UNIVERSIDADE FERNANDO PESSOA

Quando a Tecnologia Viola a Intimidade: O Caso dos Deepfakes Sexuais

[When Technology Violates Privacy: The Case of Sexual Deepfakes]

Projeto de Graduação

1o Ciclo de Criminologia

Inês Isabel da Silva Sales 2022122003

Trabalho efetuado sobre orientação da

Professora Doutora Ana Sani

2024/2025

Quando a Tecnologia Viola a Intimidade: O Caso dos Deepfakes Sexuais

[When Technology Violates Privacy: The Case of Sexual Deepfakes]

Projeto de Graduação

1o Ciclo de Criminologia

Inês Isabel da Silva Sales 2022122003

Trabalho efetuado sobre orientação da

Professora Doutora Ana Sani

2024/2025

“Do one thing every day that scares you."

Eleanor Roosevelt

Agradecimentos

À minha orientadora, Professora Doutora Ana Sani, expresso a minha profunda gratidão pela orientação ao longo desta última etapa da licenciatura e pelo apoio e suporte oferecido durante a execução do meu projeto de graduação.

À Associação Portuguesa de Apoio à Vítima, em especial, à Dra. Marlene Fonseca, por me ter proporcionado uma excelente primeira experiência profissional na área e por ter aumentado o meu crescente interesse pela vitimologia.

À minha família, o meu pilar inabalável. Em especial à minha mãe, Carmo, a mulher mais forte e resiliente que conheço, que sempre me ensinou a nunca deixar que me derrubassem. Aos meus irmãos, Leandro, Patrícia e Zé Luís, por me incentivarem a correr atrás do que eu gosto e apoiarem sempre as minhas decisões. Agradeço-vos por me terem moldado na pessoa que sou hoje e por me proporcionarem sempre um lugar a que chamo Casa. Vocês são a minha força. Ao Amândio por ter acreditado na minha capacidade como aluna e futura profissional, e por me ter apoiado na concretização desta etapa tão especial da minha vida. A vocês, mil obrigados.

Aos meus companheiros de curso e amigos para a vida, Eva, Lara e Bernardo, que tornaram esta jornada muito mais leve e divertida. Obrigada por cada risada partilhada, por cada pequeno-almoço a estudar, por cada noite de estudo e por nunca me deixarem desistir. Levarei sempre a vossa amizade no coração.

Ao Zé, o meu porto seguro. Obrigada por cada abraço e palavras nos momentos de desespero e por nunca largares a minha mão, mesmo quando o caminho parecia impossível. Amo-te e agradeço por me amares e por me ajudares a alcançar o meu sonho. Obrigada por acreditares em mim quando eu própria duvidei das minhas capacidades.

Por último, mas do fundo do meu coração, às estrelinhas mais brilhantes do céu, o meu pai, António, e o meu avô, José. Sei que onde quer que estejam, têm acompanhado cada passo meu com orgulho e lágrimas nos olhos. Foram vocês que me deram força para continuar. Esta vitória também é vossa.

Resumo

O presente projeto de graduação, visa em primeiro lugar definir e contextualizar os deepfakes sexuais não consentidos (DSNC) do ponto de vista histórico, legal e social, tendo como proposta de estudo investigar os discursos, motivações e racionalizações de indivíduos que produzem ou divulgam DSNC em comunidades digitais. A DSNC é estudada como uma nova forma de violência de género baseada em imagem.

Atualmente, a disseminação dos DSNC é alarmante. Entre 2019 e 2023, registou-se um aumento de 550% no número de vídeos manipulados disponíveis online, alojados em mais de 100 sites especializados, sendo entre 96% e 98% de todos os deepfakes identificados de teor sexual (Home Security Heroes, 2023).

A partir do enquadramento teórico foi orientada a proposta de estudo. Esta proposta, entre junho de 2025 até junho de 2027, é voltada a compreender discursos, motivações e dinâmicas de comunidades que produzem e compartilham DSNC, a partir de instrumentos como inquéritos e entrevistas, para fundamentar intervenções informadas e eficazes.

Além disso, os resultados antecipados permitiram delinear perfis sociodemográficos dos autores de DSNC, mapear as técnicas de neutralização moral utilizadas para justificar a violação de imagens íntimas e identificar os mecanismos comunitários que reforçam e normalizam estas práticas.

Ademais, o trabalho pretende não só iluminar quem produz e difunde DSNC e porquê, mas também procura colmatar a lacuna científica existente sobre os perpetradores de DSNC e contribuir para políticas públicas de prevenção, legislação digital e proteção das vítimas.

Palavras-chave: Deepfake, Inteligência Artificial, Pornografia, Abuso Sexual Baseado em Imagens

Abstract

The present graduation project first seeks to define and contextualise non-consensual sexual deepfakes (NCSF) from historical, legal and social perspectives. Its empirical component is designed to investigate the discourses, motivations and rationalisations of individuals who produce or disseminate NCSF in digital communities. In this regard, NCSF are treated as a new, image-based form of gender-based violence.

Today, the spread of NCSF is alarming. Between 2019 and 2023, the number of manipulated videos available online grew by 550 %, hosted on more than 100 specialised sites; 96–98 % of all detected deepfakes are sexual in nature (Home Security Heroes, 2023).

Guided by this theoretical framework, the study proposal, between June of 2025 and June of 2027, aims to understand the discourses, motivations and community dynamics behind the production and sharing of NCSF. It will rely on online surveys and semistructured interviews to generate evidence-based recommendations for effective interventions.

The anticipated results include outlining the perpetrators' sociodemographic profiles, mapping the moral-neutralisation techniques used to justify the violation of intimate images and identifying the community mechanisms that reinforce and normalise these practices.

Ultimately, the project intends not only to shed light on who produces and spreads NCSF and why, but also to fill the scientific gap regarding NCSF perpetrators and to inform public policies on prevention, digital legislation and victim protection.

Keywords: Deepfake, Artificial Intelligence, Pornography, Image-Based Sexual Abuse

Índice

INTRODUÇÃO

O presente Projeto de Graduação, nomeado de "Quando a Tecnologia Viola a Intimidade: O Caso dos Deepfakes Sexuais" constitui num dos requisitos necessários para a obtenção do grau de Licenciatura em Criminologia, pela Universidade Fernando Pessoa.

A escolha deste tema fundamenta-se na rápida evolução dos deepfakes em específico dos deepfakes sexuais não consensuais (DSNC), bem como no interesse em perceber os motivos e os perfis envolvidos na criação, consumo e divulgação destes materiais.

Vivemos numa era marcada por avanços tecnológicos sem precedentes, em que a linha entre o real e o fabricado se torna cada vez mais ténue e difícil de traçar. Para corroborar esta ideia o Supremo Tribunal Federal (2024) brasileiro afirma que "Algumas deepfakes são tão sofisticadas que o desmascaramento exige uma análise profissional, com o apoio de técnicas e ferramentas de perícia." (p.11).

Os DSNC representam uma forma emergente de violência digital em que a Inteligência Artificial (IA) é utilizada para sobrepor rostos, frequentemente de mulheres , em vídeos pornográficos sem autorização. O fenómeno é particularmente relevante para a criminologia, uma vez que permite aprofundar a compreensão de novas formas de violência de género contra as mulheres e explorar a ineficiência ou ausência de respostas institucionais e a impunidade dos perpetradores desta. De acordo com alguns investigadores, o avanço da IA, reduziu drasticamente as barreiras técnicas, permitindo que utilizadores sem qualquer tipo de qualificação técnica produzam vídeos ou imagens com um elevado grau de realismo em minutos, em um computador doméstico (Chesney & Citron, 2019; Karaboga et al., 2024).

Portanto, o objetivo principal deste trabalho é analisar a relação entre a IA e a delinquência, com especial enfoque na criação e disseminação de DSNC e na forma como a escassa literacia digital e aceitabilidade social deste conteúdo limitam o reconhecimento público do dano e dificultam a atuação penal e preventiva. Neste sentido, torna-se fundamental estudar os impactos desta nova realidade digital, de modo a compreender um fenómeno tão complexo como atual, para futuramente desenvolver planos de prevenção ou intervenção eficazes.

Relativamente à estrutura do presente trabalho, este organiza-se em dois capítulos principais. Em primeiro lugar, é apresentado o primeiro capítulo que aborda os conceitos fundamentais relacionados com a IA e deepfakes, com principal enfoque nos DSNC. Nomeadamente, é abordado o processo de criação, impactos, perfis de vítimas e perpetradores, métodos de combate e deteção, entre outros. Em segundo lugar, é apresentado o segundo capítulo, que expõe os elementos metodológicos da investigação, incluindo a proposta de estudo, a amostra, os instrumentos, os procedimentos e os resultados esperados. Por último, é exibida a conclusão, as referências bibliográficas que sustentam o desenvolvimento teórico e metodológico do trabalho, e os anexos.

CAPÍTULO I – ENQUADRAMENTO TEÓRICO

1.1. Inteligência Artificial (IA)

Ao longo da história, a humanidade sempre buscou ferramentas que facilitassem o dia a dia. Esse fenómeno acontece desde a invenção da roda até a revolução industrial, nos quais o progresso foi moldado por necessidades práticas e pelo desejo constante de superação de limites. Com o passar dos séculos, o desenvolvimento de máquinas, motores e sistemas automatizados transformou profundamente a sociedade e as suas dinâmicas sociais, económicas e laborais (Oliveira & Barroco, 2023).

Na realidade, a IA mesmo antes de ser o que é atualmente já era algo abordado e fantasiado em filmes, bandas desenhadas e séries de ficção científica. Há décadas, que o ser humano procura ensinar máquinas a raciocinar como ele próprio, e foi a partir da década de 50 que ideia da IA começou a ser estudada cientificamente. O primeiro artigo a discutir essa possibilidade foi escrito por Alan Turing, em 1950. Seis anos depois, durante a conferência de Darthmouth, John McCartthy, deu o nome de Inteligência Artificial a essa tecnologia, sendo este o marco inicial da área (Bhutani & Sanaria, 2023).

Foi a partir dos anos 2000, com o aumento do poder computacional, disponibilidade massiva de dados e avanços em algoritmos de aprendizado de máquina, que a IA evoluiu significativamente (Jiang et al., 2022). De facto, muitas das tarefas que antes eram realizadas por humanos estão a ser substituídas pela robótica. Como em todas as revoluções, estas mudanças transformam expressivamente a forma como enfrentamos os desafios (Okolie, 2023).

Alguns autores realçam que a IA é algo complexo e amplo, e, por isso, existem ainda inúmeras definições para o termo. Além disso, certas profissões sentem-se intrigadas pelas definições dadas ao fenómeno, sendo, desta forma, uma definição não consensual ou única (Angeli et al., 2019; Vieira, 2024). Por exemplo, a IA, de acordo com João Maia (2018) é a criação de uma inteligência que iguala ou até mesmo supera o intelecto humano em algumas características.

Dito isto, esta é uma realidade presente no nosso quotidiano, e a sociedade está cada vez mais dependente desse domínio. É importante realçar que a IA está integrada em dispositivos como a Siri, a Alexa e até mesmo em serviços streaming como Spotify e

Netflix – e tudo indica que esta se tornará a normalidade do nosso dia a dia (Morais & Castelo Branco, s.d.).

Além disso, esta ferramenta tem desempenhado um papel importante em vários setores da sociedade e no quotidiano das pessoas, mas esta pode ser relacionada com a expressão "faca de dois gumes" simbolizando a ideia de que a IA não apresenta apenas impactos benéficos, mas também levanta questões importantes que precisam de ser estudadas e controladas (Danry et al., 2022).

Por um lado, esta nova tecnologia está ligada a inúmeros benefícios, muitos dos quais passam despercebidos. Ademais, Angeli et al. (2019) realça que a IA oferece "novas soluções a cada dia, agindo neste meio com cada vez mais efetividade, e impulsionando as empresas a diminuírem seus custos e aumentarem a qualidade de seus produtos." (p.11). Outros autores identificaram benefícios em áreas como medicina, justiça, apoio militar, entre outros (Vieira, 2024).

Por outro lado, inúmeros especialistas demonstram preocupação com o rumo que esta tecnologia pode tomar, especialmente questões relacionadas com os direitos humanos, à liberdade pessoal, ao desemprego, aos direitos de autor e ao risco de guerras cibernéticas (Maia, 2018). Assim, esta ferramenta também pode ser utilizada para fins menos éticos, como o cibercrime, a violação da privacidade e autonomia das pessoas, a perseguição cibernética e o bullying. Estas práticas podem resultar em perdas financeiras, danos à reputação e, em casos extremos, perda de vidas (Burrel & Fourcade, 2021; Januário, 2024). Por conseguinte, esta quando utilizada em excesso ou mal utilizada também pode proporcionar riscos e problemas (Floridi, 2018). Segundo Moreira & Ribeiro (2024):

> É preciso, por exemplo, compreender como a inteligência artificial opera e como ela pode ser influenciada por vieses culturais, econômicos e políticos. As pessoas também precisam desenvolver habilidades para trabalhar com a inteligência artificial e adaptar-se às mudanças rápidas que ocorrem na sociedade da informação (p.8).

Dessa forma, torna-se evidente que o desafio não reside no desenvolvimento tecnológico, mas também na capacidade da sociedade de acompanhar esse avanço de forma crítica, ética e consciente. Contudo, é urgente capacitar as pessoas para que desenvolvam competências que lhes permitam interagir e trabalhar com estas tecnologias, tornando-se

agentes ativos e não apenas espectadores passivos na sociedade digital em constante transformação. Diante disso, torna-se urgente investigar os perigos e incertezas associadas a este uso de tecnologia.

1.2. Deepfakes

A rápida evolução das tecnologias de IA tem promovido transformações significativas na produção e manipulação de conteúdo digital. Recentemente, a IA ganhou destaque mundial ao demonstrar a capacidade de gerar textos e manipular e criar áudios, fotos e vídeos sintéticos altamente realistas – uma prática conhecida como Deepfake (Patel et al., 2023).

Segundo Ruiter (2021): "A tecnologia deepfake refere-se a técnicas de aprendizado de máquina que podem ser usadas para produzir arquivos de vídeo ou áudio com aparência e som realistas de indivíduos fazendo ou dizendo coisas que não necessariamente fizeram ou disseram." (p. 2).

Os deepfakes são considerados uma das formas mais nocivas da desinformação pois conseguem enganar facilmente as pessoas, sejam pessoas com conhecimento sobre o tema ou não (Prado, 2021). Esta ferramenta tem como objetivo "intensificar conflitos e debates existentes, minar a confiança nas instituições estatais e incitar a raiva e as emoções em geral. A erosão da confiança provavelmente tornará a atividade policial mais difícil." (Europol, 2022, p. 6).

É importante realçar que a manipulação de conteúdos, apesar de se ter tornado algo popular em 2017, devido a um utilizador da aplicação Reddit que se chamava "deepfake" que publicava vários vídeos de teor sexual com o rosto de atrizes famosas, não é uma prática exclusiva do século XXI (Meskys et al., 2020). O conteúdo manipulado recua à antiguidade e à Idade Média, com a falsificação de pergaminhos, cartas, moedas e joias, ou seja, a falsificação de conteúdos e objetos quando era vista como uma mais-valia era realizada para proveito próprio (Karaboga et al. 2024).

Efetivamente, a sua origem está relacionada com a história da fotografia, dos vídeos e das gravações de som, mesmo antes de esses se tornarem digitais. Assim, a primeira fotografia considerada falsa remonta a 1840, com a imagem intitulada de "Noye" de Hippolyte Bayard. Ainda no mesmo século, surgiram várias outras fotografias que recorriam a

técnicas de dupla exposição, com o intuito de representar fantasmas (Kaswan et al., 2023; Kietzmann et al., 2020).

A Europol (2020) ao apresentar que os deepfakes podem ser agrupadas em cinco categorias demonstra a complexidade do tema. Dito isto, é possível identificar: a substituição facial, acontece quando o rosto de uma pessoa é sobreposto ao outro; a reencenação facial, na qual são alteradas expressões faciais para simular afirmações não proferidas; a geração de rostos sintéticos completamente fictícios; síntese de fala, que cria áudios falsificados com recurso a algoritmos; e os *shallowfakes*, que recorrem a métodos de edição mais simples e rudimentares, não baseados em tecnologias IA (Europol, 2022; Karaboga et al., 2024).

1.2.1. O processo de criação

Os deepfakes são produtos da IA gerados por meio de tecnologias de *deep learning* que manipulam pixéis ou sons, permitindo até a criação de rostos em filmes e vozes em gravações já existentes (Junior & Hessel, 2021).

A criação desta tecnologia baseia-se em técnicas de aprendizagem automática, especificamente em redes neuronais artificiais e nas Generative Adversial Networks (GANs) que utilizam duas redes neurais para criar conteúdos convincentes. Essas duas redes são dois componentes importantes para a sua criação. O nome desses componentes é gerador e discriminador e, ambos, trabalham em conjunto de forma competitiva, isto é, o primeiro tenta criar conteúdos falsos realistas e o segundo avalia a autenticidade desses conteúdos, comparando-os com imagens reais e com o seu conhecimento sobre movimentos humanos. Esta interação é benéfica na medida que eles aperfeiçoam a sua capacidade de gerar imagens e vídeos convincentes (Ruiter, 2021).

É importante realçar que este processo é possível devido à alimentação de uma grande quantidade de material visual que consiste em imagens e vídeos. Estes materiais permitem o modelo de geração de deepfakes aprender as características do rosto alvo, em específico, a recolha e o pré-processamento de dados, incluem uma variedade de poses, expressões, condições de iluminação e fundos, que levam à reprodução digital hiper-realista e sincronizada do rosto (Kaswan et al., 2023; Ruiter, 2021).

Ao longo do tempo, a sofisticação dos deepfakes aumentou significativamente graças aos ciclos de feedback entre investigadores que criam esta tecnologia e os que desenvolvem mecanismos para a detetar. Assim, esta constante produção e aperfeiçoamento resultou em materiais cada vez mais difíceis de detetar, com o potencial de distorcer opinião pública, enganar indivíduos e prejudicar reputações (Odeh, 2024; Ruiter, 2021).

Além disso, existe uma facilidade de acesso muito grande, na medida que existem sites e aplicações gratuitas que oferecem ajuda na criação e manipulação de imagens e áudios. Na verdade, é preocupante, que pessoas com nenhuma ou pouca formação técnica facilmente consigam produzir este tipo de conteúdo (Molina & Berenguel, 2022).

1.2.2. O impacto

O surgimento de deepfakes traçou um panorama misto, por um lado, criou aplicações úteis principalmente no setor de entretenimento, mas, por outro lado, trouxeram ameaças cada vez mais sofisticadas e difíceis de detetar, como o uso de aplicações de criação de áudio deepfake para cometer fraudes (Karaboga et al., 2024).

Esta tecnologia torna mais fácil retratar alguém a dizer ou fazer algo que nunca disse ou fez, levantando uma série de questões éticas importantes. Ao longo dos tempos alguns autores afirmaram que os deepfakes apesar de serem uma mais-valia em algumas áreas, noutras suscitam preocupações éticas consideráveis. Neste contexto, torna-se essencial analisar os impactos dessa tecnologia, ponderando benefícios e desafios (Citron & Chesney, 2019; Meskys et al, 2020; Odeh, 2024).

1.2.2.1. Benefícios

Os deepfakes podem ser utilizados como uma mais-valia em diversos campos como a aprendizagem, a terapia, a saúde e o entretenimento (Danry et al., 2022).

Relativamente a área do entretenimento e da arte, é realçado por Junior e Hessel (2021) “vários benefícios relacionados com o uso dos deepfakes para criar experiências personalizadas para cada utilizador como como ter uma conversa com um artista ou ser saudado por uma figura ilustre do passado ao adentrar em uma galeria.” (p.84). Como exemplo desta ideia, o *Dalí Museum* surpreendeu os visitantes ao recriar o artista, Salvador Dali, para fazer uma visita guiada ao espaço e contar a sua vida a partir da sua arte (Dalí Museum, 2019).

Complementando, na área da saúde, os deepfakes têm sido utilizados para a reabilitação da fala, isto é, a reconstrução da própria voz do doente, nomeadamente em casos de sequelas de AVC[1] ou ELA[2], a partir da atividade cortical e de gravações antigas do paciente, restituindo a identidade vocal e a maior naturalidade conversacional (Metzger et al., 2025).

1.2.2.2. Malefícios

Apesar dos benefícios, a utilização indevida de deepfakes representa um desafio crescente para a sociedade e para o sistema penal. A manipulação digital tem sido utilizada como instrumento de chantagem, extorsão, roubo de identidade, incitamento à violência contra minorias, reforço de narrativas extremistas ou terroristas e humilhação pública, com consequências devastadoras para a integridade psicológica das vítimas, como ansiedade, depressão, stress pós traumático e isolamento social (Chesney & Citron, 2019; Patel et al., 2023; UNICRI, 2025; UNFPA, 2025a).

Efetivamente, a popularização desta tecnologia, anteriormente restrita a profissionais de indústria, tornou-se acessível, tendo sido, nos últimos anos, apropriada por agressores sexuais "como um instrumento de vitimização feminina" (Rodrigues, p.3, 2023).

De facto, esta ferramenta tornou-se poderosa, na medida, que consegue influenciar ou distorcer a verdade, tanto na esfera política como na social, ou seja, quando é utilizada dessa maneira transforma-se numa extensão de *fake news* (Molina & Berenguel, 2022). A própria Europol (2022) alerta para o impacto significativo que os deepfake podem ter no trabalho policial e no sistema judicial, exemplificando com situações em que vídeos adulterados mostram falsamente suspeitos a sair do local do crime, conduzindo perseguição de indivíduos inocentes ou a utilização de provas audiovisuais manipuladas em tribunal, comprometendo a integridade do processo penal.

Em síntese, o realismo dos deepfakes dificulta o combate à desinformação uma vez que os métodos tradicionais de verificação de fontes e da veracidade tornam-se menos eficazes (Stanciu & Ciuperca, 2024). Os constantes avanços tecnológicos vão tornar cada vez mais desafiante a deteção destes conteúdos adulterados, sendo, por isso, necessário

[1] Acidente Vascular Cerebral
[2] Esclerose Lateral Amiotrófica

criar e desenvolver novas bases de dados e métodos de deteção sofisticados para conseguir lutar contra esta realidade e recorrer a uma verificação completa das provas digitais de modo a mostrar que são fiáveis e autênticas (Europol, 2022; Korshunov & Marcel, 2019).

1.3. Deepfake sexuais não consensuais (DSNC)

O relatório da Home Security Heroes (2023) analisa o fenómeno dos deepfakes a partir de uma amostra robusta de 95820 vídeos manipulados, 85 canais especializados e mais de 100 sites. Neste foi aferido que entre 2019 e 2023 houve um aumento de 550% de volume de deepfakes disponíveis na internet, o que evidencia a velocidade com que esta ferramenta evolui.

Com efeito, a rápida criação de sites, comunidades e aplicações, como DeepNude[3], facilitou a partilha, venda e criação deste tipo de conteúdo, isto é, o que antes demorava horas, hoje pode levar apenas minutos para ser criado (Home Security Heroes, 2023; Rodrigues, 2023).

O Relatório Anual de Segurança Interna (RASI) (2024) embora não fale diretamente de deepfakes sexuais afirma que "…a *Sextortion*, ou tentativa de extorsão com a ameaça de exposição de imagens de teor íntimo da vítima, mantém-se como uma ameaça, representando cerca de 8% dos incidentes…" (p.111).

É importante também evidenciar que a facilidade de acesso e massiva disponibilização a conteúdos pornográficos explícitos tem criado um basto banco de dados informal que alimenta a contínua criação de deepfakes sexuais (Newton & Stanfill, 2020). Por isso, a forma atual mais comum dos deepfakes são os de conteúdo sexual, representando entre 96% e 98% de todos os deepfakes encontrados online, vitimando maioritariamente mulheres (End Violence Against Women, 2024; Home Security Heroes. 2024; Umbach et al., 2024).

[3] "are named after an app entitled DeepNude that enabled users to upload images of clothed women that would then be 'stripped' of their clothing as the app would match a woman's face to a nude body. While the original application was removed, many versions of it subsist online, including a popular version on messaging application Telegram." (Lalonde, 2022, p.2)

A diferença entre deepfakes sexuais consentidos e não consentidos é o consentimento da pessoa retratada na criação. Em termos normativos, este requisito segue a mesma ideia do artigo 79.º do Código Civil português

> 1. O retrato de uma pessoa não pode ser exposto, reproduzido ou lançado no comércio sem o consentimento dela… 3. O retrato não pode, porém, ser reproduzido, exposto ou lançado no comércio, se do facto resultar prejuízo para a honra, reputação ou simples decoro da pessoa retratada.

Assim, os deepfakes sexuais consentidos são todos os quais o indivíduo retratado autoriza a criação de partilha ou exibição do conteúdo, geralmente em contexto de relacionamento ou com consentimento explicito, enquanto os DSNC consistem na produção e disseminação de conteúdos sexualmente explícitos manipulados através de IA, geralmente pela substituição do rosto de uma vítima sobre o corpo de outra pessoa em vídeos ou imagens pornográficas (Brigham et al., 2024; Flynn et al., 2022; Mcglynn et al., 2022).

Os DSNC são classificados como uma modalidade de abuso sexual de imagem baseada em tecnologia e uma nova forma de violência sexual digital, onde a intimidade é violada não por uma exposição real, mas por uma simulação visual altamente realista (Brigham et al, 2024; Flynn, et al, 2022; Mcglynn et al., 2022).

Conforme Ferreira (2024) “o sexo feminino é o alvo, o que apenas significa que esta é a nova e crescente forma de violência contra as mulheres.” (p. 8). Dito isto, estes representam uma violência contra as mulheres baseada no género pois afetam desproporcionalmente as mulheres (Convenção do Conselho da Europa para a Prevenção e o Combate à Violência contra as Mulheres e a Violência Doméstica, 2011).

1.3.1. Perpetradores

De acordo com o estudo realizado por Umbach et al (2024), a prevalência de perpetradores de deepfake sexuais varia entre países. Na realidade são poucos os que afirmam ter desenvolvido este tipo de deepfakes, aproximadamente 1.8% de indivíduos na amostra global relataram ter criado.

Contrariamente, alguns países exibem taxas mais altas do que a média global. Entre eles é possível identificar os Estados Unidos da América (EUA) com 2.6% e a Austrália com

2.4%, tendo sido os mesmos onde ocorreram mais atos de criação ou ameaça de criação de deepfake (Umbach et al., 2024).

Quanto ao género dos perpetradores, os homens tendem a estar mais envolvidos do que as mulheres como criadores. Adicionalmente, evidencia-se uma maior taxa de ameaça em divulgar deepfakes pornográficos do que a de indivíduos que efetivamente criam ou partilham o conteúdo (Umbach et al., 2024).

Outro estudo suíço indica que homens com menos de 35 anos têm uma maior probabilidade de criar ou consumir estes conteúdos sintéticos (Xu et al., 2025).

1.3.1.1. Motivações

As razões que levam indivíduos a produzirem e disseminarem deepfakes de caráter sexual são múltiplos, entre eles é possível evidenciar o ganho monetário, a vingança, a gratificação sexual, a chantagem e a prova de masculinidade ou de poder (Han et al., 2025; Okolie, 2023).

A motivação central identificada em estudos recentes é o prazer e satisfação de fantasias eróticas, por vezes devido à falta de conteúdo erótico que enquadre nos parâmetros do comprador. Estes casos são claramente associados aos de deepfakes pornográficos de celebridades femininas (Han et al., 2025; Okolie, 2023). Dessa forma, os DSNC funcionam como uma extensão tecnológica do voyeurismo e da fantasia sexual, permitindo que os espectadores assistam a conteúdos eróticos amplificados por meio de IA (McGlynn & Topalak, 2025).

Embora menos discutida no discurso popular, a motivação económica também exerce um papel relevante. O aumento da procura deste deepfakes, formou verdadeiros mercados online de pornografia deepfakes, como o MrDeepFakes[4], onde criadores dos DSNC vendem vídeos sob encomenda, de celebridades e de pessoas comuns, mesmo que as diretrizes do site proíbam a proliferação de conteúdos de não celebridades (Han et al., 2025).

Notavelmente, motivações de vingança pessoal tendem a ser uma exceção, pois na maioria dos casos os autores não conhecem as vítimas. No entanto, quando este tipo de casos ocorre, ou seja, quando o deepfake é utilizado com intuito de assediar, aterrorizar,

[4] Maior plataforma de deepfakes sexuais, encerrada a 5 de maio de 2025

humilhar ou prejudicar a reputação da vítima, torna-se semelhante à ideia de *revenge porn*[5] aquando distribuídos e criados por ex-parceiros (Han et al., 2025; Ferreira, 2024).

As dinâmicas de poder e masculinidade tóxica também foram identificadas como motivadores de divulgação e criação. Em outras palavras, perpetradores vêm nestes conteúdos uma arma de poder e controlo e como uma forma de afirmar o domínio sobre a vítima. Estas motivações levam normalmente a conteúdos sádicos[6] com cenários sexuais da vítima a ser punida fisicamente (Han et al., 2025).

A *sextortion*[7] também é utilizada para obter ganhos monetários ou favores sexuais (Cambrige Dicionary, 2025; Okolie, 2023). Abusadores sexuais de menores normalmente utilizam estes conteúdos para ameaçar crianças, com o objetivo de obter favores ou conteúdos sexuais reais de menores (Internet Watch Foundation, 2024).

Em alguns casos, as motivações para a divulgação deste conteúdo pode ser diferente do motivo ou pedido de criação. Seguindo esta ideia, Rodrigues (2023) chama a atenção para a partilha orientada "…por uma lógica de *slut shaming*… processo social de menosprezo e degradação à mulher que viole as expectativas culturais de conduta sexual esperada de seu gênero" (p.286). Esta ideia pode ser corroborada pelas conceptualizações de género, em específico com a feminilidade enfatizada em que muitas vezes o ser humano vê a mulher como um símbolo de castidade e pureza.

1.3.2. Vítimas

As vítimas desta tecnologia são, maioritariamente, celebridades. Na verdade, 94% das pessoas retratadas em DSNC trabalham nas diferentes áreas do setor de entretenimento, especialmente cantoras (58%) e atrizes (33%). Geograficamente, este tipo de divulgação e criação de conteúdo acontece muito associado a atrizes e cantoras sul-coreanas, seguidas de norte-americanas e japonesas (Adjer, 2024; Home Security Heroes, 2023; Rodrigues, 2023).

Apesar da maioria dos DSNC sejam relacionados a famosos, com a evolução e sofisticação da tecnologia, logo, começaram a criar conteúdos de teor sexual de pessoas

[5] "private sexual images or films showing a particular person that are put on the internet by a former partner of that person, as an attempt to punish or harm" (Cambridge Dictionary, 2025a).
[6] "sadismo… desejo de fazer sofrer o objeto sexual e é considerada uma perversão sexual, já que este desejo substitui o fim normal de uma relação sexual e é a única forma do indivíduo obter prazer" (Infopedia, 2025).
[7] "the practice of forcing someone to do something, particularly to perform sexual acts, by threatening to publish naked pictures of them or sexual information" (Cambridge Diocionary, 2025b).

mais próximas como professoras, alunas, vizinhas, ex-parceiras entre outros (Rodrigues, 2023).

A percentagem global de indivíduos que relataram ser vítimas foi de aproximadamente 2.2%, tendo uma taxa maior em países como Austrália (3.7%), Coreia do Sul (3.1%) e México (2.9%). A maioria dos relatos refere-se a casos onde conteúdo foi criado e partilhado online, sem consentimento e de forma anónima (Umbach et al., 2024).

Os DSNC não são uma realidade apenas relacionada a maiores de idade. Um dos cenários destacados pode ser a utilização desta tecnologia por abusadores sexuais de crianças para criar imagens ou vídeos sexuais onde a voz, corpo e cara pareçam de uma criança para a sua satisfação sexual (U.S. Department od Homeland Security, 2021). Efetivamente, no período de 9 de março a 7 de abril, foram encontrados 9 vídeos deepfake em fóruns dark web que sobrepõem o rosto de uma criança a pornografia adulta (Internet Watch Foundation, 2024).

1.3.3. Relações vítima-perpetrador

A relação entre vítima e perpetrador influência significativamente as atitudes em relação à criação de conteúdo não consentido (Brigham et al., 2024).

Efetivamente, na maioria dos casos, os autores do conteúdo não apresentam qualquer ligação ou relacionamento próximo com a vítima, como hackers ou qualquer pessoa que procure ganho financeiro ou reconhecimento. Dessa forma, são poucos os incidentes que envolvam contactos próximos entre a vítima e autor. Todavia, quando essa acontece os autores costumam ser ex-companheiros ou ex-parceiros íntimos (Brigham et al., 2024; Meskys et al., 2020; Umbach et al., 2024).

1.3.4. Normalização da criação

Segundo Skyes e Matza (1957) na sua teoria de técnicas de neutralização, os delinquentes recorrem a estratégias cognitivas que lhe possibilitam violar normas sem perder a autoimagem de “pessoa” normal”.

É importante realçar que são escassas as pesquisas que estudam a perceção do público a cerca desta ferramenta, contudo, alguns afirmam que nos fóruns de deepfakes, observam-se por parte de consumidores e autores: negações de existência de uma vítima, apelos à ausência de danos ou transferência de responsabilidade para a tecnologia ou para a vítima (Brigham et al., 2024; Han et. al, 2024; Karaboga et al., 2024; Umbach et. Al, 2024).

No geral, 47% do público tem uma atitude positiva enquanto 36.8% atitude negativa e 16.1% atitude neutra em relação aos deepfakes. Em contrapartida, os deepfakes de conteúdo adulto reúnem mais atitudes negativas (47.5%) do que positivas (42%) evidenciando sentimentos negativos como raiva, medo e tristeza por parte dos participantes (Xu et al., 2025).

Contudo, desde 2022, o Reddit tem observado um aumento significativo a favor de comunidades dedicadas à criação e partilha de deepfakes pornográficos. O aumento deveu-se, em parte, pelas políticas de moderação pouco ajustadas para a nova problemática. Por isso, é possível identificar no discurso dos participantes, uma frequência elevada de vocabulário misógino[8] e vulgar (Gamage et al., 2023).

Apesar de vários participantes concordarem com a criminalização dos deepfakes sexuais, os mesmos reportam consumir conteúdo pornográfico fabricado de celebridades (Umbach et al. 2024).

A aceitação não é homogênea pois quando se aborda o género é possível observar que os homens tendem a aceitar ou a classificar como menos graves do que mulheres (Umbach et al. 2024).

Deveras, existe maior aceitação quando não há intenção maliciosa pois sugere que algumas pessoas podem ser motivadas pela curiosidade, diversão ou exploração de novas formas de comunicação. Por exemplo, em relacionamentos amorosos, em que já houve relações íntimas, é considerado aceitável a criação desde que não partilhem o conteúdo, sendo até considerado engraçado ou lisonjeador (Brigham et al., 2024).

Realmente, existem artigos que afirmam que 74% dos perpetradores não sente culpa de ter divulgado ou criado deepfakes pornográficos, muitas vezes proveniente e relacionada à impunidade desta conduta (Home Security Heroes, 2024).

Sem dúvida existe uma maior aceitação por parte dos usuários quando a criação e o consumo são motivados por satisfação de fantasias sexuais ou entretenimento, do que quando são criados para magoar ou humilhar terceiros. Igualmente, existe significativamente mais aceitação quando as vítimas são celebridades do que quando são "pessoas normais" (Umbach et al., 2024).

[8] "feelings of hating women, or the belief that men are much better than women" (Cambridge Diocionary, 2025c)

Na verdade, alguns consumidores deste conteúdo afirmam que a sua existência pode ser considerada uma contribuição para a sociedade (Brigham et al., 2024; Han et al., 2025; Umbach et al., 2024).

1.3.5. Métodos de combate e deteção

A sofisticação crescente da tecnologia deepfake exige uma resposta multidimensional, que envolva tanto medidas tecnológicas, organizacionais e jurídicas. De acordo com Westerlund (2019), existem quatro formas de combater deepfakes: "1) legislação e regulamentação, 2) políticas corporativas e ações voluntárias, 3) educação e treinamento, e 4) tecnologia anti-deepfake que inclui deteção de deepfake, autenticação de conteúdo e prevenção de deepfake." (p.44).

Assim, o possível aumento da disseminação e normalização da prática reforça a necessidade de regulamentação e de estratégias de enfrentamento, ou seja, esta ameaça emergente requer atenção multidisciplinar envolvendo educação, legislação e ações sociais para mitigar seus efeitos nocivos, especialmente na proteção de vítimas vulneráveis e na conscientização social (Junior & Heller, 2021; Umbach et al., 2024).

1.3.5.1. Tecnologia

As plataformas digitais onde os deepfakes sexuais circulam, como YouTube, Instagram, TikTok, Reddit e X, tem um papel central, dado que é nelas que os conteúdos manipulados são geralmente divulgados e disseminados (Europol, 2022). O estudo de Newman et al. (2024), revela que mais de 30% dos utilizadores da Internet consomem notícias através do Facebook e do Youtube. Assim, o desafio aqui é remover os conteúdos (Junior & Heller, 2021).

Portanto, é essencial, para prevenir, detetar e definir normas e responsabilizações claras, a existência de uma colaboração entre governos, empresas, tecnologias e Organizações Não Governamentais (Europol, 2022).

Várias empresas têm procurado adotar políticas que proíbam a partilha e publicação destes conteúdos com base na sua intencionalidade e impacto social (UNICRI, 2025). Por exemplo, a Meta (2025), proprietária do Instagram e do Facebook, estabelece parcerias para avaliar a veracidade de conteúdos em tempo real e remove conteúdos fabricados

quando estes apresentam probabilidade elevada de causar danos físicos iminentes ou interferir diretamente em processos políticos.

Neste sentido, surgiram iniciativas focadas na proteção de vítimas de conteúdos íntimos reais ou adulterados, divulgados sem consentimento. Entre essas, destacasse a StopNCII.org, desenvolvida pela Revenge Porn Helpline (RPH) em parceria com algumas plataformas digitais. Esta ferramenta baseia-se na criação de um "hash" – código único e irreversível gerado a partir da imagem ou vídeo original – que bloqueia o envio ou publicação de uma imagem em plataformas digitais. Desta forma, a tecnologia atua como um mecanismo preventivo eficaz que assegura a privacidade e proteção da vítima, sem que o conteúdo original precise sequer ser carregado online (StopNCII.org, 2025).

A utilização de marca de água digital é uma técnica eficaz na deteção em tempo real de conteúdos manipulados, que garanta uma resistência à adulteração e ofereça meios de rastreabilidade legal (Lai et al., 2025). Seguindo esta ideia, o Google desenvolveu o SynthID que durante o processo de criação de uma imagem por tecnologia IA incorpora uma marca de água, invisível a olho humano, que funciona como um identificador digital de conteúdo IA permitindo o reconhecimento do conteúdo pelas plataformas digitais. Com isso, é possível aplicar medidas como bloquear a publicação ou adicionar uma etiqueta informativa que esclareça a sua origem artificial, contribuindo assim para a transparência e o combate à desinformação visual (Heikkilä, 2023).

Complementarmente, empresas como a Sensity AI (2024) desenvolveram sistemas especializados na deteção de vídeos pornográficos falsificados, através de análises técnicas detalhadas e ficheiros multimédia. A plataforma possui uma equipa técnica que monitoriza a *web* e *dark web*, reunindo a maior base de dados conhecida de deepfakes identificados, o que permite antecipar novas ameaças.

1.3.5.2. Legislação e Políticas

No começo, os deepfakes eram sobretudo montagem pouco realistas, por isso, durante muito tempo o Direito Penal não observava esta tecnologia como uma ameaça (Rodrigues, 2023). Porém, diante dos impactos que os deepfakes sexuais trazem atualmente, muitos ordenamentos jurídicos passaram a criminalizar a criação ou divulgação de conteúdos manipulados com o intuito de os conter.

Na realidade a End Violence Against Women (2024) sublinha que sem delito próprio, as vítimas enfrentam lacunas legislativas que podem dificultar a responsabilização. Também, a ausência de culpa, por parte dos autores, proveniente da falta de responsabilização, reforça a importância da criação de políticas e campanhas de sensibilização (Home Security Heroes, 2024).

De acordo com a teoria clássica da dissuasão penal, a criminalização acompanhada de penas poderá ter um efeito preventivo geral, desincentivando potenciais infratores pela perceção do risco de punição (Beccaria, 1764). A existência de legislação contra deepfakes com consequências de pena de prisão ou multas elevadas podem dissuadir a divulgação e criação desses. A título exemplificativo desta ideia, após o anúncio da nova ofensa no Reino Unido dois sites de pornografia deepfake bloquearam o acesso a utilizadores britânicos, demonstrando um efeito inibidor imediato da legislação proposta (Levy, 2023).

Diversos países, como a Coreia do Sul, Reino Unido e EUA já regularizaram os deepfakes. Em 2025, foi aprovada quase por unanimidade, nos EUA, a "TAKE IT DOWN ACT", uma lei federal especificamente dirigida à proibição da circulação de conteúdos íntimos não consensuais, reais ou sintéticos, incluindo deepfakes pornográficos (United States Congress, 2025).

No espaço europeu, em 2024, a União Europeia (UE) "a fim de proteger eficazmente as vítimas de tais comportamentos…" (Parlamento Europeu & Conselho da União Europeia, 2024, p.4) aprovou a Diretiva UE 2024/1385, que estabelece padrões mínimos para a criminalização de certas formas de violência, obrigando os Estados-Membros a tipificarem a criação de divulgação de conteúdos sexualmente explícitos falsificados sem consentimento, até 2027. Paralelamente, outro regulamento, estabeleceu requisitos de transparência para sistemas de IA que geram conteúdos sintéticos, incluindo deepfakes. Especificamente, os artigos 50.º a 53.º estabelecem que tais conteúdos devem ser devidamente identificáveis através de marcações específicas (Parlamento Europeu & Conselho da União Europeia, 2024).

O estudo nacional de Moreira (2025) afirma que embora alguns casos os deepfakes possam enquadrar-se em tipos legais já existentes do Código Penal (CP), ainda subsistem lacunas, sendo por isso necessário a criação de leis que contemplem os danos particulares

causados pelos deepfakes sexuais. Por exemplo, na Lei 26/2023, artigo 193.º "Devassa através de meio de comunicação social, da Internet ou de outros meios de difusão pública generalizada", ainda existem falhas, como a falta de menção clara a conteúdos manipulados .

Ainda no contexto nacional, coloca-se o desafio de avaliar a autenticidade de conteúdos digitais apresentados como prova. Por conseguinte, torna-se imprescindível que, os Órgãos de Polícia Criminal (OPC) e demais intervenientes da justiça, desenvolvam competências técnicas especializadas na deteção de deepfakes e adotem métodos validados para a certificação da autenticidade das provas digitais, garantindo assim a integridade do processo penal e evitando decisões judiciais baseadas em provas adulteradas (Europol, 2022).

Como refere McGlynn (2024) é o facto do consentimento, isto é, não podemos criminalizar a publicação, criação ou partilha de deepfakes que tiveram o consentimento da pessoa em questão. Não obstante, esta autora salienta que a implementação desta legislação específica encontra a oposição sob o argumento de que são meras fantasias sexuais, invocando-se a liberdade de expressão.

Em suma, Delfino (2019) chama atenção para as lacunas da lei afirmando que "…qualquer pessoa que tenha aparecido numa imagem digital pode "estrelar" pornografia contra a sua vontade e, atualmente, a lei não oferece nenhum recurso claro ou direto para impedir isso." (p.890). Mesmo países que já proíbem esta divulgação realçam quão desafiante é identificar os autores de crime o que gera um sentimento de impunidade (Ferreira, 2024).

1.3.5.3. Consciencialização Social

Umbach et al, (2024) realça que embora alguns países tenham leis específicas a efetividade dessas leis é limitada devido à baixa conscientização publica, dificuldades na comprovação de violações e desafios na aplicação jurídica. Logo, há uma necessidade urgente de ações educativas para informar o público sobre riscos, consequências e danos, especialmente entre homens jovens dada a sua maior propensão para se envolver nesta conduta e a sua menor perceção de gravidade (Umbach et al., 2024).

Dessa forma, programas educativos aumentam significativamente a capacidade de identificar conteúdos manipulados, promovendo uma sociedade mais crítica e resiliente

face à desinformação e tomadas de decisão informadas. Hoq et al. (2025) demonstraram que, após uma intervenção educativa, os participantes melhoraram significativamente a sua precisão na deteção de deepfakes.

A UE publicou orientações específicas destinadas a professores e educadores com o objetivo de combater a desinformação, promover a literacia digital e assegurar um mundo digital seguro aos jovens ao sensibilizar e promover uma utilização responsável. Embora não trate exclusivamente da problemática de deepfakes pornográficos, aborda diretamente os desafios colocados pela manipulação digital e a disseminação de conteúdos enganosos, como parte integrante da desinformação online. As orientações incentivam o desenvolvimento de competências críticas, como a verificação de fontes, o reconhecimento de manipulações visuais e o uso responsável da tecnologia de modo a ajudar os jovens a reconhecer deepfakes e outras formas de desinformação. Além disso, o documento enfatiza a necessidade de formar educadores para falar sobre estas temáticas nas práticas curriculares (Comissão Europeia, 2022).

Além disso, desenvolver campanhas dirigidas a grupos vulneráveis, como seniores ou comunidades com pouco acesso a educação, é fulcral para evitar que se tornem alvos fáceis da desinformação. Consequentemente, a falta de familiaridade com os meios digitais e a baixa literacia tem sido apontada como um dos fatores críticos que facilitam a proliferação de deepfakes, pois é a população sem competências digitais que tende a partilhar e acreditar em conteúdos manipulados compartilhando-os, alimentando o que se pode considerar uma verdadeira epidemia de desinformação (Mokadem, 2023).

Para além das campanhas formais e educacionais, projetos culturais e artísticos tem demonstrado ser ferramentas preciosas na sensibilização social. Segundo Godulla (2022) intervenções artísticas que exploram uma tecnologia deepfake contribuem para o desenvolvimento do pensamento crítico e criam uma resposta emocional que muitas vezes ultrapassa os efeitos das abordagens convencionais, ou seja, expor ao público a problemática de forma criativa e provocadora, pode aumentar a consciência sobre os riscos.

CAPÍTULO II – PROPOSTA DE ESTUDO

O segundo capítulo foca-se na exposição da proposta de estudo. Dito isto, aqui serão apresentados os objetivos gerais e específicos, a descrição da metodologia utilizada, a

amostra, os instrumentos e procedimentos necessários para realizar o estudo e, por fim, a identificação dos resultados esperados.

2.1. Objetivos Gerais e Específicos

Os objetivos do projeto são as metas que se esperam alcançar e realizar na execução do futuro plano de ação. Assim, seguindo a meta SMART[9], foi estabelecido um objetivo geral claro e relevante, como quadro de referência do que se espera do projeto (Doran, 1981).

Neste sentido, o objetivo geral desta proposta de estudo seria compreender, entre junho de 2025 até junho de 2027, os discursos, lógicas, motivações e dinâmicas sociais presentes em comunidades digitais que criam e compartilham DSNC.

Para que o objetivo principal do trabalho seja concretizado, é necessário estabelecer objetivos mais específicos. Assim, foram estipulados os seguintes objetivos como base de orientação para a sua execução:

1. Explorar os principais argumentos e racionalizações utilizadas pelos autores e divulgadores para justificar a criação ou partilha de DSNC.
2. Caraterizar o perfil sociodemográfico (idade, género, país, escolaridade) dos perpetuadores
3. Investigar as motivações explicitas e implícitas que levam a criação e disseminação destes conteúdos.
4. Descrever as dinâmicas grupais e os mecanismos de validação entre pares presentes nos espaços online.
5. Examinar a presença de elementos misóginos e discursos de desumanização das vítimas nos discursos.

2.2. Método

A fase metodológica de um trabalho científico corresponde ao momento em que se estabelecem os procedimentos de recolha, organização e análise de dados. Aqui é

[9] SMART é uma sigla inglesa que significa: specific (específico); measurable (mensurável); achievable (alcançável); realistic (realista); time-bound (com prazos determinados) (Doran, 1981)

assegurada a congruência lógica entre as perguntas de investigação, o enquadramento teórico e os resultados.

De acordo com os objetivos estabelecidos e à escassez de literatura empírica, a investigação adota uma abordagem qualitativa de caráter exploratório, fundamentada em entrevistas semiestruturadas, para obter dados ricos e contextuais que não seriam acessíveis por meio de métodos puramente documentais. Diferentes dos estudos quantitativos em que os dados vêm em forma de número o estudo qualitativo concentra-se em entender um fenómeno a partir de várias formas como observações, focus groups, documentos e multimédia entre outros (Theodorson & Theodorson, 1969).

Quanto ao conceito de pesquisa exploratória, refere-se a um estudo cujo objetivo é familiarizar-se sobre um fenómeno a ser investigado, para que o estudo principal subsequente possa ser elaborado com maior compreensão e precisão (Theodorson, & Theodorson, 1969).

2.2.1. Amostra/Participantes

A amostra incidirá sobre perpetradores de deepfakes sexuais não consensuais, definidos como indivíduos que conscientemente manipularam ou divulgaram conteúdos manipulados de carácter íntimo sem o consentimento da pessoa retratada.

Relativamente ao tamanho da amostra, prevê-se recrutar entre quinze e vinte participantes, com idade mínima de 18 anos, que falem português ou inglês, que relatem ter criado ou divulgado em comunidades digitais como o Reddit, o Telegram e o X.

Dada a dimensão virtualmente ilimitada da amostra e a natureza qualitativa da investigação, opta-se por uma amostragem intencional, em fóruns e grupos públicos das comunidades já mencionadas, tal como recomenda Patton (2015) para fenómenos emergentes em que a profundidade analítica é prioritária face à representatividade estatística. A amostragem intencional foca “na seleção de casos ricos em informação cujo estudo irá iluminar as questões em estudo” (Patton, 2015, p.230).

De modo a garantir o anonimato, os convites serão enviados por mensagens privadas, informando de forma clara a afiliação institucional do pesquisador e os objetivos gerais do estudo sobre DSNC, de modo a garantir a transparência e permitir o consentimento livre e esclarecido dos participantes.

2.2.2. Instrumentos

Para garantir rigor metodológico e rastreabilidade, a proposta de estudo, recorre a três instrumentos: um diário reflexivo, um inquérito online anónimo com perguntas fechadas e abertas, e entrevistas semiestruturadas em formato digital, articulados em sequência lógica de recolha, organização, análise e reflexividade.

A investigação irá utilizar um questionário, via plataforma *Google Forms,* que irá compreender três partes distintas. O primeiro bloco será de perguntas fechadas relativas às características sociodemográficas da pessoa, como idade, género e país. Em segundo lugar, será questionado o nível de familiaridade com a tecnologia de deepfake, isto é, frequência de uso ou ferramenta utilizada. Por fim, serão apresentadas perguntas sobre motivações e desculpabilização, com respostas de acordo com Escala *Likert*, instrumento amplamente empregado em pesquisas sobre interesses, atitudes, entre outros que mede o grau de concordância dos participantes a determinadas afirmações (1- discordo completamente; 2 – discordo; 3 – prefiro não dizer; 4- concordo; 5 – concordo completamente) (cf. Anexo A) (Silva & Costa, 2014).

Também serão utilizadas entrevistas online semiestruturadas, via Zoom pois esse apresenta um software com gravação local e armazenamento seguro. Para cada sessão terá de ser garantido um computador com ligação estável à Internet, fones com microfone de alta sensibilidade para captação nítida de áudio e ambiente discreto e silencioso para evitar interferências. O roteiro de entrevista (cf. Anexo B) combinará perguntas orientadoras previamente definidas que abordem o processo técnico de criação de deepfakes, os canais de disseminação preferidos, as perceções de risco e benefício, as estratégias de neutralização moral utilizadas, com espaço para exploração livre de temas emergentes. A duração da entrevista será aproximadamente de 60 minutos e será iniciada com um lembrete verbal dos termos do consentimento informado, reforçando o anonimato e a voluntariedade.

Adicionalmente, irá ser utilizado um diário reflexivo digital encriptado, para o registo diário sobre o decorrer de entrevistas, dilemas éticos e observações sobre possíveis vieses. Esta prática confere transparência interpretativa, autenticidade e rigor ao estudo (Ortlipp, 2008).

Além dos instrumentos, é importante identificar, que serão necessários alguns materiais de apoio, como formulários digitais de consentimento, matrizes Excel para registar dados de cada participante e software de transcrição automática para uma maior rapidez no processo, seguida de uma revisão manual para impedir erros na transcrição.

Em síntese, os instrumentos e materiais garantem que a recolha de dados seja sistemática, transparente, segura e suficientemente flexível para captar tanto padrões gerais, a partir do questionário, como narrativas profundas, a partir das entrevistas semiestruturadas, assegurando a qualidade e a validade dos resultados.

2.2.3. Procedimentos

Embora não exista contacto direto com participantes humanos identificáveis, a natureza sensível do fenómeno impõe especiais cuidados éticos. Posto isto, inicialmente, é essencial observar as diretrizes éticas consagradas na comunidade científica internacional, assegurando, desde o primeiro momento, a proteção do anonimato e a confidencialidade dos dados recolhidos (ALLEA, 2023). Além disso, é fundamental pedir um parecer à Comissão de Ética da Faculdade de Ciência Humanas e Sociais da Universidade Fernando Pessoa (UFP), antes de começar a investigação.

Somente após aprovação da UFP é que a coleta de dados e a investigação começa. Tendo isso em conta, o processo inicia-se com a divulgação do convite e do link para o questionário nas comunidades-alvo, acompanhando de uma breve apresentação ética e garantia de anonimato, com duração de aproximadamente 10 minutos. Os participantes que indicarem disponibilidade serão convidados a participar nas entrevistas semiestruturadas, agendadas por mensagens privadas.

Antes de iniciar cada participante tem de assinar o termo de consentimento livre e esclarecido em formato digital, que especifica a voluntariedade, o direito à desistência e as medidas de confidencialidade. Como já mencionado, a investigadora seguirá o roteiro da entrevista semiestruturada, incentivando descrições detalhadas do processo técnico de geração de deepfakes, dos canais de partilha e das estratégias de justificação moral.

É importante evidenciar, que todas as sessões serão gravadas, transcritas e submetidas a análise temática. De modo a seguir as regras da comunidade científica, será criado pseudónimos dos identificadores, ou seja, será dado um código aos usernames dos

utilizadores dos quais vão ser retirados os comentários, como acontece no estudo de Brigham et al. (2024).

Ao longo do processo, o diário reflexivo será atualizado, registando-se decisões analíticas, dilemas éticos e perceções pessoais, de modo a promover a transparência do processo de pesquisa, tornando visível a natureza construída dos resultados e os fatores pessoais do pesquisador que influenciam o estudo (Ortlipp, 2008).

2.3. Resultados Esperados

O presente projeto de graduação configura-se como uma proposta metodológica para futura investigação académica, pelo que nenhum dado foi efetivamente recolhido ou analisado.

É importante destacar que, relativamente ao estudo da autora, foi possível averiguar que existe pouca literatura focada em quem cria e compartilha este tipo de conteúdos, sendo a projeção de possíveis resultados uma etapa desafiante. Todavia, a revisão preliminar da literatura e o desenho do protocolo de estudo permitem antecipar os seguintes resultados.

No que concerne o perfil sociodemográfico dos perpetradores, tanto a nível do consumo como de criação, espera-se que seja maioritariamente homens entre os 18-35 anos, em sites como o Reddit (Gamage et al., 2023; Umbach et al., 2024; Xu, et al, 20225). Além disso, as principais motivações, poderão variar entre gratificação sexual, poder ou ganhos económicos, como já foi evidenciado no estudo de Han et al. (2023) e de Brigham et al. (2024).

Ademais, é possível formar um inventário sistemático de técnicas de neutralização usadas por atores e divulgadores de DSNC, nomeadamente a cerca da negação do dano ao culpar a vítima e da utilização deste tipo de conteúdo como uma arte ou como uma contribuição para a sociedade (Han et al., 2025; Umbach et al., 2024). Também, se prevê mecanismos de apoio e reforço comunitário, como uso de linguagem misógina e trocas de elogios ao "realismo" e solicitações de encomendas personalizadas, que contribuem para a proliferação e normalização desses conteúdos (Gamage et al., 2023; Han et al., 2025).

Dito isto, apesar de se prever práticas de neutralização e aceitação de deepfakes sexuais, é provável consumidores e criadores de deepfakes sexuais reconhecerem alguma

necessidade de lei quando se fala de pessoas normais e considerarem excessivas leis contra deepfakes de celebridades (Umbach et al. 2024).

Efetivamente, espera-se que este trabalho possa cumprir os seus objetivos. Todavia, importa reconhecer possíveis limitações que poderão afetar a amplitude e a generalização dos resultados quando a investigação vier a ser concretizada. Por exemplo, as conversas em grupos privados podem conter dinâmicas distintas dos que são estudados e a dificuldade de identificar e distinguir o autor do divulgador, dificultando a criação de "perfil de perpetrador".

CONCLUSÃO

No campo dos deepfakes, percebe-se que a evolução das GANs e o acesso facilitado a ferramentas de criação tornaram os conteúdos sintéticos realistas e difíceis de distinguir dos genuínos. Isso potencializa a desinformação e amplia o impacto psicológico das manipulações, especialmente quando se trata de DSNC.

Importa evidenciar a urgência de observar os DSNC como um fenómeno criminológico que une a violência de género com a criminalidade virtual. Reconhecendo esta realidade, o presente projeto de graduação procurou preencher uma lacuna significativa na investigação criminológica ao deslocar o foco dos DSNC das vítimas para os próprios perpetradores.

Através de um desenho metodológico qualitativo e exploratório, que combina um questionário online e entrevistas semiestruturadas, o estudo procura entender as motivações, racionalizações morais e dinâmicas de legitimação que sustentam a produção e partilha destes conteúdos.

Como todos os projetos este apresenta certas limitações, especialmente no que toca a componentes incontroláveis, como o número de voluntários na participação dos questionários e entrevistas. Posto isto, a principal limitação da proposta apresentada e a mais significativa é que o recrutamento em fóruns públicos do Reddit, X ou Telegram introduz viés de autosseleção e dificulta a verificação da identidade e do real envolvimento dos sujeitos na criação ou divulgação de DSNC, podendo misturar perpetradores ativos com meros observadores.

Em suma, embora o presente projeto se limite a apenas na apresentação de uma proposta de estudo, espera-se que seja observada por investigadores como uma possível hipótese de estudo de modo a preencher a lacuna identificada aquando do enquadramento teórico – a escassez de estudos centrados no agressor – e sustentar e aperfeiçoar ações futuras de política criminal, educação digital e proteção das vítimas num contexto de rápida evolução tecnológica. Em geral, como o criador e consumidor do deepfake não receberam muita atenção na literatura é sugerido e incentivado o desenvolvimento de pesquisas direcionadas a esses autores.

REFERÊNCIAS BIBLIOGRÁFICAS

ALLEA. (2023). The European code of conduct for research integrity. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/european-code-of-conduct-for-research-integrity_horizon_en.pdf

Angeli, P. H. D., Colodette, L., Oliveira, P. H. S. D., & Silva, A. B. D. (2019). A evolução da inteligência artificial e a substituição do trabalho humano. *Revista Ambiente Acadêmico*, *5(1)*, *7–25*.

Beccaria, C., & Voltaire. (1764). An essay on crimes and punishments. *W. C. Little and Company*. https://oll.libertyfund.org/titles/voltaire-an-essay-on-crimes-and-punishments

Bhutani, A., & Sanaria, A. (2023). The past, present and future of artificial intelligence. *Journal of Management Research & Technology*. Advance online publication. https://doi.org/10.1177/jmrt.231199305

Brigham, N. G., Flynn, A., Freeman, K., & Parry, E. (2024). "Violation of my body:" Perceptions of AI-generated non-consensual (intimate) imagery. https://doi.org/10.48550/arXiv.2406.05520

Cambridge Dictionary. (2025a). Revenge porn. Cambridge Dictionary. https://dictionary.cambridge.org/pt/dicionario/ingles/revenge-porn

Cambridge Dictionary. (2025b). Sextortion. Cambridge Dictionary. https://dictionary.cambridge.org/pt/dicionario/ingles/sextortion

Cambridge Dictionary. (2025c). Misogyny. Cambridge Dictionary. https://dictionary.cambridge.org/dictionary/english/misogyny

Chesney, R., & Citron, D. (2019). Deepfake and the new disinformation war: The coming age of post-truth geopolitics. *Foreign Affairs*. https://perma.cc/TW6Z-Q97D

Comissão Europeia, Direção-Geral da Educação, da Juventude, do Desporto e da Cultura. (2022). Orientações para professores e educadores sobre o combate à desinformação e a promoção da literacia digital através da educação e da formação. *Serviço das Publicações da União Europeia*. https://doi.org/10.2766/218432

Convenção do Conselho da Europa para a Prevenção e o Combate à Violência Contra as Mulheres e a Violência Doméstica. (2011). *Convenção de Istambul*. https://rm.coe.int/168046253d

Cordeschi, R. (2007). AI turns fifty: Revisiting its origins. *Applied Artificial Intelligence*, *21(4–5)*, *259–279*. https://doi.org/10.1080/08839510701252304

Dalí Museum. (2019). Dalí lives (via artificial intelligence). https://thedali.org/exhibit/dali-lives/

Daniel Advogados. (2022). Relatório de impacto de inteligência artificial: AIIA – Metodologia Daniel. https://www.daniel-ip.com

Danry, V., et al. (2022). AI-generated characters: Putting deepfakes to good use. *CHI Conference on Human Factors in Computing Systems (pp. 1–5)*. Association for Computing Machinery. https://doi.org/10.1145/3491101.3503736

Delfino, R. A. (2019). Pornographic deepfakes: The case for federal criminalization of revenge porn's next tragic act. *Fordham Law Review, 88(3), 887–942*. https://ir.lawnet.fordham.edu/flr/vol88/iss3/2

Doran, G. T. (1981). There's a S.M.A.R.T. way to write management's goals and objectives. *Management Review, 70(11), 35–36*. https://community.mis.temple.edu/mis0855002fall2015/files/2015/10/S.M.A.R.T-Way-Management-Review.pdf

End Violence Against Women. (2024). Government criminalises creation of deepfakes, but with a major loophole. https://www.endviolenceagainstwomen.org.uk/government-criminalises-creation-of-deepfakes-but-with-a-major-loophole/

Europol. (2022). Facing reality? Law enforcement and the challenge of deepfakes: An observatory report from the Europol Innovation Lab. *Publications Office of the European Union*.
https://www.europol.europa.eu/publications-events/publications/facing-reality-law-enforcement-and-challenge-of-deepfakes

Ferreira, A. N. (2024). A difusão de pornografia deepfake não consentida: A nova forma de violência contra a mulher. Associação Portuguesa de Mulheres Juristas. https://www.apmj.pt/files/121/Estudos-Premiados/498/A-difusao-de-pornografia-deepfake-nao-consentida--a-nova-forma-de-violencia-contra-a-Mulher---Ana-Neto-Ferreira.pdf

Floridi, L., et al. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds & Machines, 28, 689–707*. https://doi.org/10.1007/s11023-018-9482-5

Flynn, A., et al. (2022). Deepfakes and digitally altered imagery abuse: A cross-country exploration of an emerging form of image-based sexual abuse. *The British Journal of Criminology*, *62(6)*, *1341–1358*. https://doi.org/10.1093/bjc/azac012

Gamage, D., Ravinthran, H., & Sasahara, K. (2023). Moral intuitions behind deepfake-related discussions in Reddit communities. *ArXiv.org*. https://doi.org/10.48550/arXiv.2303.06216

Godulla, A. (2022). Deepfake art: Artistic interventions in the age of synthetic media. *Media-N: Journal of the New Media Caucus*, *18(3)*, *1–16*.

Han, C., Li, A., Kumar, D., & Durumeric, Z. (2024). Characterizing the MrDeepFakes sexual deepfake marketplace. arXiv. https://arxiv.org/html/2410.11100v2

Heikkilä, M. (2023). Google DeepMind has launched a watermarking tool for AI-generated images. *MIT Technology Review*. https://www.technologyreview.com/2023/08/29/1078620/google-deepmind-has-launched-a-watermarking-tool-for-ai-generated-images/

Hoq, A., Rahman, M., Chowdhury, M., & Uddin, M. (2025). Feedback and education improve human detection of image manipulation on social media. *Media Research*, *37(2)*, *55–70*. https://doi.org/10.37016/mr-2020-175

Infopédia. (2025). Infopédia. Porto Editora. https://www.infopedia.pt/artigos/

Internet Archive Help Center. (2025). Wayback Machine general information. Internet *Archive Help Center*. https://help.archive.org/help/wayback-machine-general-information/

Internet Watch Foundation. (2024). AI CSAM report update: Prompt—From fantasy to photo-realistic reality. *Internet Watch Foundation*. https://www.iwf.org.uk/media/nadlcb1z/iwf-ai-csam-report_update-public-jul24v13.pdf

Januário, T. F. X. (2024). Inteligência artificial e abuso sexual baseado em imagens: Uma análise de deepfakes não consensuais à luz do direito penal português. *Revue Internationale de Droit Pénal*, *95(2)*, *483–497*. https://doi.org/10.466/1263-7

Jiang, Y., Li, X., Luo, H., Yin, S., & Kaynak, O. (2022). Quo vadis artificial intelligence? *Discover Artificial Intelligence, 2(1), Article 22*. https://doi.org/10.1007/s44163-022-00022-8

Junior, F., & Hessel, A. M. D. G. (2021). Entre ver e crer. *TECCOGS: Revista Digital de Tecnologias Cognitivas, 23, 79–89*. https://doi.org/10.23925/1984-3585.2021i23p79-89

Karaboga, M., Frei, N., Puppis, M., Vogler, D., Raemy, P., Ebbers, F., Runge, G., Rauchfleisch, A., de Seta, G., Gurr, G., Friedewald, M., & Rovelli, S. (2024). Deepfakes und manipulierte Realitäten - Technologiefolgenabschätzung und Handlungsempfehlungen für die Schweiz. vdf. https://doi.org/10.5281/zenodo.11643644

Kaswan, K. S., Malik, K., Dhatterwal, J. S., Naruka, M. S., & Govardhan, D. (2023). Deepfakes: A review on technologies, applications and strategies. *International Conference on Power Energy, Environment & Intelligent Control (PEEIC) (pp. 292–297)*. IEEE. https://doi.org/10.1109/PEEIC59336.2023.10450604

Kietzmann, J., Lee, L.W., McCarthy, I.P. and Kietzmann, T.C. (2020). Deepfakes: Trick or treat?. *Business Horizons, 63(2), pp. 135–146*. https://doi.org/10.1016/j.bushor.2019.11.006

Korshunov, P., & Marcel, S. (2019). Vulnerability assessment and detection of deepfake videos. *International Conference on Biometrics (ICB) (pp. 1–6)*. IEEE. https://doi.org/10.1109/ICB45273.2019.8987375

Lage, M. C., & Godoy, A. S. (2008). O uso do computador na análise de dados qualitativos: Questões emergentes. *Revista de Administração Mackenzie*, 9(4), 75–98. https://doi.org/10.1590/S1678-69712008000400006

Lai, Z., Kim, J., & Park, S. (2025). Enhancing deepfake detection: Proactive forensics techniques using digital watermarking. *Computers, Materials & Continua*, *82(1)*, *73–102*. https://doi.org/10.32604/cmc.2024.059370

Lalonde, D. (2022). Policy options on non-consensual deepnudes and sexual deepfakes. *Learning Network Brief*, *39*, *1–8*. https://www.gbvlearningnetwork.ca/our-work/briefs/brief_39.html

Levy, S. (2023). The biggest deepfake porn website is now blocked in the UK. WIRED. https://www.wired.com/story/the-biggest-deepfake-porn-website-is-now-blocked-in-the-uk/

Maia, J. (2018). O pós-humano: Ideias e problemáticas sobre a criação da inteligência artificial. https://hdl.handle.net/10316/95474

McGlynn, C. (2024). Creating sexually explicit deepfakes without consent: Options for law reform. https://e87dab74-be98-4bb1-83c5-05251d2bc6f4.usrfiles.com/ugd/e87dab_1676da131ec64bd08c34f9fbe7fb6845.pdf

McGlynn, C., Rackley, E., & Johnson, K. (2022). Shattering lives and myths: A report on image-based sexual abuse. *Durham University*. https://www.tandfonline.com/doi/epdf/10.1080/23268743.2019.1675091

McGlynn, C., & Toparlak, R. T. (2025). The "new voyeurism": Criminalizing the creation of deepfake porn. *Journal of Law and Society*. Advance online publication. https://doi.org/10.1111/jols.12527

Mesky, E. et al. (2020). Regulating Deep Fakes: Legal and Ethical Considerations. *Journal of Intellectual Property Law & Practice*, *Volume 15*, *Issue 1*, *January 2020*, *Pages 24–31*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3497144

Meta. (2025). Media manipulated media. https://transparency.fb.com/engb/policies/community-standards/manipulated-media/

Molina, A. C., & Berenguel, O. L. (2022). Deepfake: The evolution of fake news. *Research, Society and Development*, *11(6)*, *e56211629533*. https://doi.org/10.33448/rsd-v11i6.29533

Morais, F. D. B., & Castelo Branco, V. R. (n.d.). A inteligência artificial: Conceitos, aplicações e controvérsias. In Anais do XX Simpósio Internacional de Ciências Integradas da UNAERP – Campus Guarujá.

Mirsky, Y., & Lee, W. (2021). The creation and detection of deepfakes: A survey. *ACM Computing Surveys*, *54(1)*, *Article 8*. https://doi.org/10.1145/3425780

Mokadem, S. S. E. (2023) "The Efecct of Media Literacy on Misinformation and Video Detection. *Arab Media & Society*, *Issue 35*. https://doi.org/10.70090/SM23EMLM

Moreira, J. R., & Pereira Ribeiro, J. B. (2023). Letramento e competência informacional e as relações éticas na gestão da informação e do conhecimento no contexto da inteligência artificial. *Brazilian Journal of Information Science: Research Trends*, *17*, *e023047*. https://doi.org/10.36311/1981-1640.2023.v17.e023047

Moreira, M. J. (2025). A utilização de deepfakes no domínio jurídico penal: Uma nova realidade. https://inteligenciaartificialhoje.pt/wp-content/uploads/2025/04/A-utilizacao-de-deepfakes-no-dominio-juridico-penal_-uma-nova-realidade_.pdf

Newman, N., Fletcher, R., Robertson, C. T., Eddy, K., & Nielsen, R. K. (2024). Digital news report 2024. *Reuters Institute for the Study of Journalism*, *University of Oxford*. https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2024

Odeh, A. (2024). Unmasking Deepfakes: Advances in Fake Video Detection. *Revue d'Intelligence Artificielle*, *38(4)*, *1119-1131*. https://doi.org/10.18280/ria.380407

Oliveira, F. A. F. & Barroco, S. M. S. (2023). REVOLUÇÃO TECNOLÓGICA E SMARTPHONE: CONSIDERAÇÕES SOBRE A CONSTITUIÇÃO DO SUJEITO CONTEMPORÂNEO. *Psciologia em estudo*, *28*, *e51648*. https://doi.org/10.4025/psicolestud.v28i0.51648

Okolie, C. (2023). Artificial intelligence-altered videos (deepfakes), image-based sexual abuse, and data privacy concerns. *Journal of International Women's Studies*, *25(2)*, *Article 11*. https://vc.bridgew.edu/jiws/vol25/iss2/11

Ortlipp, M. (2008). Keeping and using reflective journals in the qualitative research process. *The Qualitative Report*, *13(4)*, *695–705*. https://doi.org/10.46743/2160-3715/2008.1579

Parlamento Europeu & Conselho da União Europeia. (2024, 14 de maio). Diretiva (UE) 2024/1385 do Parlamento Europeu e do Conselho, de 14 de maio de 2024, relativa ao combate à violência contra as mulheres e à violência doméstica. *Jornal Oficial da União Europeia*, *L 1385*, *1–36*. https://pessoas2030.gov.pt/legislacao/diretiva-ue-2024_1385-do-parlamento-europeu-e-do-conselho/

Patel. Y. et al. (2023). "Deepfake Generation and Detection: Case Study and Challenges". IEEE Access, vol. 11, pp. 143296-143323. https://doi.org/10.1109/ACCESS.2023.3342107

Patton, M. Q. (2015). Qualitative research & evaluation methods (3rd ed.). Sage. https://aulasvirtuales.files.wordpress.com/2014/02/qualitative-research-evaluation-methods-by-michael-patton.pdf

Prado, M. P. (2021). Deepfake de áudio: Manipulação simula voz real para retratar alguém dizendo algo que não disse. TECCOGS – *Revista Digital de Tecnologias Cognitivas*, *23*, *45–68*. https://revistas.pucsp.br/index.php/teccogs/article/download/55977/37926/169047

Relatório Anual de Segurança Interna. (2024). Relatório anual de segurança interna (RASI) 2024. *República Portuguesa*. https://www.portugal.gov.pt/pt/gc24/comunicacao/documento?i=relatorio-anual-de-seguranca-interna-rasi-2024

Rodrigues, P. G. (2023). Deepfakes pornográficas não consensuais: A busca por um modelo de criminalização. *Revista Brasileira de Ciências Criminais*, *199*, *277–311*. https://doi.org/10.5281/zenodo.8380977

Ruiter, A. (2021). The distinct wrong of deepfakes. *Philosophy & Technology*, *34*, *1311–1332*. https://doi.org/10.1007/s13347-021-00459-2

Sensity AI. (2024). State of deepfakes 2024. Sensity AI. https://5865987.fs1.hubspotusercontent-na1.net/hubfs/5865987/SODF%202024.pdf

Silva, S. D. & Costa, F. J. (2014). Mensuração e Escalas de Verificação: uma Análise Comparativa das Escalas de Likert e Phrase Completion. *PMKT – Revista Brasileira de Pesquisas de Marketing, Opinião e Mídia, São Paulo, Brasil, v. 15, p. 1-16.* https://revistapmkt.com.br/wp-content/uploads/2022/01/1_Mensuracao-e-Escalas-de-Verificacao-uma-Analise-Comparativa-das-Escalas-de-Likert-e-Phrase-Completion-1.pdf

Stanciu, A. & Ciuperca, E. (2024). Can Deepfakes Benefit the Metaverse in an Era of Disinformation? Insights from a Systematic Review. https://doi.org/10.1016/j.ifacol.2024.07.125

StopNCII.org. (2025). Stop non-consensual intimate image abuse. https://stopncii.org/

Supremo Tribunal Federal. (2024). Guia ilustrado contra as deepfakes. Data Privacy Brasil; Supremo Tribunal Federal. https://portal.stf.jus.br/desinformacao/doc/Guia%20ilustrado%20Contra%20DeepFakes_ebook%20(1).pdf

Sykes, G. M., & Matza, D. (1957). Techniques of neutralization: A theory of delinquency. American Sociological Review, 22, 664–670. https://doi.org/10.2307/2089195

Texas Legislature. (2019). An act relating to the creation of the criminal offense of using deepfake video to influence the outcome of an election (S.B. 751, 86th Leg., Reg. Sess.). https://capitol.texas.gov/tlodocs/86R/billtext/html/SB00751F.htm

Thakur, R., & Rohilla, R. (2020). Recent advances in digital image manipulation detection techniques: A brief review. *Forensic Science International, 312, 110311*. https://doi.org/10.1016/j.forsciint.2020.110311

Theodorson, G. A., & Theodorson, A. G. (1969). A modern dictionary of sociology. C. E. Merrill.

https://ia903407.us.archive.org/33/items/in.ernet.dli.2015.137525/2015.137525.A-Modern-Dictionary-Of-Sociology.pdf

Umbach, R., Henry, N., Beard, G., & Berryessa, C. (2024). Non-consensual synthetic intimate imagery: Prevalence, attitudes, and knowledge in 10 countries. *Association for Computing Machinery*. https://doi.org/10.1145/3613904.3642382

Umur Aybars Çiftçi, Demir, İ., & Yin, L. (2023). Deepfake source detection in a heartbeat. *The Visual Computer*, *39*, *2193–2207*. https://doi.org/10.1007/s00371-023-02981-0

United Nations Population Fund. (2025a). 16 days of activism against gender-based violence. https://www.unfpa.org/pt/thevirtualisreal

United Nations Population Fund. (2025b). Technology-facilitated gender-based violence: A growing threat. https://www.unfpa.org/TFGBV

United Nations Interregional Crime and Justice Research Institute. (2025). Malicious uses and abuses of artificial intelligence. https://unicri.org/node/3278

United States Congress. (2025). Tools to Address Known Exploitation by Immobilizing Technological Deepfakes on Websites and Networks Act (S. 146, 119th Cong.). https://www.congress.gov/bill/119th-congress/senate-bill/146

U.S. Department of Homeland Security. (2021). Increasing threat of deepfake identities. https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf

Vieira, L. (2024). Responsabilidade civil por danos causados por sistemas de inteligência artificial. *Handle.net*. http://hdl.handle.net/11067/7869

Xu, Z., et al. (2025). Public perception towards deepfake through topic modelling and sentiment analysis of social media data. *Soc. Netw. Anal. Min.* 15, 16 https://doi.org/10.1007/s13278-025-01445-8

ANEXOS

Anexo A – Exemplo de questionário

No âmbito do presente Projeto de Graduação, nomeado de "Quando a Tecnologia Viola a Intimidade: O Caso dos Deepfakes Sexuais", que visa a obtenção do grau de Licenciatura em Criminologia, pela Universidade Fernando Pessoa. O presente questionário, tem como objetivo recolher informações iniciais sobre o perfil, motivações e padrões de uso de ferramentas deepfake entre indivíduos que criaram ou partilharam deepfakes sexuais não consensuais.

Este projeto foi aprovado pela Comissão de Ética da Universidade Fernando Pessoa, de modo a assegurar o respeito dos direitos aos direitos, à privacidade, e ao bem-estar de todos os participantes demonstrando publicamente o compromisso ético do estudo.

Com isto, solicita-se a sua colaboração nesta fase, relembrando que é voluntária, anónima e confidencial. O preenchimento deste inquérito tem a duração estimada de 10 minutos. Além disso, evidencia-se que não existem respostas certas ou erradas, apenas a sua experiência e opinião importam.

Antes de prosseguir, pedimos que leia atentamente as informações fornecidas sobre a voluntariedade da participação, o anonimato dos dados e o seu direito de desistir a qualquer momento, sem necessidade de justificativa. Se decidir participar, solicito que responda "Sim" à última questão, o que equivale ao seu consentimento informado para utilização anónima dos dados neste projeto.

Agradeço sinceramente a sua colaboração.

Inês Sales – 2022122003@ufp.edu.pt

Concorda em participar neste inquérito?

- ☐ Sim
- ☐ Não

I – Caracterização demográfica

Qual é a sua nacionalidade? _____

Qual é a sua idade?

- 18-25
- 26-35
- 36-50
- >50

Qual é o seu género?

- Masculino
- Feminino
- Não binário
- Outro: _____

II – Experiência e envolvimento com ferramentas de Deepfake e contexto de compartilhamento

Qual o seu envolvimento técnico?

- Crio deepfakes
- Divulgo deepfakes
- Ambos

Há quanto tempo utiliza software de criação manipulação de vídeos ou imagens?

- <1 ano
- 1-3 anos
- >3 anos
- Nunca usei

Em média, com que frequência utiliza essas ferramentas?

- Todos os dias
- 1 vez por semana
- 1 vez por mês
- Outro: _____

Em qual destes canais costuma publicar ou compartilhar deepfakes?

- Grupos privados de Telegram

- ☐ Subreddits
- ☐ Perfis ou páginas no X
- ☐ Fóruns especializados (ex: MrDeepFakes)
- ☐ Mensagens diretas
- ☐ Outro: _____

Quais dos seguintes programas ou plataformas já utiliza para criar deepfakes?

- ☐ FaceSwap
- ☐ DeepNude
- ☐ Zao
- ☐ Aplicação online gratuita. Qual?_____
- ☐ Outro:_____

Tem algum conhecimento sobre alguma lei ou política de plataformas que proíbam DSNC?

- ☐ Sim
- ☐ Não

Se sim, cite qual(is):_____

[…]

III. Motivações

Por favor, assinale o grau de concordância [1- discordo completamente; 2 – discordo; 3 – prefiro não dizer; 4- concordo; 5 – concordo completamente]

	1	2	3	4	5
"Crio deepfakes para obter gratificação erótica"					
"Partilho deepfakes como forma de entretenimento"					
"Procuro controlo e poder ao produzir o conteúdo"					
"Busco reconhecimento ao produzir este conteúdo"					

"Crio ou partilho deepfakes em troca de ganhos financeiros"					
["…"]					

V. Disponibilidade para entrevista

Se concorda em continuar para uma entrevista online de cerca de 60 minutos, insira um pseudónimo ou e-mail (opcional): _____

Anexo B – Exemplo de roteiro de entrevista

Duração: 60 minutos

Objetivo: explorar em profundidade o processo de criação partilha e racionalização de DSNC

1. **Abertura e consentimento**
   a. Reforçar anonimato e voluntariedade.
   b. Confirmar assinatura do termo de consentimento.
   c. Explicar breve visão geral do que será perguntado.
2. **Perfil e Contexto**
   a. "Conte-me, há quanto tempo e por que começou a usar ferramentas de deepfake"
   b. "Como aprendeu a utilizá-las? Teve ajuda de alguém ou aprendeu por conta própria?"
   c. …
3. **Processo Técnico de Criação**
   a. "Descreva passo a passo o seu fluxo de trabalho quando cria um deepfake (ex: software, tempo gasto, fontes de imagem/áudio)"
   b. "Que critérios usa para escolher a face e o corpo que manipula?"
   c. "Como avalia a "qualidade" e o "realismo" antes da partilha efetiva do vídeo?"
   d. …
4. **Canais de Difusão e Dinâmicas Comunitárias**
   a. "Em quais plataformas costuma publicar ou partilhar deepfakes? Alguma razão para escolher essas?"

b. "Como funciona a interação com outros usuários nesses espaços? Pode dar um exemplo de como recebe feedback (elogios, críticas, pedidos)?"
c. "Existe algumas normas de partilha grupo para partilhar esse conteúdo?"
d. …

5. Motivações e Justificativas

a. "Quais foram ou são suas principais motivações ao criar e partilhar deepfakes?"

b. "Como justifica para si mesmo ou para os outros a prática de compartilhar esse conteúdo sem consentimento?"

c. "Já discutiu com alguém a legalidade ou a ética disso? O que costuma responder?"

d. …

6. Encerramento

a. "Ainda há algo que gostaria de acrescentar e que não perguntamos?"

b. "Agradeço a sua participação. Se desejar revisar algum trecho da transcrição ou retirar alguma informação, avise-me em até sete dias."