regras e políticas em blog diretamente no servidor da rede social chamado “Central de Ajuda”. Segundo a empresa, as regras presentes visam garantir que todas as pessoas possam ter dentro do x uma conversa pública livre e segura. Na central de ajuda ao acessar a página você encontra página de regras, a página é segmentada por assuntos que podem ser acessados em diferentes com regras sobre cada assunto.

Respectivamente o assuntos contemplados pela seção de Regras e políticas são: Geral, Integridade e autenticidade da plataforma, Segurança e crimes Virtuais, Propriedade intelectual, Diretrizes de uso da plataforma, configurações de cota, diretrizes para autoridades policiais, pesquisa e experiências e recursos específicos do país.

Para nos atermos à nossa proposta de pesquisa as seções que foram utilizadas nesta monografia são “Geral”, “Segurança e crimes virtuais” “Diretrizes de uso da plataforma”. Dentro de cada seção existem subdivisões por assuntos que são abordados pela empresa e suas regras dentro da Rede Social.

Na seção “geral” os assuntos abordados estão distribuídos em tópicos. O tópico de segurança no uso de plataformas de redes sociais abrange medidas para proteger os usuários contra comportamentos nocivos e ilegais. Entre as principais diretrizes, destacam-se:

1. Discurso de ódio: proibição de ameaças e ofensas violentas.
2. Entidades violentas: não é permitido promover atividades de grupos de ódio.
3. Exploração sexual de menores: tolerância zero.
4. Abuso/assédio: proibição de bullying ou incitação.
5. Conduta de ódio: ataques com base em identidade pessoal são proibidos.
6. Mídias sensíveis: limitações rigorosas quanto ao conteúdo explícito.
7. Suicídio e crimes: proibição de incentivo ao suicídio e uso para atividades ilegais.

A plataforma X estabelece diretrizes rigorosas quanto à privacidade, proibindo a divulgação de informações pessoais sem consentimento, a exposição de material íntimo sem permissão e o acesso não autorizado a contas alheias. No quesito autenticidade, é

vedada a manipulação de informações, a fraude em processos cívicos e o uso de identidades falsas. Também há proibições quanto à disseminação de mídias manipuladas e violação de direitos autorais e marcas registradas. Além disso, o uso de publicidade de terceiros em vídeos é restrito, com penalidades aplicáveis em caso de violações.

Na seção "Segurança e Crimes Virtuais", a plataforma X apresenta suas políticas de maneira detalhada, abordando temas como conduta e propagação de ódio, assédio e abuso, e mídia sensível. Além disso, há regras específicas sobre entidades que propagam violência, autores de ataques violentos e glorificação da violência. Também são tratadas políticas sobre produtos ilegais ou regulamentados, discurso violento, informações privadas, automutilação e suicídio, nudez não consensual, organizações violentas, informações de perfil abusivas, e exploração sexual de menores.

5.3 As ferramentas de gerenciamento de comunidade do Bluesky

O Bluesky tem em sua equipe de criação um dos mesmo criadores do X (na época Twitter). Ele foi lançado em sua versão Beta em 2023 e em sua versão oficial e aberta em fevereiro de 2024. Apesar de ter características semelhantes ao X, sendo também uma rede de microblogging, apresenta diferenças que impactam diretamente a comunidade de usuários dentro do seu sistema: a rede é federada. Através de um protocolo de código aberto, chamado pela empresa de Protocolo AT (AT Protocol), a rede social se propõe a não ter o monopólio do seu código, deixando disponível para que federações possam surgir e criar novas versões desta mesma rede e novos *feeds* personalizados de acordo com o interesse de um grupo ou usuário. Mediante isto, vamos listar as ferramentas que afetam gerenciamento de conteúdo, interação com usuário e *feeds* personalizados.

5.3.1 Ferramentas de interação com conteúdos no Bluesky

As redes sociais abordadas neste estudo apresentam semelhanças e diferenças em relação às ferramentas e utilidades oferecidas aos usuários. Apesar dessas diferenças, é possível identificar ferramentas voltadas ao gerenciamento de conteúdo produzido e consumido.

A Primeira ferramenta que diz respeito ao gerenciamento de conteúdo é a Configurações de interação de post. Através dessa ferramenta, as pessoas que irão

publicar podem customizar quem pode interagir em seu post. Essa ferramenta funciona clicando no botão “Interact With Someone” após o clique no menu de interação apresentada o usuário pode desativar ou ativar as citações para aquele post específico e decidir quem pode responder seu post. As configurações de interação de posts podem ser alteradas após a publicação.

Figura 17 - Configuração de interação de post

Configurações de interação de postagem

Personalize quem pode interagir com esta postagem.

Configurações de citações

Citações habilitadas ☒

Configurações de resposta

Permitir respostas de:

Todos ✓ **Ninguém**

Ou combine estas opções:

Usuários mencionados

Usuários seguidos

Usuários em "Quero"

Salvar

Fonte: *Bluesky*

A segunda ferramenta é a Silenciar Thread. Ao utilizar essa ferramenta o usuário pode evitar que um thread aparece em seu *feed* ou notifique caso esse conteúdo seja atualizado. É possível escolher essa configuração clicando no menu post no qual o usuário deseja mostrar e selecionando para mutar.

Figura 18- Menu de post - Silenciar Thread

Fonte: *Bluesky*

A Terceira ferramenta é o Silenciar palavras e tags (Figura 19). Os posts que aparecem na tela do usuário podem ser multados baseados nos textos, nas tags ou nos dois juntos. Ao clicar na ferramenta, uma nova tela aparece para que o usuário adicione a palavra ou tags que deseja mudar, escolha o tempo em que essa palavra ou tag será multada e decidir se os usuário que utilizam essa tag podem ser excluídos dos seus seguidores.

Figura 19 - Adicionar palavras e tags silenciadas

 A screenshot of a mobile application screen titled 'Adicionar palavras/tags silenciadas'. Below the title is a paragraph of explanatory text: 'As postagens podem ser silenciadas com base em seu texto, suas tags ou ambos. Recomendamos evitar palavras comuns que aparecem em muitas postagens, pois isso pode resultar em nenhuma postagem sendo exibida.' Below the text is a text input field with the placeholder 'Digite uma palavra ou tag'. Underneath is a 'Duração:' section with four radio button options: 'Para sempre' (selected), '24 horas', '7 dias', and '30 dias'. Below that is a 'Silenciar em:' section with two radio button options: 'Texto e tags' (selected, with a document icon) and 'Apenas Tags' (with a hashtag icon). At the bottom is an 'Opções:' section with a checkbox labeled 'Excluir usuário que você segue'. A large blue button with the text 'Adicionar +' is at the very bottom.

Fonte: *Bluesky*

Outra das ferramentas disponibilizadas pelo Bluesky é o Mostrar mais ou Mostrar menos. Através dessas opções o usuário que gosta de um tipo de post pode clicar no botão “mostrar mais desses” para que o algoritmo coloque mais posts do mesmo tipo no *feed*, e caso clique em “mostrar menos desses” para que os tipos de posts sejam mostrados com menos frequência.

Além de poder decidir quem pode interagir com o seu post, como foi citado anteriormente, existe outra ferramenta que diz respeito à interação das pessoas com o conteúdo: Se remover de um repost. Com esta ferramenta, se alguém postou um conteúdo de outro usuário comentando algo sobre ele, o usuário repostado pode escolher se retirar desta publicação. Ao escolher se remover o conteúdo aparece indisponível e com uma mensagem de “Removido pelo autor” (fig 20).

Figura 20 - Removido pelo autor



Fonte: própria do autor

A sexta ferramenta é Denúncia de postagem (Figura 20). Ao clicar no conteúdo indesejado e abrir o menu presente no post, o usuário escolher reportar o post, indicando ao Bluesky que aquele post precisa ser revisto. Ao clicar no botão da ferramenta, uma tela se abre oferecendo opções de motivos pelos quais o post deve ser revisto. As opções incluem posts enganadores, spam, conteúdo sexual indesejado, comportamento antisocial - que pode ser classificado como intolerância, trolls da internet e assédio, ilegal e urgente que inclui violação da lei e dos termos de serviço e outros. Ao escolher uma das opções a plataforma abre uma página onde é possível escolher para qual serviço de moderação de comunidade a denúncia será feita, podendo escrever um texto explicativo sobre a situação e o motivo de acreditar que o conteúdo é prejudicial.

Figura 21 - Denúncia de post**Denunciar esta postagem**

Por que esta postagem deve ser analisada?

Postagem Enganosa

Falsificação de identidade, desinformação ou alegações falsas >

Spam

Menções ou respostas excessivas >

Conteúdo Sexual Indesejado

Nudez ou conteúdo adulto sem rótulo aplicado >

Comportamento Anti-Social

Assédio, intolerância ou "trollagem" >

Illegal e Urgente

Violações flagrantes da lei ou dos termos de serviço >

Fonte: Bluesky

Como utilidade dentro do menu de moderação, é possível encontrar o Filtro de conteúdo, através dele, o usuário pode habilitar o recebimento de conteúdo adulto, além de avaliar a possibilidade de receber conteúdos de nudismo sem cunho sexual, as opções para o recebimento deste tipo de conteúdo são “mostrar”, “avisar” e “esconder”.

Figura 22 - Filtros de conteúdo**Content filters**

Enable adult content

Disabled ☐Adult content can only be enabled via the Web at bsky.app.**Non-sexual Nudity**

E.g. artistic nudes.

Show

Warn

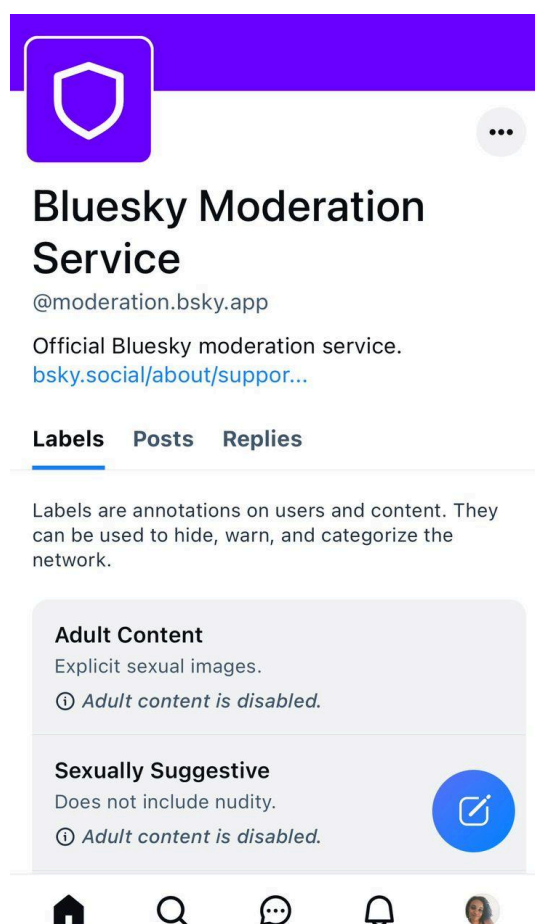
Hide

Advanced**Bluesky Moderation Service**Official Bluesky moderation service.
<https://bsky.social/about/support/c...> >

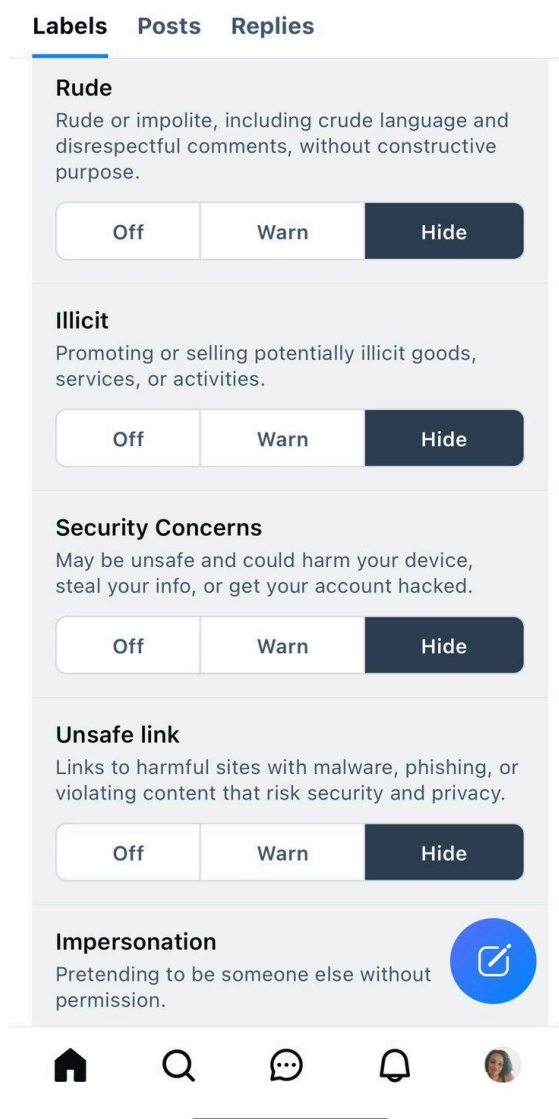
Fonte: Bluesky

O Serviço de Moderação do Bluesky é uma ferramenta de configuração avançada. Em configurações avançadas o usuário pode fazer configurações mais detalhadas para que conteúdos que contém etiquetas especificar não sejam mostrados para o usuário ou antes de serem mostrados forneçam um aviso. As etiquetas oferecidas pelo serviço e moderação, semelhante a algumas ferramentas vistas anteriormente, também apresentam as opções “esconder”, “comunicar” e “mostrar”, apesar disso os rótulos são mais detalhados e podem ser personalizados por cada usuário individualmente. Os rótulo da ferramenta incluem conteúdo adulto, mídias gráfica automutilação, conteúdo sensível, extremista, intolerante, ameaças, rude, ilícito, preocupação com segurança, links não confiáveis, representação, desinformação, golpes, Farm de engajamento, spam, rumores e conteúdo enganador.

Figura 23 - O serviço de moderação do Bluesky

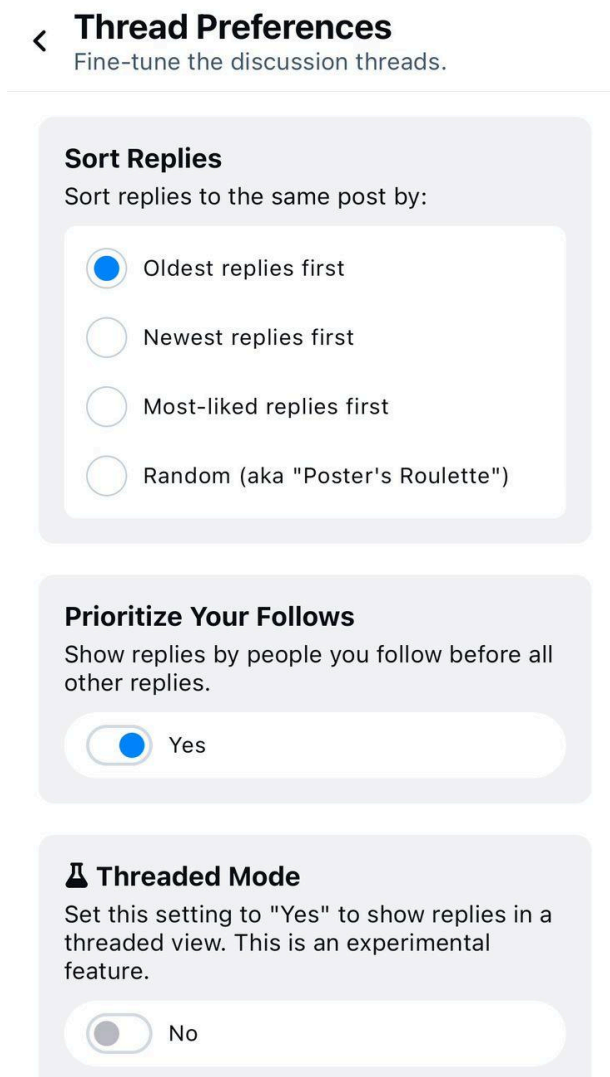


Fonte: Bluesky

Figura 24 - Etiquetas de moderação

Fonte: Bluesky

Outra utilidade disponibilizada pela plataforma é a Preferências de Threads (Figura 25). Através dessa ferramenta, o usuário pode afinar notificações e interações com um post threads, mostrando ou não a interação ou discussões que estão acontecendo.

Figura 25 - Preferências de Thread

Fonte: Bluesky

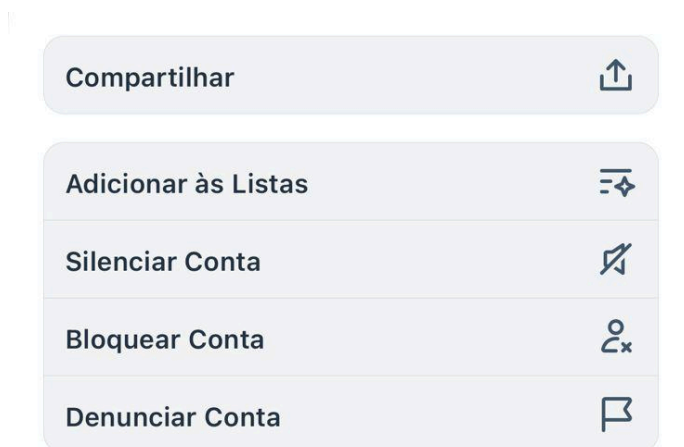
5.3.2 Ferramentas de interação com o usuário

No Bluesky, diferentemente do X, não é possível criar uma conta privada. Avaliamos como um usuário pode evitar ou restringir o contato com pessoas indesejadas dentro da plataforma. As ferramentas listadas estão ligadas a interação de usuário e não necessariamente seu conteúdo.

A primeira ferramenta que vamos elencar é o Bloquear Conta (Figura 26). Através dessa ferramenta, o usuário pode bloquear a conta que desejar, após isso, a conta bloqueada não poderá responder publicações, mencionar o usuário que a bloqueou nem interagir de nenhuma forma com essa conta. O usuário que bloqueou não terá mais

os conteúdos do que foi bloqueado e a outra parte não receberá mais em seu *feed* conteúdos seus. Após o bloqueio o usuário pode acessar uma lista com todos os usuários que bloqueou.

Figura 26 - Bloquear conta do Bluesky



Fonte: Bluesky

A segunda ferramenta é o Silenciar Conta. Com esta aplicação o usuário poderá silenciar a conta que deseja impedindo que posts e comentários em outros posts feitos pela conta bloqueada apareçam em sua tela inicial de postagens. Após o silenciamento não são geradas nenhuma lista com usuários silenciados, mas aparece uma indicação ao entrar no perfil de que o usuário foi silenciado como podemos ver na Figura 27. É possível dessilenciar este usuário a qualquer momento.

Figura 27 - Conta Silenciada



Fonte: Bluesky

A Denúncia de Conta é a terceira ferramenta listada. Através desta aplicação, a pessoa que teve um interação indesejada ou se sentiu violada com esta interação pode

solicitar que o perfil do usuário seja avaliado pela equipe de moderação do Bluesky. Ao clicar na opção Denúncia de conta, uma tela se abre onde o usuário pode escolher entre as opções “conta enganosa”, “frequentemente posta conteúdo indesejado”, “nome ou descrição viola os padrões da comunidade” e “outros” (Figura 28).

Figura 28 - Denunciar este usuário

Denunciar este usuário
Por que este usuário deve ser analisado?

- Conta Enganosa**
Falsificação de identidade ou alegações falsas sobre identidade ou filiação
- Frequentemente Posta Conteúdo Indesejado**
Spam; menções ou respostas excessivas
- Nome ou Descrição Viola os Padrões da Comunidade**
Termos utilizados violam as diretrizes da comunidade
- Outro**
Outro problema

Precisa denunciar uma violação de copyright? [Ver detalhes](#)

Fonte: Bluesky

A quarta ferramenta é a Lista de usuário (Figura 29). Com essa ferramenta o usuário pode criar listas públicas e compartilháveis de assunto variáveis. Através dessas listas novos *feeds* podem ser criados.

Figura 29- Lista de usuário

Fonte: Bluesky

A quinta ferramenta é a Lista de Moderação (Figura 30), que permite ao usuário bloquear ou silenciar múltiplas contas de uma só vez. Caso as listas sejam utilizadas de forma maliciosa, como para perseguição ou envio de spam, o time de gerenciamento de comunidade do Bluesky poderá desativá-las. Além disso, os usuários têm a opção de denunciar listas consideradas abusivas.

Figura 30 - Lista de moderação

Fonte: *Bluesky*

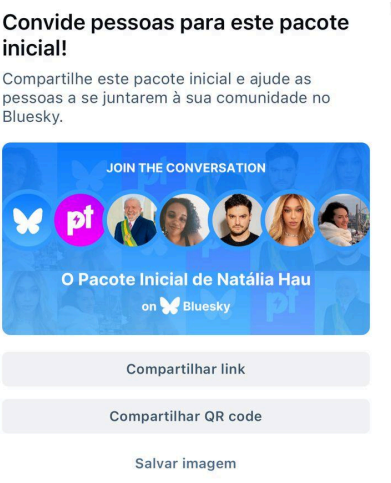
Pacote Inicial é outra das funções disponibilizadas pelo Bluesky. Através da ferramenta, o usuário pode criar um kit inicial para novos usuário, com assunto e temas específicos e pessoas que tem a ver com esse tema, assim, um novo usuário ao entrar e seguir o Pacote Inicial, seguirá todo os seguidores presentes neste pacote, criando uma comunidade inteiramente personalizada logo em seu primeiro momento. As listas iniciais podem ser denunciadas pelos usuários.

Figura 31 - Pacote inicial



Fonte: Bluesky

Figura 32 - Compartilhar Pacote inicial



Fonte: Bluesky

A sétima ferramenta é as mensagens diretas, através dela os usuários podem se comunicar entre si de forma privadas, porém a Rede Social oferece configurações de chat, onde o usuário pode definir e alterar qualquer momento as permissões de mensagem podendo escolher entre “todos”, usuários que eu sigo” e “ninguém”, independente do momento da alteração, conversas que atuais podem continuar.

Figura 33 - Mensagens diretas configurações de chat

< **Configurações do Chat**

Permitir novas mensagens de

Todos ☐

Usuários que eu sigo ☒

Ninguém ☐

Você pode continuar conversando,
independentemente da configuração que
escolher.

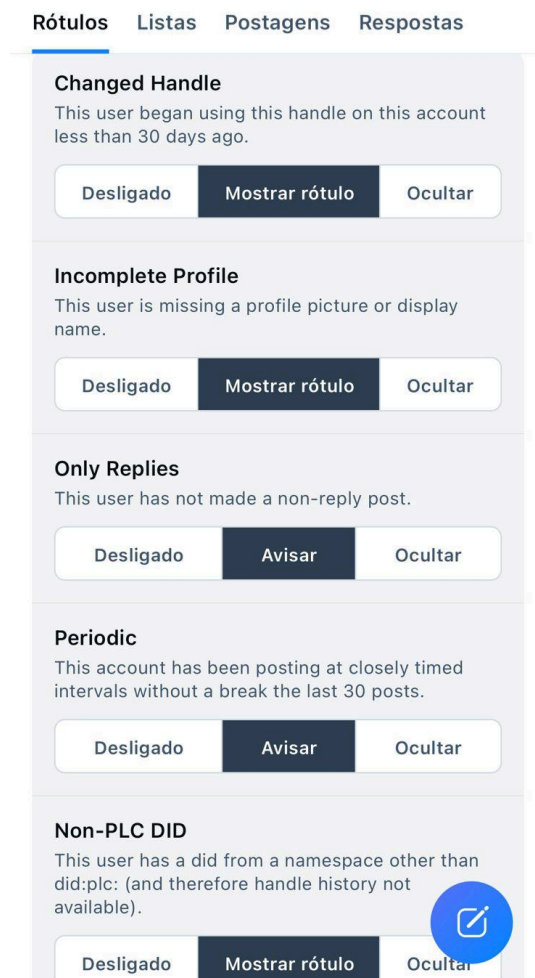
Sons de Notificação

Habilitado ☒

Desabilitado ☐

Fonte: Bluesky

A última ferramenta são os rótulos, com esta ferramentas qualquer usuário pode criar uma rotuladora automática, se baseando no código da Rede Social para criar novos rótulos de conteúdo. Ao se tornar um rotulador uma conta de rotulador será criada e todas as pessoas que seguirem irão adicionar os rótulos em sua usabilidade podendo configurá-los de acordo com suas necessidades (Figura 34).

Figura 34 - Lista de rótulos

Fonte: Bluesky

5.3.3 Feeds do Bluesky

Para a personalização da usabilidade do usuário nesta rede, ao invés de oferecer comunidades como sua rede social irmã, o Bluesky oferece uma opção diferente: os *Feeds*. *Feeds* customizados é um ferramenta que fornece algoritmos personalizados aos usuários por meio do AT Protocol. Essa funcionalidade permite que os usuários escolham suas próprias linhas do tempo, baseadas em algoritmos, fazendo com que um *feed* específico esteja à disposição do usuário.

Os *feeds* customizados funcionam através de um servidor que recebe uma solicitação do servidor do usuário e retorna uma lista de postagens, com a opção de anexar alguns metadados.

Como o Bluesky não implementa comunidades como no X, os *feeds* foram utilizados para analisar um conjunto de pessoas e conteúdos relacionados. Foram mantidas as temáticas de cultura pop, com um *feed* destinado a fãs da Marvel e o outro *feed* com temática científica tendo em grande parte dos usuários inseridos em um contexto acadêmico. Através do uso nesses dois *feeds* foi possível listar algumas ferramentas específicas.

Uma das ferramentas é a regras de uso do *feed*. Nessas regras ao entrar no *feed* é possível encontrar um post fixo no topo indicando quais são as regras de uso do *feed* escolhido pelo usuário. Apesar de existir esta forma de normatização, não é preciso de aprovação dos administradores ou criadores do *feed* para que os usuários interajam e recebam as publicações feitas internamente.

A segunda ferramenta disponibilizada é a denúncia de *feed*. Esta ferramenta se comporta como uma ferramenta de denúncia de perfil, oferecendo as opções de denúncias devido ao nome ou descrição que viole os padrões da comunidade, comportamento anti-social e comportamento ilegal e urgente.

Hashtags de uso é uma forma do post aparecer no *feed* específico no qual o usuário está participando, cada criador de *feed* determina qual será a palavra, hashtag ou emoji específico para que o seu *feed* seja utilizado. Ademais, o uso se dá agregando as ferramentas de interação com conteúdo e de interação com o usuário citados anteriormente.

5.4 Termos de uso e diretrizes da comunidade do Bluesky

Apesar de ser uma rede consideravelmente nova, o Bluesky também tem regras e documentos normativos e de governança para os usuários em sua plataforma. Porém essa documentação normativa não se organiza em forma de documento e sim em publicações em sua plataforma oficial.

5.4.1 Termos de uso do Bluesky

O Termo de Uso do Bluesky é chamado de Termo de serviço e está hospedado no site externo da empresa em formato de postagem de blog. Logo no topo da página a empresa diz que seus documentos de governança se aplicam somente ao Bluesky e não

as redes federadas que estão dentro do protocolo At, incluindo os Sites e o Aplicativo Bluesky. Caso o usuário utilize outra aplicação de rede social no Protocolo AT que não seja o Bluesky Social (referida como “Aplicação do Desenvolvedor”), os desenvolvedores dessa aplicação fornecerão os termos e condições que regerão a experiência do usuário.

Os termos abrangem, inicialmente, quem pode usar o Bluesky Social. Em seguida, fazem uma análise sobre o conteúdo gerado pelos usuários. São abordadas as proibições gerais e as medidas de aplicação, tema da rescisão de contas e os procedimentos de apelação. Além disso, o texto discute a questão dos links para sites ou recursos de terceiros, observando as implicações de se acessar conteúdos externos à plataforma.

O documento destaca a necessidade de se conectar com as diretrizes da comunidade e ao aceitá-lo, então o usuário terá concordado com as regras que estão descritas nas diretrizes. Sobre conteúdos produzidos na plataforma, o produto do conteúdo deve se responsabilizar totalmente por seus conteúdos. O utilizador deve compreender e concordar que pode estar exposto a conteúdo de usuário que seja impreciso, não confiável, falso, desatualizado, incompleto, inapropriado, inadequado para crianças ou de outra forma inadequado ao seu propósito.

5.4.2 Diretrizes da comunidade do BlueSky

No Bluesky, a documentação segue o mesmo perfil da apresentação dos termos de uso, de uma forma mais informal sendo apresentada através de publicações de blogs em um setor em seu site chamado “suporte”. Segundo a empresa, por oferecer diversos tipos de serviço por consequência de sua característica federada, sendo o Bluesky só um desses serviços, as diretrizes só dizem respeito ao Bluesky, deixando claro que qualquer outro serviço dentro de sua federação através do Protocolo AT terá suas próprias diretrizes.

Um dos seus pontos introdutórios na documentação da plataforma é sua proposta de criar comunidades que respeitam as preferências individuais e se adaptam de acordo com as necessidades de seus usuários. As Diretrizes de comunidade da plataforma se dividem em três pontos: Empoderar o usuário, Cultivar um ambiente acolhedor e melhorar como *feedbacks*.

A plataforma afirma que as Diretrizes da comunidade foram desenhadas para promover uma experiência agradável e segura. A empresa afirma que sua moderação é feita de forma simples e discreta durante o uso do usuário e também através de filtros de conteúdo disponíveis na rede social. Apesar disso, o usuário que receber um conteúdo, mensagem ou perfil que está quebrando as regras, deve realizar a denúncia para a plataforma.

Como primeiro ponto abordado pela documentação, a empresa deixa como regra o “Não Use Bluesky Social para quebrar a lei e causar danos a outros”. O Bluesky começa citando o exemplo do tipo de material que não pode ser distribuído do uso da rede social com a não promoção, ou distribuição de material de grupos de ódio ou organizações terroristas proibidas. Além disso, a disseminação de material que envolva abuso sexual infantil ou qualquer conteúdo que retrate menores em contexto sexualmente explícito é expressamente vedada. A exploração sexual e o tráfico humano, incluindo qualquer tentativa de distribuir, participar ou normalizar o abuso sexual infantil, são práticas inaceitáveis. Incluindo também a troca de interações predatória de adultos com menores de idade .

O comércio de substâncias ou produtos ilegais, bem como a distribuição de informações pessoais privadas sem o devido consentimento, configuram práticas ilícitas. Do mesmo modo, a invasão ou o acesso não autorizado a sistemas, a prática de fraudes para ganho financeiro e o uso de técnicas como spam e phishing, com o objetivo de prejudicar a experiência de outros usuários, são estritamente proibidos.

No que tange à propriedade intelectual, é necessário respeitar as leis de direitos autorais, marcas registradas e outras formas de propriedade intelectual. A interferência em processos eleitorais, como a disseminação de desinformação ou a intimidação de participantes, constitui uma grave violação. Não se deve compartilhar informações enganosas falsamente atribuídas a candidatos ou incitar a perturbação dos processos eleitorais.

Em seu segundo ponto diz respeito a interação de usuários com usuários, envolvendo o respeito ao tratar e interagir com outros perfis. A documentação da rede social estabelece diretrizes que proíbem exemplos claros de comportamentos inadequados, como o assédio ou abuso direcionado a indivíduos ou grupos específicos. Entre esses comportamentos, destaca-se o assédio sexual e o assédio vinculado ao gênero, os quais são terminantemente proibidos. Também é proibido promover ódio e extremismo escolhendo pessoas e grupos como alvo de ataques por suas características

de raça, gênero, religião, etnia, nacionalidade, deficiência e orientação sexual. Além disso, também é proibido a publicação de ameaças de violência, danos físicos ou conteúdos que encorajam ou glorificam a violência. Também não é permitido encorajar suicídio ou automutilação, mostrar descrição de violência excessiva, *gore* e tortura, desmembramento e atividade sexual não consensual.

A utilização de aplicativos e serviços digitais deve incluir as normas e padrões estabelecidos pelos desenvolvedores de cada plataforma. No âmbito do Protocolo AT, os desenvolvedores têm a liberdade de definir diretrizes comunitárias específicas para o uso de seus aplicativos e serviços. O descumprimento dessas diretrizes pode acarretar sanções, como a restrição do acesso ou uso desses serviços.

Para assegurar o cumprimento das diretrizes da comunidade do Bluesky, é necessário o envolvimento de toda a comunidade, incluindo desenvolvedores, criadores, usuários e visitantes. A plataforma disponibiliza um sistema de denúncia, tanto no aplicativo quanto online, que facilita a detecção de violações das diretrizes. Os usuários podem denunciar infrações por meio da funcionalidade de Denunciar posts, contas ou mensagens diretas. As denúncias são avaliadas de acordo com as diretrizes da Bluesky, e a plataforma se reserva o direito de decidir quando e como agir, podendo impor restrições ao uso do serviço em caso de violações.

A documentação da plataforma abrange especificamente, no terceiro ponto, o uso da ferramenta Rótulos, uma das ferramentas que a comunidade do Bluesky tem para controlar sua experiência e ajudar na classificação de conteúdo. As diretrizes recomendam que os rótulos sejam claros e não abusivos, baseando-se em conteúdo e comportamento observáveis. A aplicação dos rótulos deve ser feita de maneira imparcial, ética e consistente, visando minimizar vieses, o que aumenta a confiança dos usuários da ferramenta de rotulagem. É proibido o uso de rótulos de maneira abusiva, com o objetivo de assediar ou direcionar ataques a contas, grupos ou características protegidas.

6. ANÁLISE

Com base nos dados coletados na metodologia previamente e etapa anterior mencionada, foram desenvolvidos quadros para análise e comparação dos crimes de assédio online, doxing, revenge porn e perseguição. Cada ferramenta será listada e

organizada segundo as seguintes classificações: coíbe e pune os crimes mencionados neste texto.

Os dados, crimes e classificações foram organizados por meio de uma análise comparativa qualitativa, com uma breve explicação da função de cada ferramenta. Os quadros estão separados de acordo com cada rede social estudada e a finalidade de cada ferramenta, seja para interação com conteúdo ou com usuários. Assim, foram gerados quatro quadros contendo os dados necessários para a análise.

Na primeira visualização (Quadro 1) é possível verificar a disposição das ferramentas de interação de conteúdo da comunidade da rede social X em relação com os crimes avaliados neste trabalho. Cinco ferramentas foram listadas a fim de terem sua utilidade avaliada.

Quadro 1- Análise das ferramentas de interação com conteúdo do X

Ferramenta	Assédio online	Doxxing	Perseguição	Revenge Porn
Filtro de Palavras	Coibir	Não interfere	Não interfere	Não interfere
Filtro de notificações	Coibir	Não interfere	Coibir	Não interfere
Mostrar com menos frequência	Coibir	Não interfere	Não Interfere	Não interfere
Denunciar post	Punir	Punir	Punir	Punir
Quem pode responder?	Coibir	Não interfere	Coibir	Não interfere

Fonte: própria do autor.

Ao analisar o filtro de palavras, é possível verificar sua compatibilidade com a possibilidade de coibir o assédio online, como este crime acontece de várias formas sendo uma delas o ataque direto a um grupo ou minoria, o filtro de palavras diminui a possibilidade de tais tipos de postagens chegarem no *feed* de uma usuária, personalizando a experiência e criando o uso personalizado e seguro. Mediante aos outros crimes listados, não foi verificado a possibilidade de punir ou coibir a sua realização com esta ferramenta.

Ao observar a ferramenta Filtro de notificações vemos que ela pode interferir no momento de impedir que o crime de assédio online aconteça, tendo base que as tentativas de contato com pessoas desconhecidas com potencial de cometer esse crime pode ser limitada, evitando o contato indesejado com esses conteúdos ou pessoas. Além disso, o uso desse recurso pode ser relevante na prevenção e punição de casos de perseguição virtual. A perseguição, no contexto online, caracteriza-se pelas tentativas insistentes do agressor de estabelecer contato com a vítima por diferentes meios, o que gera uma sensação de insegurança no espaço digital em que a pessoa está inserida.

Na ferramenta Mostrar com menos frequência temos a possibilidade de coibir o crime de assédio, já que o usuário não terá acesso com a mesma frequência de posts do mesmo teor, atacando mulheres ou até mesmo conteúdos de cunho sexual indesejado.

A ferramenta de denúncia de posts tem se mostrado eficaz em todas as instâncias virtuais analisadas nesta pesquisa. Por meio dela, é possível evitar que conteúdos que se enquadram como assédio continuem sendo propagados, além de causar consequências para o usuário que realizou o post danoso, a publicação pode ser denunciada como “Abuso ou assédio”. No caso do crime de Doxxing, a denúncia pode evitar que as informações pessoais divulgadas alcancem mais pessoas, o conteúdo pode ser denunciado como “Discurso violento” ou “Privacidade”.

Em perseguição o conteúdo pode ser denunciado como “Ódio”, “Discurso Violento” ou “Abuso e Assédio”. A denúncia tem o potencial de eficácia para que a pessoa perseguida não tenha acesso a mais posts de seu agressor e que os posts existentes sejam retirados da plataforma, caso confirmado a denúncia pelo X. Já em Revenge Porn, a denúncia se faz importante para que o conteúdo sendo divulgado sem o consentimento da vítima seja retirado, o conteúdo pode ser denunciado como “Privacidade”, “Falsa identidade” e “ódio”.

A ferramenta Quem pode responder? se enquadrrou como eficaz para coibir dois tipos de violência virtual: o assédio online e a perseguição. Através da ferramenta é possível decidir quem pode interagir com as publicações do usuário, diminuindo a possibilidade de que pessoas desconhecidas ou indesejadas interajam com conteúdo, seja em forma de assédio ou ameaças de perseguição.

As ferramentas de interação com conteúdos não necessariamente geram punição para os quartos das postagens que são retiradas do ar ou sofrem punição de entrega, as

ferramentas de interação com usuários (Quadro 2) tem o potencial de limitar a interação de pessoa com pessoa.

Quadro 2- Análise das ferramentas de interação com usuários X

Ferramenta	Assédio online	Doxxing	Perseguição	Revenge Porn
Conta privada	Coibir	Coibir	Coibir	Não interfere
Filtro de mensagens	Coibir	Não interfere	Coibir	Não interfere
Interação de post	Coibir	Coibir	Coibir	Não interfere
Posts protegidos	Coibir	Não interfere	Coibir	Não interfere
Silenciar	Coibir	Não interfere	Coibir	Não interfere
Bloquear	Coibir	Não interfere	Coibir	Não interfere
Denunciar	Punir	Punir	Punir	Punir

Fonte: própria do autor.

Ao observar as ferramentas de interação com usuários é possível perceber que elas aparecem em maior quantidade, no Quadro 2 verificamos um cenário um pouco diferente visto que as ferramentas aqui listadas têm um impacto direto na interação pessoa com pessoa.

Ao definir a conta como privada o usuário aumenta a sua proteção de dados, diminuindo a possibilidade de ter informações privadas a disposição de qualquer usuário. Além disso, o assédio online pode ser evitado através da privacidade de conta, já que só quem o usuário autoriza poderá seguir, interagir e receber as publicações feitas pela usuária. Seguindo a mesma lógica, a ferramenta tem potencial de diminuir a perseguição. Visto que somente pessoas selecionadas terão acesso ao perfil da usuária.

O Filtro de mensagens é uma ferramenta com o potencial de interferir diretamente com o crime de assédio online, visto que através desse filtro o usuário pode escolher de quais perfis deseja receber mensagem, essa personalização pode ser feita para pedidos privados e abertos e pode até mesmo selecionar somente uma pequena parcela de seguidores. Ao verificar o crime de perseguição, a lógica do Assédio online

também é aplicável visto que o agressor terá acesso limitado ou nenhum acesso a caixa de mensagens privadas do outro usuário.

Com a ferramenta Interação de post os Crimes de Assédio online e perseguição podem ser evitados por limitar o acesso ao conteúdo dos usuário malicioso contra uma conta, jpa em doxxing, a configuração de não permitir que qualquer usuário faça repost do conteúdo, pode evitar não totalmente a divulgação de informações e dados pessoais, diretamente através do repost com comentário.

Com os posts protegidos é possível diminuir as chances para que assédio online e perseguição aconteçam, visto que somente seguidores poderão receber as publicações feitas pela usuária.

Com a ferramenta silenciar, o usuário deixa de receber interação e publicações de pessoas nas quais ele acredita produzir conteúdos que se enquadram como Assédio online ou resposta insistentes de uma só pessoa, se configurando Perseguição.

Ao observar a ferramenta Bloquear, podemos verificar que ela tem o potencial de coibir os crimes de assédio e perseguição, visto que o usuário limita todo o acesso da conta, posts, perfil e mensagens diretas, além de nenhuma publicação feita pela pessoa bloqueada chegará em quem a bloqueou, porém, entendendo que o crime de perseguição tem como característica a tentativa insistente de acesso a vítima, a ferramenta de bloqueio serve como uma punição leve por excluir totalmente o usuário de ter acesso a pessoa que o bloqueou.

Através do denunciar todos os crimes podem ser punidos e evitados, desde de que a denúncia seja aprovada pela plataforma.

Os quadros acima (Quadro 1 e Quadro 2) abrangem o ambiente virtual total da Rede social, porém o X também oferece lugares exclusivos e fechados, às comunidades, por isso foi necessário listar (Quadro 3) as ferramentas exclusivas de comunidade oferecidas pelo X.

Quadro 3- Análise das ferramentas nas comunidades do X

Ferramenta	Assédio online	Doxxing	Perseguição	Revenge Porn
Regras de comunidade	Coibir	Coibir	Coibir	Coibir
Filtro de spam	Não interfere	Coibir	Coibir	Coibir

Ferramenta	Assédio online	Doxxing	Perseguição	Revenge Porn
Palavras chaves	Coibir	Coibir	Coibir	Coibir
Remover membros	Coibir	Coibir	Coibir	Coibir
Aplicação de regras da comunidade	Punir	Coibir	Coibir	Coibir

Fonte: própria do autor.

Na aba de Regras da comunidade, as regras podem mudar, mas caso deseje os moderadores podem criar leis que abrangem os crimes citados para evitar que aconteçam dentro da comunidade. No caso de acesso, os moderadores devem deixar claro sobre como uns usuário dentro da comunidade podem ou não podem agir entre si, no caso de doxxing, as comunidades que foram avaliadas nesta pesquisa, tinham regras clara sobre as postagens serem somente sobre os tópicos que aquela comunidade se propõe, sem fugir dos temas discutidos. Tentativas de interação insistentes podem ser impedidas pelas ferramentas de interação com usuários (Quadro 2) ou por regras anti Spam. Quando ao crime da pornografia de vingança (Revenge Porn), os moderadores podem adicionar regras que proibam o conteúdo pornográfico seja ele produzido por si mesmo ou terceiros.

Ao analisar a eficácia da ferramenta Filtro de Spam, em caso de Doxxing o crime viola as diretrizes das comunidades e podem ser classificados como spam. No caso de perseguição, posts repetitivos e tentativas de contato com uma potencial vítima através da comunidade também pode ser considerada assédio. No caso de Revenge Porn o post também pode ser considerado spam visto que foge de dentro dos tópicos propostos pela comunidade.

As palavras chaves se aplicam ao ajudar os administradores de comunidade a identificarem conteúdos ofensivos onde pode se enquadrar o Assédio online, além disso identificar informações pessoais como número de documentos e divulgação de endereços como acontece no doxxing e imagens indesejadas que pode incluir conteúdo pornográfico como no Revenge porn.

Ao avaliar a ferramenta de remoção de membros, é possível aplicá-la como forma de coibir e punir usuários em todos os delitos mencionados nesta pesquisa. Já no caso da ferramenta de Aplicação de Regras da Comunidade, além de servir como punição, ela atua também como um mecanismo que impede que os crimes continuem

acontecendo. No entanto, é importante considerar que, para que medidas cautelares sejam aplicadas às contas, os conteúdos abusivos ou que violem as regras da comunidade precisam ser devidamente identificados. Como consequência, essas medidas são implementadas para corrigir comportamentos inadequados

Quadro 4- Análise das ferramentas de interação com conteúdo do Bluesky

Ferramenta	Assédio online	Doxxing	Perseguição	Revenge Porn
Interação de post	Coibir	Não interfere	Coibir	Não Interfere
Mutar Thread	Coibir	Não interfere	Não interfere	Não interfere
Mutar palavras e tags	Coibir	Não interfere	Não interfere	Não interfere
Mostrar menos	Coibir	Não interfere	Não Interfere	Não interfere
Remover-se de conteúdo	Coibir	Coibir	Coibir	Não interfere
Reporte de post	Punir	Punir	Punir	Punir
Filtro de conteúdo	Coibir	Não interfere	Não interfere	Não interfere
Serviço de moderação	Coibir	Não interfere	Não interfere	Não interfere
Preferências de threads	Coibir	Não interfere	Coibir	Não interfere

Fonte: própria do autor.

No Quadro 4, as ferramentas que foram elencadas foram organizadas em uma forma clara para facilitar esta análise.

A ferramenta de Interação de post, tem como funcionalidade limitar o acesso deliberado de usuários a qualquer conteúdo protegido, o dono da publicação consegue decidir a qualquer momento o tipo de pessoa que ele deseja que tenha a liberdade de interagir com a com sua publicação. Por sua natureza de preservar a personalidade do usuário, a ferramenta se apresenta capaz de coibir a contecimentos de crimes como assédio online e perseguição.

Ao continuar a análise, é possível elencar que na ferramenta Mutar Thread e tags os crimes de assédio e perseguição são afetados diretamente ao serem impedidos de

continuar acontecendo, devido a caracterização do crime envolver contato indesejado com as possíveis vítimas.

Devido a natureza do crime de assédio inclui ser exposto a imagens indesejadas, a ferramenta Mostrar menos pode funcionar como forma de coibir a realização ou perpetuação do deste crime.

Já ao avaliar a ferramenta Remover-se de conteúdo, é possível ver sua utilidade em interferir nos crimes de assédio, perseguição doxing sejam afetados, ao abordar o doxing, a ligação da imagem do perfil com os dados pessoais podem ser evitados.

O filtro de conteúdo serve como empecilho para que conteúdos indesejados apareçam para o usuário.

A próxima ferramenta é o reporte de post, onde o usuário pode denunciar uma publicação devido ao seu conteúdo, existindo a possibilidade de que todos os crimes sejam afetados através do impedimento de que continuem acontecendo e da possível penalização e suspensão de contas que serve como punição aos usuários infratores das diretrizes de comunidade e possíveis criminosos.

O filtro de conteúdo é uma ferramenta na qual o usuário pode escolher que tipo de conteúdo deseja receber e sem desejar receber conteúdos voltados para maiores de dezoito anos, diminuindo o risco do usuário receber conteúdos de cunho sexual indesejado, interferindo diretamente no crime de assédio.

Ao avaliar o serviço de moderação do Bluesky, todos os crimes listados nesta monografia podem encontrar moderação, visto que através da ferramenta, cada teor de conteúdo pode ser indicado e comunicado para os moderadores do serviço, rotulando o tipo de conteúdo produzido, fazendo com que a conta seja desativada.

A ferramenta Preferência de Threads usuários podem evitar o crime de assédio e perseguição, deixando de interagir e receber notificações de post, se retirando das discussões que estão acontecendo em um conteúdo.

Quadro 5 - Análise das ferramentas de interação com usuário do Bluesky

Ferramenta	Assédio online	Doxxing	Perseguição	Revenge Porn
Bloquear conta	Punir	Punir	Punir	Punir
Mutar conta	Coibir	Não Interfere	Coibir	Não Interfere

Ferramenta	Assédio online	Doxxing	Perseguição	Revenge Porn
Denúncia de conta	Coibir	Coibir	Coibir	Punir
Lista de usuários	Coibir	Não Interfere	Não Interfere	Não Interfere
Lista de moderação	Coibir	Coibir	Coibir	Coibir
Pacote inicial	Coibir	Não Interfere	Não Interfere	Não Interfere
Configuração de mensagens diretas	Coibir	Não Interfere	Coibir	Não Interfere
Rótulos	Coibir	Não Interfere	Coibir	Não Interfere

Fonte: própria do autor.

No Quadro 5 estão sendo avaliadas 8 ferramenta de interação de usuário com usuário, apresentando diferenças para o mesmo tipo de ferramenta do x, começando com a citação da ferramenta Bloquear a conta, que vai ajudar o usuário de assédio e perseguição a não receber conteúdos de um usuário indesejado. Apesar disso, não existem demais interferências a respeito da punição e coibição dos outros crimes investigados nesta monografia.

A ferramenta Mutar Conta, apresenta a mesma utilidade de bloquear, interagindo e afetando na realização dos crimes que foram citados acima.

A denúncia de conta do bluesky funciona como uma forma de punir, visto que a pessoa infratora pode ter sua conta excluída.

A Lista de Usuário é uma ferramenta criada para gerar uma navegação personalizada, podendo ser eficaz no momento de evitar o crime de assédio, visto que a usuária não terá contato inicial com pessoas fora de sua lista e seu *feed* específico, sem receber publicações feitas pelo possível agressor.

Com a Listas de moderação o crime de perseguição e assédio podem ser evitados, através da ferramenta a usuária pode adicionar em sua listas perfis suspeitos e para manter bloqueados e mutados.

Através do pacote inicial, um novo usuário pode personalizar sua navegação, utilizando a lista personalizando por um amigo e tendo acesso ao mesmo tipo de *feed* que a amiga tem, evitando o recebimento de publicações indesejadas em seus *feeds*.

Na configuração de mensagens diretas, a usuária tem como limite quem pode mandar mensagens internas, incluindo a possibilidade de limitar o acesso de pessoas que são seguidoras mútuas, sem que a amizade seja desfeita.

Através do Rótulo, o usuário consegue perceber e se proteger de rótulos de pessoas recentes, ou que mudam muito de nome de usuário e fotos de perfil, e sinalizando muitas informações relevantes para que possíveis usuárias identifiquem comportamentos suspeitos por parte dos seus colegas. A ferramenta pode ser uma boa ferramenta quando o assunto é perseguição e assédio, mas sem que ela avalie sozinha.

Quadro 6 - Análise das ferramentas dos feeds do Bluesky

Ferramenta	Assédio online	Doxxing	Perseguição	Revenge Porn
Regras de uso do <i>feed</i>	Não Interfere	Não Interfere	Não Interfere	Não Interfere
Denúncia de <i>feed</i>	Punir	Punir	Punir	Punir
Hashtags de uso	Não Interfere	Não Interfere	Não Interfere	Não Interfere

Fonte: própria do autor.

Como explicado anteriormente, o Bluesky não se organiza da forma no qual o X se estrutura. Ao conferir a descrição regras do *feed* e comparar com a análise (Quadro 6), elas não se mostraram eficazes em nenhum dos crimes visto que a regras não são pré-requisitos para participar do X de um *feed*, e qualquer pessoa pode se somar a tarefa.

Em contrapartida, tentativas de organizações de *feeds* que ofendem as regras da comunidade ou são de alguma forma criminosas pode ser impedida e punido através da denúncia para a equipe de moderação. essa ferramenta pode afetar todos os crimes, visto que os *feed* que propagam notícias, informações pessoais e revenge porn podem ser punidos por acontecerem dentro dos *feed* que foram criados com um tom maldoso.

E por fim, a ferramenta hashtag de uso não interfere em nada na realização ou punição de todos os crimes citados dentro da plataforma.

Na Tabela 1, são apresentados os dados referentes à quantidade de ferramentas disponíveis em cada rede social analisada, categorizadas como mecanismos para coibir, punir ou não interferir nos crimes investigados. Esses dados permitem uma compreensão mais detalhada sobre a efetividade de cada plataforma. De maneira geral, observa-se que a rede X dispõe de 17 ferramentas, enquanto a Bluesky apresenta 20. Contudo, um aspecto determinante na avaliação das plataformas não é apenas o número de ferramentas, mas a qualidade e a efetividade de suas funcionalidades na prevenção e punição das infrações abordadas nesta pesquisa.

Tabela 1 - Levantamento quantitativo de ferramentas e crimes X e Bluesky

	X			Bluesky		
Crime	Coíbe	Pune	Não interfere	Coíbe	Pune	Não interfere
Assédio online	13	2	2	15	3	2
Doxxing	7	2	8	3	3	14
Revenge Porn	5	2	10	1	4	15
Perseguição	12	2	3	8	3	9

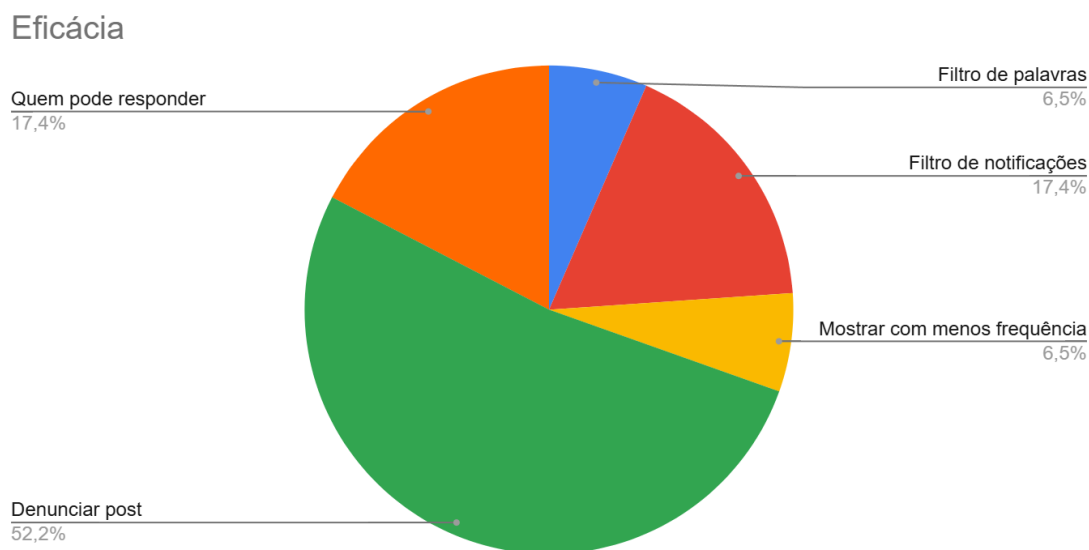
Fonte: própria do autor

Ao analisar os dados levantados, discussões sobre sua efetividade e qualidade das ferramentas de cada uma das redes serão discutidas, para que possamos averiguar sua interferências nos crimes abordados nesta presente pesquisa.

7. RESULTADOS E DISCUSSÕES

Após a análise das ferramentas e sua interferência nos crimes de assédio, doxxing, perseguição e revenge porn, algumas discussões serão levantadas para fins de conclusão do trabalho, como os dados de cada rede foram divididos em três quadros, vamos abordar resultados gerais e resultados específicos avaliando quantitativamente a eficácia de cada quadro de ferramentas.

Gráfico 1 - Porcentagem de eficácia das ferramentas de interação do conteúdo do X



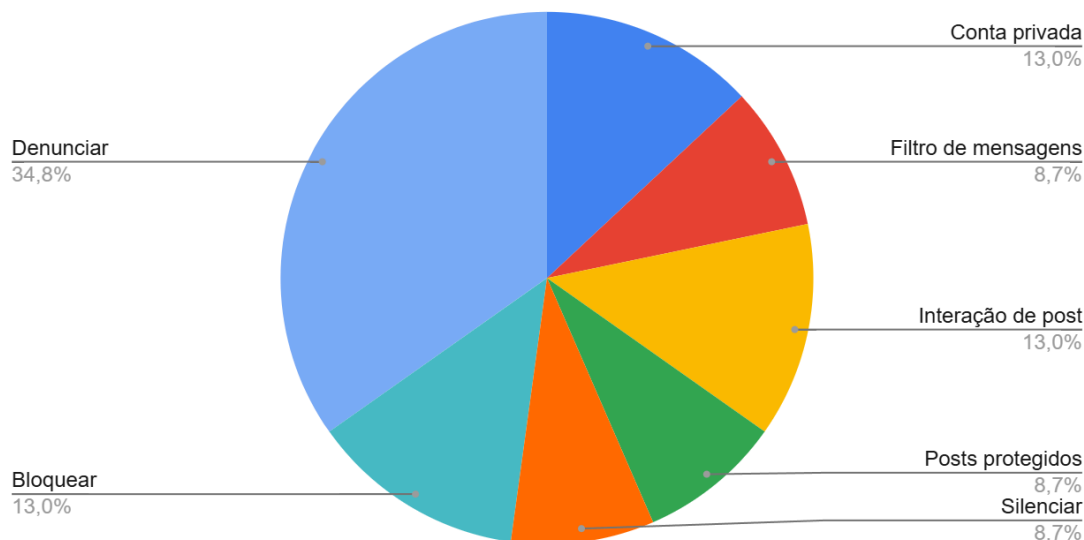
Fonte: própria do autor

No Gráfico 1, observa-se que apenas uma das ferramentas demonstrou interferência em todos os crimes abordados na análise. Além disso, os crimes de doxxing e revenge porn não podem ser coibidos preventivamente pelas ferramentas listadas (Quadro 1); essas infrações apenas podem ser punidas após sua ocorrência, uma vez que o conteúdo já precisa estar online para que a denúncia seja realizada.

As ferramentas que demonstraram maior interferência foram respectivamente: Denunciar post, Filtro de Notificações e Quem Pode Responder. O Denunciar post funciona em todas as instâncias para coibir e punir todos os tipos de crimes listados, já o Filtro de notificações tem o potencial de interferir nos crimes de assédio e perseguição, tendo em vista a possibilidade do usuário limitar seu acesso a tentativas de contatos e posts indesejadas de usuários maliciosos.

Gráfico 2 - Porcentagem de eficácia das ferramentas de interação com usuário do X

Eficácia

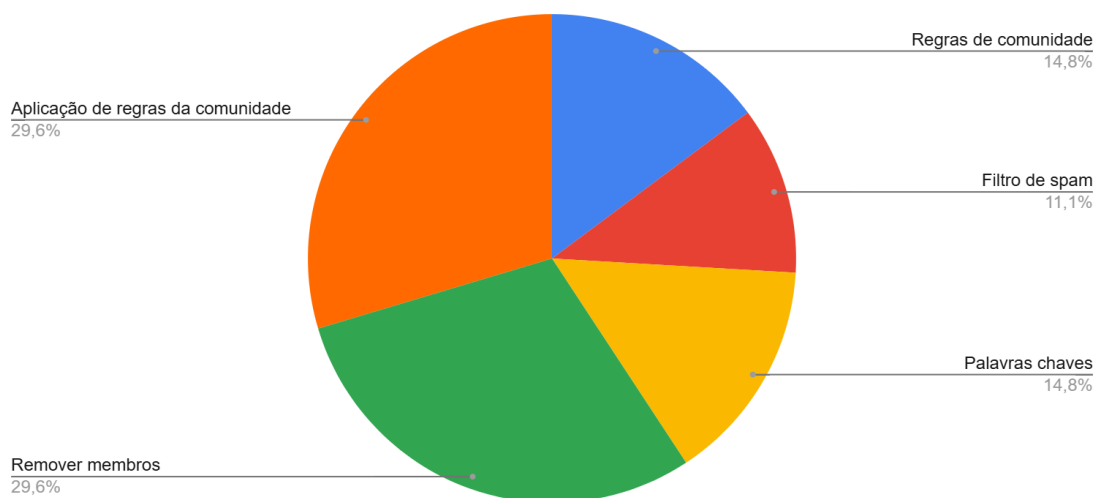


Fonte: própria do autor

No Gráfico 2, a ferramenta de denúncia é a que apresenta maior potencial de interferência nos crimes estudados acima, além disso ter a conta privada também apresenta um potencial expressivo para que avanços de pessoas desconhecidas e indesejadas sejam evitados desde o começo. As ferramentas com menor expressividade foram o Silenciar, Posts protegidos e Filtro de mensagens, por só alcançarem interferência somente em casos de Assédio e Perseguição.

O Bloquear e Conta Privada oferece proteções semelhantes se a Pessoa que utiliza a conta privada já utilizar a interação posts configurando de forma que proteja seus dados.

Gráfico 3- Porcentagem de eficácia das ferramentas de comunidade do X

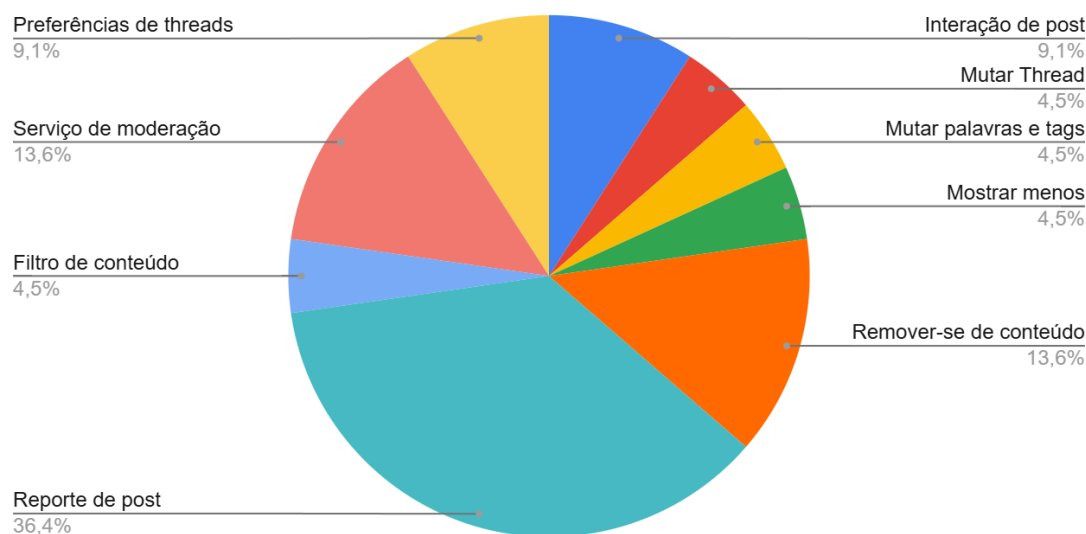


Fonte: própria do autor

O Gráfico 3 diz respeito às ferramentas inseridas dentro das comunidades do X para que a moderação e interação de usuários seja segura, a maioria dessas utilidades servem para que moderadores possam punir e aplicar as regras. A ferramenta de Regras influencia diretamente se a pessoa poderá ou não entrar e fazer parte da comunidade. Apesar disso, a ferramenta apresenta eficácia moderada, deixando que as ferramentas de moderação tenham maior peso. À aplicação de regras, como impedir que os usuários punidos sejam bloqueados de interagir ou até mesmo expulsar os membros da comunidade que cometem infrações, se mostra mais eficaz.

Gráfico 4 - Porcentagem de eficácia das ferramentas de interação do conteúdo do Bluesky

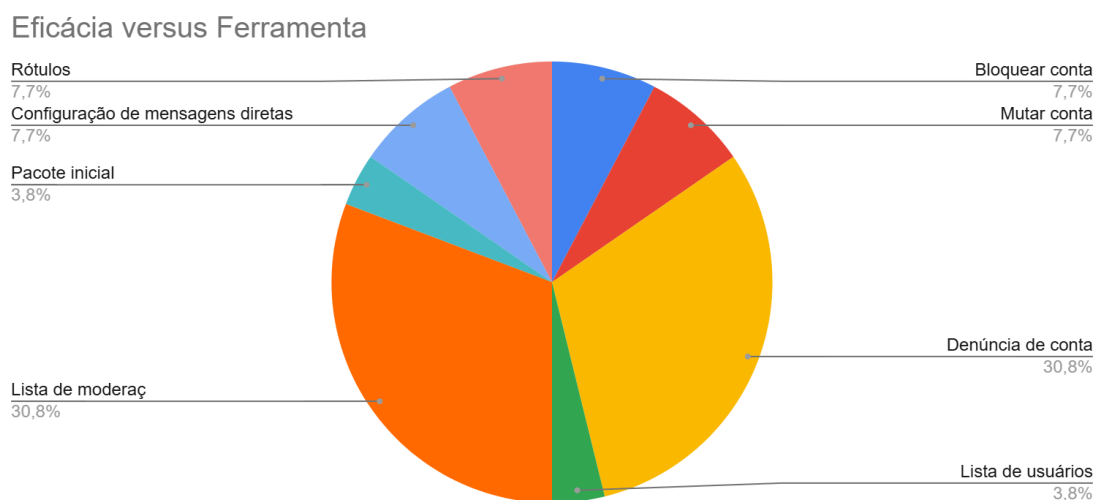
Eficácia versus Ferramenta



Fonte: própria do autor

Como é possível averiguar no Gráfico 4, se comparado com as ferramentas de interação com conteúdo do X, o Bluesky oferece mais opções de ferramentas específicas para personalizar a interação de conteúdos dos usuários. Mas apresentando um comportamento semelhante ao X (Gráfico 1), a ferramenta de maior relevância ao combater e coibir os crimes deste estudo também é a ferramenta de denúncia de Post. Além disso, no Bluesky é possível perceber que a opção de “Remover-se de conteúdo” apresenta uma boa possibilidade de intervir.

Já ferramentas como Mutar palavras e tags, mostrar menos, Mutar Thread e Filtro de Conteúdo apresentam um impacto menor, alcançando e interferindo só duas instâncias de somente dois crimes listados (Quadro 4).

Gráfico 5 - Ferramentas de interação com usuários no Bluesky

Fonte: própria do autor

A Denúncia de Contas e Lista de Moderação (Gráfico 5), apresentam quase o mesmo princípio bloquear usuários indesejados de sua rede, já a lista de usuários e pacote inicial apesar de não terem grande impacto em termos de punição são boas formas de evitar que casos de assédio e perseguição aconteçam.

Como o Bluesky não apresenta nenhuma comunidade, sendo sua forma equivalente os *feeds*, dentro desta avaliação somente três ferramentas específicas para o *Feed* foram encontradas. Das três ferramentas elencadas, somente a Denunciar um *feed*, se mostrou eficaz para combater *feeds* que apresentam conteúdos específicos.

Ao avaliar as ferramentas das duas Redes Sociais e suas usabilidades, tendo em mente os quatro crimes contra a mulher citados anteriormente, é possível discutir a efetividade de cada rede para proteção contra crimes contra mulheres. Entende-se ainda que existem crimes que não são acobertados e tratados de forma individual. Entre os crimes mencionados, o doxxing e o revenge porn foram identificados como os mais desassistidos por essas ferramentas. Esses crimes, por envolverem a divulgação de informações ou imagens sem a autorização da vítima, só podem ser punidos mediante denúncia; o simples bloqueio do infrator não possui impacto significativo, uma vez que o conteúdo permanece acessível online. Os crimes de assédio e perseguição são os crimes que melhores ferramentas de gerenciamento de comunidades dessas redes sociais, a ponto de evitar e punir no ambiente online a realização.

No contexto das ferramentas de gerenciamento de comunidades, o Assédio e perseguição, são crimes que podem receber mais interferência. Observando os documentos disponíveis de ambas as plataformas, constata-se que possuem um posicionamento semelhante quanto à punição de usuários. Contudo, apenas a rede X exige parâmetros legais rigorosos para divulgar informações necessárias à instauração de um processo criminal.

Quanto à efetividade das ferramentas de comunidade do X e *Feed* do Bluesky é possível afirmar que o X possui melhores recursos de moderação. No entanto, estes só estão disponíveis para seus moderadores, deixando que a interação de usuário com usuário seja denunciada a seus moderadores para que assim sejam aplicadas as regras da comunidade. Já os *Feeds* do Bluesky podem ser criados e moderados pelos seus criadores, porém seus moderadores não exercem controle direto sobre o que os conteúdos adicionados e não podem impedir que ninguém utilize.

Já ao falarmos em interação com conteúdos os bluesky oferece mais ferramentas para personalizar a navegação do usuário, criando um ambiente único para cada um que se registrar na plataforma, quanto a efetividade no combate e denúncia de crimes ainda deixa a desejar por não ter botões de fácil acesso denúncias para órgãos policiais. Porém ao comparar as ferramentas de denúncias contidas na plataforma com as de sua concorrente, elas têm acesso mais prático. Apesar disso, as ferramentas ligadas com bloqueio são bem mais projetadas e de fácil alcance no X.

8. CONSIDERAÇÕES FINAIS

Este estudo se propôs a fazer uma análise comparativa das ferramentas de notificação e controle comunitário em duas redes sociais, identificando e classificando ferramentas de moderação e suas utilidades diante situações de assédio online, revenge porn, perseguição e Doxxing.

Durante a pesquisa, foi constatado que alguns crimes encontram dificuldades na definição de ferramentas que possam efetivamente funcionar como punição ou como inibidor de sua ocorrência. Além disso, bloqueios e desativações de contas representam as punições mais leves que essas plataformas podem oferecer, já que cabe às autoridades competentes investigar e processar tais crimes, mediante denúncias por parte das vítimas.

Em síntese, nenhuma das redes apresentou a possibilidade de tornar-se um ambiente totalmente seguro para que mulheres utilizassem sem que fosse importunadas. O X fornece mais recursos para moderação e organização das comunidades, enquanto o Bluesky oferece ferramentas adicionais de personalização de conteúdo, proporcionando um certo grau de controle sobre o *feed* e ajudando a evitar o acesso a conteúdos indesejados. No que tange à interação entre usuários, o Bluesky se destaca por oferecer mecanismos que auxiliam na identificação de possíveis ameaças, como os rótulos de conteúdo.

Esse trabalho espera contribuir para elucidação de requisitos, futuras ferramentas e políticas de proteção às mulheres em redes sociais. As lacunas apontadas aqui não atendem toda a complexidade da realidade das interações em rede. Mas, ao menos, nos crimes mais tipificados há uma direção para melhoria apontada pelas contribuições deste trabalho.

Para uma possível escolha de rede de microblogging, as usuárias precisarão considerar a usabilidade das plataformas, avaliando quais recursos podem ser mais benéficos para suas necessidades. Com a arquitetura federada de redes, há um grande potencial para que essas políticas possam melhorar a sua aplicação e qualidade. Mas atualmente isso não acontece.

9. REFERÊNCIAS

AGÊNCIA BRASIL. Safernet registrou 74 mil denúncias de crimes de ódio na web em 2022. Disponível em: <https://agenciabrasil.ebc.com.br/radioagencia-nacional/direitos-humanos/audio/2023-02/safernet-registrou-74-mil-denuncias-de-crimes-de-odio-na-web-em-2022>. Acesso em: 07 de novembro de 2023.

BRASIL. [Constituição (1940)]. Código Penal - Decreto - Lei nº 2.848. [S. l.: s. n.], 1940.

BRASIL. DECRETO-LEI Nº 2.848 nº 2848, de 7 de dezembro de 1940. DOS CRIMES CONTRA A PESSOA e Crimes contra a honra. [S. l.], 1940.

CASTELLS, Manuel. A Sociedade em Rede. [S. l.: s. n.], 1996.

D'ANDRÉA, C. F. DE B. Pesquisando plataformas online: conceitos e métodos. [s.l.] EDUFBA, 2020.

DIJCK, José Van; POELL, Thomas; WAAL, Martijn. The Platform Society: Public values in a connective World. In: GOVERNING a responsible platform society. [S. l.: s. n.], 2018.

FIORILLO, Celso A. P.; CONTE, Christiany Pegorari. Crimes no meio ambiente digital: E a sociedade da informação. [S. l.: s. n.], 2016.

GILLESPIE, Tarleton. Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media. [S. l.: s. n.], 2018. 288 p. ISBN 978030017130.

JUSBASIL. In: Doxxing: Essa forma de exposição de dados é crime?. [S. l.], 30 mar. 2024. Disponível em: <https://www.jusbrasil.com.br/artigos/doxxing/858971038>. Acesso em: 29 mar. 2024.

JUSBASIL. In: O que é Revenge Porn ou Pornografia de Vingança e porque você deve saber como combater este tipo de ato. [S. l.], 30 mar. 2024. Disponível em: <https://www.jusbrasil.com.br/artigos/o-que-e-revenge-porn-ou-pornografia-de-vinganca-e-porque-voce-deve-saber-como-combater-este-tipo-de-ato/831302225#:~:text=O%20a%20pode%20ser%20enquadrado,respons%C3%A1vel%20por%20publicizar%20as%20imagens>. Acesso em: 29 mar. 2024.

MCLUHAN, Marshall. Os Meios de Comunicação Como Extensões do Homem: (understanding Media). [S. l.: s. n.], 1964.

NAÇÕES UNIDAS. Objetivos de Desenvolvimento Sustentável: ODS 5 - Igualdade de Gênero. 2015. Disponível em: <https://sdgs.un.org/goals/goal5>. Acesso em: 15 set. 2024.

OLIVEIRA JÚNIOR, José Oswaldo de. A SBED e a Declaração de Lima de 2022: garantia de atenção às crianças e adolescentes com dor e às suas famílias. Brazilian

Journal of Pain, v. 5, n. 3, 2022. Disponível em: <https://doi.org/10.5935/2595-0118.20220050-pt>. Acesso em: 7 nov. 2023.

OLIVEIRA, Kathleen Maria Silva de. O que é considerado um crime na internet? In: JUSBRASIL (Brasil) (org.). O que pode ser considerado crime na internet?. [S. l.], 2022. Disponível em: <https://www.jusbrasil.com.br/artigos/o-que-pode-ser-considerado-crime-na-internet/1487045194#:~:text=Todos%20aqueles%20crimes%20que%20ocorrem,danificar%20os%20dados%20dos%20usu%C3%A1rios>. Acesso em: 19 mar. 2024.

ONU. 73% das mulheres que estão conectadas já sofreram violência online. Disponível em: <https://news.un.org/pt/story/2015/09/1525411>. Acesso em: 07 de novembro de 2023.

ORIHUELA, Jose Luiz. Twitter y el boom del microblogging: Perspectivas del Mundo de la Comunicación. [S. l.: s. n.], 2007. cap. 43, p. 2-3.

PEREIRA, L. G.; DA SILVA, M. N.; SOUZA, V. P. de; REZENDE, Y. C. Hostilidade em jogos online: perspectiva feminina. Múltiplos Olhares em Ciência da Informação, Belo Horizonte, v. 7, n. 2, 2018. Disponível em: <https://periodicos.ufmg.br/index.php/moci/article/view/17033>. Acesso em: 9 maio. 2024.

PIRES, Bruna Cristina Assali. Comportamento tóxico de jogadores de League of Legends pela ótica de Winnicott. 2021. Trabalho de Conclusão de Curso (Graduação em Psicologia) - Faculdade de Ciências Humanas e da Saúde da Pontifícia Universidade Católica de São Paulo, São Paulo, 2021.

RAINIE, Lee; WELLMAN, Barry. Networked: The New Social Operating System. [S. l.: s. n.], 2012.

The Blue Sky Team. Moderation in a Public Commons. The latest from Bluesky. Disponível em: <https://blueskyweb.xyz/blog/6-23-2023-moderation-proposals>. Acesso em: 07 de novembro de 2023.

TURKLE, Sherry. Alone Together: Why We Expect More from Technology and Less from Each Other. [S. L.]: Basic Books, 2011. 384 p.