Deepfakes e nudez não consensual: o papel da IA na ampliação da violência de gênero digital

Tema: Riscos da IA; Aspectos éticos da IA; Questões relacionadas a gênero.

Proponente: Instituto de Pesquisa em Direito e Tecnologia do Recife - IP.rec (Terceiro Setor)

Palestrantes:

NOME	ORGANIZAÇÃO	SETOR	MINIBIOGRAFIA	
Vanessa Lee Araújo	Polícia Civil	Governamental	Possui graduação em Direito pela Universidade Federal do Pará (2009). Tem experiência na área de Direito, com ênfase em Direito Penal, Processual Penal e Direito Digital. É Delegada de Polícia e atualmente exerce a função de Diretora da Divisão de Prevenção e Repressão de Crimes Tecnológicos - DPRCT.	
Clarissa Mendes Gonçalves	IP.rec	Terceiro Setor	Graduada em Ciências Sociais, Mestra e Doutora em Sociologia pela UFPE. Integrante do Núcleo de Estudos e Pesquisas em Políticas de Segurança (NEPS - UFPE). Pesquisadora e Coordenadora de Projetos na área de Inteligência Artificial no IP.rec.	
Sofia Schurig	Núcleo Jornalismo	Empresarial	Repórter de tecnologia do Núcleo Jornalismo, com experiência em cobertura de direitos humanos, segurança infantil e extremismo online. Pesquisadora na Safernet Brasil.	
Ximena Cuzcano	Situada.Online	Comunidade Científica e Tecnológica	Engenheira de Sistemas e Mestre em Cibersegurança. É pesquisadora na Situada.Online. Seu trabalho examina os desafios emergentes no espaço digital, promovendo a convergência entre tecnologia e valores éticos para um uso responsável e equitativo da tecnologia na sociedade.	









Moderador(a):

NOME	ORGANIZAÇÃO	SETOR	MINIBIOGRAFIA
Carolina Gomes Pugliesi Branco	IP.rec	Terceiro Setor	Graduanda em Direito pela Universidade Federal de Pernambuco (UFPE). Atua como pesquisadora no IP.rec, atua nas áreas de Inteligência Artificial e Regulação de Plataformas.

Relator(a):

NOME	ORGANIZAÇÃO	SETOR	MINIBIOGRAFIA
Luana Batista Araújo	IP.rec	Terceiro Setor	Graduada em Engenharia da Computação pela Unifesspa (Universidade Federal do Sul e Sudeste do Pará). Atua como pesquisadora no IP.rec (Instituto de Pesquisa em Direito e Tecnologia do Recife) nas áreas de Inteligência Artificial e Criptografia. Bolsista do Programa Youth Brasil 2023 (CGI.br/NIC.br).

Estruturação do workshop

Objetivos (propostos e atingidos): O presente workshop teve como objetivo compreender o fenômeno das deepfakes pornográficas, enquanto manifestação da violência de gênero, no contexto latino americano, mapeando tendências comuns e particularidades. Discutiu-se, com uma abordagem multissetorial, como a tecnologia vem sendo utilizada para atacar a dignidade sexual de mulheres e meninas na região, destacando as dimensões: a) geracionais, uma vez que tem se tornado comum episódios dessa natureza em ambiente escolar e b) políticas, já que as deepfakes têm sido usadas como forma de descredibilizar candidatas a cargos públicos. Por fim, o painel levantou possíveis estratégias para enfrentar o problema na região, adentrando em iniciativas que tenham o intuito de combater esse fenômeno, com foco especial no Brasil, que tem, ao menos, 3 projetos de leis principais que tratam a questão: o PL 5394/23 da Deputada Erika Kokay (PT/DF), o PL 5467/23 da Deputada Camila Jara (PT/MS) e o PL 5342/23 do Deputado Marcelo Álvaro Antônio (PL/MG).

Resultados (propostos e atingidos): O workshop delineou o problema das deepfakes não-consensuais contendo nudez explícita de mulheres, assim, refletindo sobre uma das manifestações mais frequentes de ciberviolência de gênero da atualidade. Nesse sentido, foram abordados, de forma mais detida, as ocorrências em escolas, assim como os casos dirigidos a









mulheres na esfera pública, considerando as especificidades desses dois contextos de vitimização. A partir das discussões realizadas, foram alcançados os seguintes resultados: (1) Caracterização e compreensão dos casos mais frequentes de deepnudes, contribuindo para o aprofundamento do conhecimento sobre o tema e suas diferentes manifestações; (2) Promoção de um diálogo intersetorial e regional, conectando experiências e ocorrências observadas tanto no Brasil quanto em outros países da América Latina; (3) Produção de insumos para a formulação de políticas públicas de prevenção, com foco em estratégias eficazes de contenção e educação digital; (4) Oferecimento de recomendações para responder adequadamente ao problema uma vez que ele ocorre, focando na perspectiva da proteção às vítimas.

Justificativa em relação à governança da Internet: Em 2017, um vídeo pornográfico falso envolvendo a atriz Gal Gadot foi publicado por um usuário do Reddit chamado de "deepfakes", com isso popularizou-se o uso do termo. A expressão passou a ser utilizada para se referir a qualquer imagem, áudio ou vídeo falso produzido por computador que usam modelos de aprendizagem de máquina. Como especialistas vêm apontando, essa tecnologia tem sido utilizada para atacar, em especial, mulheres e meninas, não se limitando a apenas atrizes e cantoras famosas, mas afetando qualquer uma que tenha fotos compartilhadas na Internet. Nesse contexto, tem sido cada vez mais comum notícias sobre o uso não consentido de imagens para criar deepfakes pornográficas. A título de exemplo, alunos de uma escola em Chorrilhos (Peru) alteraram fotos de suas colegas para vendê-las. No México, investiga-se um estudante por alterar mais 50 mil fotos de alunas do Instituto Politécnico Nacional para vendê-las. Casos semelhantes a esses também foram registrados nas cidades de Recife, Rio de Janeiro e Belo Horizonte. Além desses incidentes, têm sido registrados também ataques como esses a candidatas ao pleito municipal de 2024 no Brasil. Esse tipo de conteúdo pode produzir consequências devastadoras na vida das vítimas, afetando desde de sua integridade física, moral e psicológica até seu emprego e relações familiares. Além disso, o tema aparece como um desafio para pensar aspectos que descambam em novos processos regulatórios da Internet em todo o mundo. Quais as implicações o tema traz para os níveis de aplicação e conexão dos provedores? Como combater e educar para um uso saudável da tecnologia na Internet? Existem contextos ilegítimos a priori, exigindo soluções de curadoria de conteúdo automatizada? Essas são só algumas das reflexões que demonstram como o tema pode afetar a governança multissetorial da Internet, devendo ser debatido de forma aberta e com seriedade, de modo a tornar a internet um espaço seguro e ético para todas as pessoas.

Metodologia e formas de participação desenvolvidas durante a atividade: A metodologia adotada envolveu uma abordagem multissetorial, promovendo a escuta de representantes do setor governamental, empresarial, científico-tecnológico e do terceiro setor. Cada painelista foi estimulado a expor sua perspectiva específica: a representante do governo apresentou experiências práticas sobre a atuação do Estado frente à disseminação de deepnudes e os mecanismos de proteção às vítimas; o setor empresarial e científico discutiu os desafios técnicos e as limitações das ferramentas de detecção e contenção; e o terceiro setor destacou as raízes socioculturais da violência de gênero digital, defendendo abordagens centradas na proteção e na educação. A dinâmica do painel seguiu a seguinte estrutura: introdução da moderação (5 minutos), exposições de cada painelista (10 minutos cada), provocações pela moderação (5







minutos), e respostas dos painelistas (3 minutos cada), encerrando com 25 minutos destinados à interação com o público. A participação da audiência foi estimulada antes e durante a atividade. A mobilização prévia nas redes sociais das entidades envolvidas buscou engajar o público por meio de conteúdos informativos. Durante o workshop, os participantes puderam interagir pelo aplicativo Mentimeter, acessível por QR Code, respondendo a enquetes em tempo real que alimentaram o debate com percepções diretas da audiência. Participantes presenciais também puderam se manifestar com o uso de microfones, enquanto o público remoto utilizou o chat da transmissão ao vivo para contribuir com questionamentos e comentários.

Síntese dos debates

O painel, mediado por **Carolina Branco**, concentrou-se na análise do uso de tecnologias de inteligência artificial deepfakes como mecanismo de perpetuação da violência de gênero, com ênfase na criação e disseminação de conteúdos de nudez não consensuais. Na abertura da sessão, a moderadora apresentou dados de 2023 indicando que 98% das deepfakes então identificadas apresentavam conteúdo de natureza sexual, sendo 99% das vítimas mulheres. Esses números evidenciam o caráter estruturalmente misógino da tecnologia em questão, que opera como uma atualização digital de formas históricas de opressão.

Carolina destacou que, embora essas imagens sejam sintéticas, seus efeitos são concretos: impactam a saúde mental, a integridade psicossocial, a vida pública e as relações profissionais e familiares das vítimas. Nesse contexto, mencionou o surgimento, entre 2023 e 2025, de ao menos oito proposições legislativas voltadas ao enfrentamento desse tipo de violência, com destaque para a sanção da Lei 15.123/2025, oriunda do Projeto de Lei 370/2024 da deputada Jandira Feghali. Essa norma alterou o Código Penal para agravar a pena do crime de violência psicológica contra a mulher quando praticado com uso de tecnologias de manipulação de imagem ou voz.

Ainda que reconhecendo a importância dessas iniciativas, a mediadora chamou atenção para um padrão: todas priorizam a via penal. Embora o direito penal tenha valor simbólico e capacidade de coerção, Carolina observou que sua aplicação deveria ocorrer em última instância (ultima ratio), e não como estratégia central. Enfatizou, nesse sentido, que uma resposta eficaz ao problema exige ações articuladas que vão além da criminalização, como o fortalecimento da perícia digital, a remoção célere de conteúdos, o suporte jurídico e psicológico às vítimas, a responsabilização civil das plataformas e, sobretudo, a promoção da educação digital crítica.

Por fim, concluiu sua introdução destacando que o enfrentamento da violência algorítmica de gênero demanda políticas públicas intersetoriais que promovam a dignidade, a autonomia e a presença plena de mulheres e meninas em todos os espaços sociais, incluindo os de representação política.

Na sequência, tiveram início as exposições das painelistas, organizadas a partir de perguntas orientadoras formuladas pela mediação.







A primeira intervenção foi realizada por **Clarissa Mendes**, representante do terceiro setor. Sua fala foi guiada pela seguinte pergunta: "Como as iniciativas do Terceiro Setor ajudam a iluminar as raízes socioculturais profundas, como a misoginia e a objetificação da mulher, que impulsionam a criação e disseminação de deepfakes com nudez não consensual, particularmente nos contextos geracionais (ambiente escolar) e políticos (contra candidatas)? Qual o papel essencial da educação digital e das ações da sociedade civil em complementar (ou mesmo pautar) as respostas legais e tecnológicas para combater esse fenômeno e suas raízes?"

Clarissa iniciou abordando as limitações metodológicas de algumas pesquisas sobre o tema, que frequentemente negligenciam marcadores interseccionais como raça, identidade de gênero e classe. Apesar disso, afirmou que os dados disponíveis já permitem identificar com clareza que o gênero é um elemento central na configuração dessa violência. Apontou, por exemplo, uma pesquisa que analisou resultados de busca no YouTube e constatou que 100% dos vídeos classificados como deepfakes envolvendo conteúdo sexual tinham mulheres. Além disso, destacou que tanto o desenvolvimento quanto o consumo dessas tecnologias são predominantemente realizados por homens, o que reforça a sua base estruturalmente misógina.

Ao discutir o caso do aplicativo DeepNude, Clarissa chamou atenção para o fato de que a ferramenta foi treinada exclusivamente com imagens de corpos femininos cisgênero, não operando com imagens masculinas. Tal configuração revela que a misoginia está presente não apenas no uso, mas na própria concepção técnica da tecnologia. Segundo ela, trata-se de tecnologias que operam dentro de um ecossistema predatório, que ignora o consentimento e monetiza a violência de gênero digital. Ela observou que figuras públicas, como atrizes, cantoras e candidatas políticas, estão particularmente expostas a esse tipo de ataque, embora disponham de maior capacidade de acessar mecanismos de remoção de conteúdo. Já mulheres anônimas, sobretudo em contextos de violência doméstica ou em espaços historicamente masculinizados, enfrentam grandes obstáculos.

Clarissa também chamou atenção para a disseminação desse tipo de prática em ambientes escolares, envolvendo adolescentes, o que revela a precocidade com que a cultura da objetificação feminina é reproduzida e, por ultimo, finalizou afirmando que o enfrentamento das deepfakes deve ir além de soluções técnicas ou penais e ser orientado por princípios de proteção, educação e transformação cultural.

A segunda fala foi de **Ximena Cuzcano**, pesquisadora do campo científico e tecnológico, à qual foi dirigida a seguinte pergunta: "Quais são os avanços tecnológicos mais promissores e as limitações atuais das ferramentas de Inteligência Artificial para a detecção e autenticação de deepfakes pornográficas, especialmente no contexto da América Latina, onde a disseminação desses conteúdos tem apresentado particularidades? Como a comunidade científica e tecnológica, a partir de sua experiência, pode agir de forma mais ética, transparente e colaborativa para desenvolver e implementar soluções técnicas (ferramentas de detecção, mecanismos de prevenção) que efetivamente mitiguem a criação e disseminação desses conteúdos, ao mesmo tempo em que promovem um uso mais responsável e seguro da tecnologia para todas as pessoas?"





Ximena iniciou sua contribuição reafirmando que tecnologias digitais não são neutras e reproduzem os valores e desequilíbrios sociais do contexto em que são desenvolvidas. Ela apontou que essas tecnologias, quando inseridas em contextos de profunda desigualdade estrutural, como o latino-americano, tendem a reproduzir e intensificar essas assimetrias.

Em seguida, a pesquisadora explicou brevemente os fundamentos técnicos das deepfakes, com foco nas redes neurais adversariais (GANs), que conferem realismo elevado a vídeos e imagens falsificados.

Apesar dos avanços na área de detecção, a palestrante observou que os criadores de deepfakes também acompanham essas inovações e atualizam seus métodos para evitar a identificação de seus conteúdos, o que gera um ciclo de evasão contínuo. Além disso, apontou diversas limitações nas tecnologias atuais: os detectores são treinados com bases de dados eurocentradas, não são eficazes na identificação de conteúdos sexualizados e não operam em plataformas de circulação amplamente utilizadas como WhatsApp e Telegram.

Com base em sua pesquisa sobre o ecossistema de aplicativos de deepfake na América Latina, Ximena relatou que muitos deles são amplamente acessíveis, inclusive a adolescentes, e frequentemente utilizam imagens sexualizadas como estratégia de atração. Apontou ainda a predominância de homens em cargos técnicos nas empresas que desenvolvem essas tecnologias, o que reforça a ausência de diversidade no processo de criação. A seu ver, não basta adaptar ferramentas técnicas, é preciso repensar os próprios fundamentos ideológicos dessas tecnologias, e concluiu defendendo uma abordagem ética e interdisciplinar, com regulação efetiva e participação ativa de mulheres no desenvolvimento de soluções.

Na terceira exposição, **Sofia Schurig**, representante do setor empresarial, respondeu à pergunta: "Quais são os casos mais relevantes e emblemáticos de deepfakes pornográficas não consensuais que você tem encontrado? A partir desses exemplos, como avaliar a capacidade atual das grandes plataformas de internet em detectar e remover esse tipo de conteúdo de forma eficaz e proativa? E quais responsabilidades e ações concretas essas empresas do setor empresarial precisam assumir com urgência para melhorar suas políticas, ferramentas tecnológicas e processos de moderação, contribuindo para mitigar a disseminação dessa forma de violência de gênero digital em larga escala?"

Sofia compartilhou experiências de monitoramento em redes sociais que demonstram a ineficácia dos sistemas atuais de moderação. Apresentou, por exemplo, casos envolvendo perfis no Instagram e grupos no Telegram que promoviam conteúdos de erotização infantil gerados por IA, os quais permaneceram ativos mesmo após diversas denúncias, sendo removidos apenas após intervenção da imprensa. Ressaltou que as plataformas mantêm uma postura predominantemente reativa, com processos de moderação automatizados e cada vez mais desprovidos de profissionais especializados em gênero, linguagem e contexto local.





Além disso, abordou o caso do modelo Stable Diffusion 1.5, treinado com imagens reais de abuso infantil, cujo uso permanece irrestrito devido à natureza aberta de seu código e base de dados. Argumentou que a responsabilização das empresas deve incluir não apenas o conteúdo final, mas também decisões estruturais relacionadas ao treinamento de modelos e à curadoria de dados e destacou que a ética digital não pode ser vista como um diferencial, mas sim como um princípio mínimo para a atuação de qualquer empresa em ambientes digitais.

Encerrando a rodada de exposições, **Vanessa Lee Araújo**, representante do setor governamental, respondeu à pergunta: "Qual é o principal gargalo prático e jurídico que a DPRCT e outras unidades policiais no Brasil enfrentam hoje ao investigar e responsabilizar os autores de deepfakes com nudez não consensual, especialmente considerando a velocidade de disseminação e as dificuldades de rastreamento em meio digital? De que forma os procedimentos de exclusão de conteúdo e proteção às vítimas, esperados do setor governamental, podem ser agilizados e fortalecidos para mitigar os impactos devastadores sobre a vida das mulheres e meninas afetadas?"

Vanessa reconheceu que, embora o direito penal possa ser necessário em determinados casos, sua efetividade depende de uma estrutura institucional que hoje é, em grande medida, insuficiente. Relatou dificuldades enfrentadas pelas autoridades na obtenção de dados junto às plataformas, que muitas vezes se eximem de cooperar sob o argumento da proteção à liberdade de expressão. Destacou ainda entraves técnicos, como a permanência do protocolo IPV4, que dificulta o rastreamento de autores de crimes digitais.

A representante destacou a crescente ocorrência de ataques a pesquisadoras e candidatas em universidades e no campo político, reforçando o uso da IA como instrumento de silenciamento e deslegitimação de mulheres em posições de liderança.

Por fim, Vanessa enfatizou a necessidade urgente de regulamentação específica e de maior articulação entre os setores público e privado para garantir respostas mais eficazes.

Ao final do painel, foram apresentados os resultados das enquetes realizadas com o público por meio da plataforma Mentimeter. Na primeira, 98% das pessoas participantes afirmaram que as plataformas deveriam ser obrigadas a remover conteúdos de deepfakes com nudez não consensual imediatamente, mesmo diante do risco de eventual censura indevida. Na segunda, 69% responderam que a distinção entre imagens reais e sintéticas não altera o impacto da violência. As debatedoras concordaram, com diferentes nuances: Clarissa destacou que a distinção pode importar para as vítimas e para a análise do fenômeno, mas que os efeitos são igualmente danosos; Ximena e Vanessa reforçaram que, independentemente da origem da imagem, o conteúdo é invasivo e tem impacto significativo; e Sofia criticou o uso do argumento da censura indevida pelas empresas, apontando a necessidade de se aplicar a mesma inteligência usada para fins comerciais à proteção de direitos fundamentais.

Ao final das exposições, foi realizado um bloco de perguntas e comentários abertos à audiência presencial e remota, com participação ativa do público por meio do YouTube e intervenções









presenciais. A rodada permitiu aprofundar temas como responsabilização das plataformas, regulação, interseccionalidade, ética no desenvolvimento tecnológico e o papel da educação na prevenção da violência de gênero digital.

Entre as perguntas enviadas pelo público remoto, Aline Rocha questionou se o controle social dos corpos femininos por meio da violência teria migrado irreversivelmente para o ambiente digital. Já Aline Costa Simões perguntou sobre a possibilidade de responsabilização jurídica das plataformas por conteúdos abusivos e a existência de mecanismos legais, especialmente quanto a danos morais. Em resposta, Vanessa Lee Araújo afirmou que o Marco Civil da Internet já prevê a possibilidade de remoção de conteúdo mediante notificação, mas apontou obstáculos significativos à sua efetividade, especialmente quando as plataformas estão sediadas em outros países. Ressaltou ainda a necessidade de cooperação internacional e destacou que, embora o Brasil seja referência normativa na área, a aplicação prática desses dispositivos ainda enfrenta desafios.

No bloco presencial, Larissa Santiago trouxe a questão da coleta de dados sobre raça, gênero e expressões religiosas nas delegacias especializadas. Vanessa respondeu que, apesar de ainda inexistente de forma estruturada, há esforços pontuais, como em sua própria unidade, para desenvolver uma taxonomia mais sensível às intersecções.

Luma, profissional da área técnica, apontou a ausência de formação ética durante a graduação, especialmente no que diz respeito a temas como racismo, gênero e impactos ambientais. Sofia Schurig respondeu reforçando a importância da formação crítica de profissionais da tecnologia, destacando a centralidade da educação nesse processo.

André Fernandes questionou a fragmentação dos projetos de lei relacionados à regulação de plataformas e inteligência artificial, mencionando proposições como os PLs 2630, 2628 e 2338. Indagou sobre os riscos de retrocessos e apagamentos no processo regulatório.

Isabela perguntou se, na perspectiva das convidadas, ainda é possível apostar na educação como ferramenta para mitigar riscos e promover o uso ético da tecnologia. Em resposta, todas as painelistas reafirmaram o papel da educação digital crítica como eixo transversal e indispensável. Vanessa mencionou o programa "Guardiões Digitais", que atua em escolas abordando temas como ética, racismo, sexismo e segurança online com crianças, responsáveis e educadores.

Mari Leal enfatizou que o sexismo está na base das violências analisadas e perguntou como cada setor pode contribuir para uma transformação social efetiva. As respostas convergiram na importância de políticas estruturantes e da atuação conjunta entre Estado, sociedade civil e empresas para promover mudanças culturais duradouras.

Por fim, Elizabeth questionou a experiência concreta das painelistas no contato com as plataformas para remoção de conteúdos. Clarissa Mendes afirmou que as plataformas continuam sendo uma via central de acesso e disseminação dos deepfakes, e que enfrentá-las é parte indissociável da resposta à violência. Destacou ainda que o problema deve ser compreendido de









forma estrutural, articulando ações de educação, letramento digital, apoio às vítimas e pressão pública. Sofia complementou que, para que as plataformas mudem, seria necessário redesenhar seus modelos de negócio, o que, em sua visão, não é uma disposição real das empresas atualmente.

A rodada demonstrou a complexidade do tema e a necessidade de abordagens interdisciplinares e intersetoriais, com forte articulação entre educação, legislação, governança digital e transformação cultural. O painel foi encerrado pela mediadora com agradecimentos às participantes e ao público.

Identificação de consensos, dissensos e pontos a aprofundar:

TIPO DE	CONTEÚDO	CONSENSO	PONTOS A
MANIFESTAÇÃO Posicionamento	As deepfakes com nudez não consensual são formas contemporâneas de violência	OU DISSENSO Consenso	APROFUNDAR Considerar marcadores interseccionais como gênero, raça, deficiência,
	de gênero com impactos concretos.		sexualidade e classe social.
Proposta	Necessidade de políticas públicas intersetoriais e não apenas de respostas penais.	Consenso	Como integrar de forma eficaz essas políticas entre setores.
Proposta	Fortalecer mecanismos de responsabilização civil das plataformas e revisar seus modelos de negócios que priorizam lucro sobre segurança.	Consenso	Viabilidade política e técnica de uma regulação mais incisiva sobre o modelo de negócio das big techs.
Posicionamento	As tecnologias de IA não são neutras e refletem desigualdades sociais e de gênero no modo como são desenvolvidas e utilizadas.	Consenso	Adoção de diretrizes éticas e equipes diversas no desenvolvimento tecnológico.
Posicionamento	A educação digital crítica é elemento essencial na prevenção e combate à violência de gênero digital.	Consenso	Como incluir de forma estruturada esse conteúdo nos currículos escolares e na formação técnica em tecnologia.
Proposta	Ampliar a coleta e uso de dados interseccionais nas delegacias e serviços públicos para compreender melhor o perfil das vítimas.	Consenso	Padronização e institucionalização da coleta desses dados nos sistemas de segurança pública.







