

## DEEPFAKES E PROVA DIGITAL NA VIOLÊNCIA DOMÉSTICA: OS RISCOS DA INTELIGÊNCIA ARTIFICIAL

Pedro Miguel Freitas 

Universidade Católica Portuguesa 

Ana Guerreiro 

University of Maia – UMaia 

**Contextualização:** Os avanços recentes da inteligência artificial generativa têm introduzido novos desafios ao sistema de justiça penal, em particular no domínio da prova digital. Entre as aplicações mais problemáticas destacam-se os deepfakes, suscetíveis de induzir em erro quanto à sua autenticidade, especialmente em contextos de maior fragilidade probatória, como os crimes de violência doméstica.

**Objetivos:** O presente estudo tem como objetivo geral compreender as perceções, experiências e práticas dos/as magistrados/as relativamente à admissibilidade e valoração da prova digital, com especial enfoque nos meios de prova audiovisual e nos riscos associados à sua eventual manipulação por tecnologias de inteligência artificial generativa.

**Métodos:** Foi adotada uma abordagem qualitativa, recorrendo-se à realização de entrevistas semiestruturadas a oito magistrados/as, com experiência profissional em processos de violência doméstica.

**Resultados:** Observaram-se limitações na deteção de elementos audiovisuais gerados ou manipulados por inteligência artificial. A autenticidade deste tipo de elementos (deepfakes) granjearam mais credibilidade que elementos autênticos. De todo o modo, embora a prova audiovisual seja considerada relevante, o depoimento da vítima continua sendo considerado a “prova rainha”.

**Palavras-chave:** Prova audiovisual; valoração; deepfakes; autenticidade; inteligência artificial.

## DEEPFAKES Y PRUEBA DIGITAL EN LA VIOLENCIA DOMÉSTICA: LOS RIESGOS DE LA INTELIGENCIA ARTIFICIAL

**Contextualización:** Los avances recientes de la inteligencia artificial generativa han introducido nuevos desafíos al sistema de justicia penal, particularmente en el ámbito de la prueba digital. Entre las aplicaciones más problemáticas destacan los deepfakes, susceptibles de inducir a error en cuanto a su autenticidad, especialmente en contextos de mayor fragilidad probatoria, como los delitos de violencia doméstica.

**Objetivos:** El presente estudio tiene como objetivo general comprender las percepciones, experiencias y prácticas de los magistrados en relación con la admisibilidad y valoración de la prueba digital, con especial enfoque en los medios de prueba audiovisual y en los riesgos asociados a su eventual manipulación por tecnologías de inteligencia artificial generativa.

**Método:** Se adoptó un enfoque cualitativo, recurriendo a la realización de entrevistas semiestructuradas a ocho magistrados con experiencia profesional en procesos de violencia doméstica.

**Resultados:** Se observaron limitaciones en la detección de elementos audiovisuales generados o manipulados por inteligencia artificial. La autenticidad de este tipo de elementos (deepfakes) ha granjeado más credibilidad que los elementos auténticos. De todos modos, aunque la prueba audiovisual se considera relevante, el testimonio de la víctima sigue considerándose la «prueba reina».

**Palabras clave:** Prueba audiovisual; valoración de la prueba; deepfakes; autenticidad; inteligencia artificial.

## DEEPFAKES AND DIGITAL EVIDENCE IN DOMESTIC VIOLENCE: THE RISKS OF ARTIFICIAL INTELLIGENCE

**Contextualization:** Recent advances in generative artificial intelligence have introduced new challenges to the criminal justice system, particularly within the realm of digital evidence. Notable among the most problematic applications are deepfakes, which are liable to mislead regarding their authenticity, especially in contexts of significant evidentiary fragility, such as domestic violence crimes.

**Objectives:** The general objective of this study is to understand the perceptions, experiences, and practices of magistrates regarding the admissibility and assessment of digital evidence, with a special focus on audiovisual evidence and the risks associated with its potential manipulation by generative artificial intelligence technologies.

**Methods:** A qualitative approach was adopted, utilizing semi-structured interviews with eight magistrates possessing professional experience in domestic violence proceedings.

**Results:** Limitations were observed regarding the detection of audiovisual elements generated or manipulated by artificial intelligence. These types of elements (deepfakes) garnered more credibility than authentic elements. Nevertheless, although audiovisual evidence is considered relevant, the victim's testimony continues to be regarded as the "queen of proofs".

**Keywords:** Audiovisual evidence; evidentiary assessment; deepfakes; authenticity; artificial intelligence.

## INTRODUÇÃO

Os avanços recentes da inteligência artificial generativa (GenAI) têm vindo a produzir transformações profundas em diversos domínios da vida social, levantando novos e complexos desafios, nomeadamente para o Direito. Entre as múltiplas possibilidades proporcionadas por estas tecnologias, destacam-se os modelos capazes de gerar ou manipular conteúdos textuais, visuais e sonoros com elevado grau de realismo, cuja utilização se generalizou para além de contextos académicos ou técnicos. A crescente acessibilidade a estas ferramentas tem intensificado o debate jurídico em torno dos riscos associados à produção de conteúdos potencialmente enganadores, com particular relevância para o funcionamento do sistema de justiça.

Um dos fenómenos mais emblemáticos neste contexto é o dos *deepfakes*, entendidos como conteúdos de imagem, áudio ou vídeo total ou parcialmente gerados ou manipulados por inteligência artificial, suscetíveis de induzir em erro quanto à sua autenticidade. Atendendo ao papel central que a prova digital assume hoje na investigação criminal, a emergência de tecnologias capazes de produzir representações hiper-realistas coloca em causa o tradicional princípio do *seeing is believing*, introduzindo um risco acrescido de manipulação probatória. Este risco revela-se particularmente desafiante em contextos de maior vulnerabilidade probatória, como é o caso dos crimes de violência doméstica.

É precisamente neste quadro que se insere o presente estudo, que visa compreender em que medida os/as magistrados/as, enquanto atores centrais do sistema de justiça penal, são suscetíveis de serem convencidos por *deepfakes* em contextos de escassa ou difícil prova, como os de violência doméstica. Através de uma abordagem qualitativa, com recurso a entrevistas semiestruturadas, pretende-se contribuir para uma reflexão crítica sobre os desafios que a inteligência artificial generativa coloca ao sistema de justiça penal e sobre a adequação das respostas institucionais face à rápida evolução tecnológica.

## 1. A INTELIGÊNCIA ARTIFICIAL E O SISTEMA DE JUSTIÇA

Os rápidos avanços da inteligência artificial generativa sentidos nos últimos anos ultrapassaram os muros académicos e de nichos sectoriais, passando a integrar o dia a dia das empresas e pessoas. Provavelmente a maior referência popular no domínio da inteligência artificial generativa é o ChatGPT. Este modelo de IA, criado pela OpenAI em 2018, com o nome

de GPT-1<sup>1</sup>, foi sendo desenvolvido (GPT-2 em 2019<sup>2</sup> e GPT-3 em 2020<sup>3</sup>) até ao lançamento para o público em geral em novembro de 2022 do ChatGPT<sup>4</sup> (GPT-3.5). Em cerca de um mês, o número de utilizadores ativos por semana (*Weekly Active Users*) atingiu a marca de 1 milhão<sup>5</sup>, estimando-se que atualmente chegue aos 800 milhões de utilizadores<sup>6</sup>. Os principais usos do ChatGPT vão desde a geração de textos, sumarização, tradução até à programação ou criação de imagens. Na verdade, nas últimas versões, o ChatGPT transitou de um mero LLM – *Large Language Model* – para um LMM – *Large Multimodal Model* –, capaz de receber e produzir texto, áudio ou imagens.

O ChatGPT é uma de muitas manifestações do que se convencionou denominar inteligência artificial generativa ou GenAI. Este é um termo genérico para certos modelos de aprendizagem automática (*machine learning*)<sup>7</sup>, conhecidos por modelos de aprendizagem profunda<sup>8</sup> (*deep learning*), que são inspirados pela biologia humana e pela estrutura do nosso cérebro humano. São compostos por redes neuronais artificiais<sup>9</sup> multicamadas (*multilayered*)

---

<sup>1</sup> BELCIC, Ivan; STRYKER, Cole. **What is GPT** (generative pretrained transformer)? IBM, [S.l.], [S.d.]. Disponível em: <https://www.ibm.com/think/topics/gpt>. Acesso em: 03 dez. 2025.

<sup>2</sup> OPENAI. Fine-tuning GPT-2 from human preferences. **OpenAI**, [S.l.], 2019. Disponível em: <https://openai.com/index/fine-tuning-gpt-2/>. Acesso em: 03 dez. 2025.

<sup>3</sup> OPENAI. OpenAI API. **OpenAI**, [S.l.], 2020. Disponível em: <https://openai.com/index/openai-api/>. Acesso em: 03 dez. 2025.

<sup>4</sup> OPENAI. Introducing ChatGPT. **OpenAI**, [S.l.], 2022. Disponível em: <https://openai.com/index/chatgpt/>. Acesso em: 03 dez. 2025.

<sup>5</sup> EDMONDS, Lauren. 3 years old, 800 million users, 29,000 prompts a second: ChatGPT's meteoric rise, by the numbers". **Business Insider**, [S.l.], 2025. Disponível em: <https://www.businessinsider.com/chatgpt-by-the-numbers-2025-11#chat-gpt-has-800-million-weekly-active-users-2>. Acesso em: 03 dez. 2025.

<sup>6</sup> SOR, Jennifer. Sam Altman touts ChatGPT's 800 million weekly users, double all its main competitors combined. **Business Insider**, 2025. Disponível em: <https://www.businessinsider.com/chatgpt-users-openai-sam-altman-devday-llm-artificial-intelligence-2025-10>. Acesso em: 03 dez. 2025.

<sup>7</sup> De acordo com a Google, machine learning é um "subconjunto da inteligência artificial que permite que um sistema aprenda e melhore autonomamente usando redes neurais e aprendizagem profunda, sem ser explicitamente programado, alimentando-o com grandes quantidades de dados" (GOOGLE. What is machine learning (ML)? **Google**, [S.l.], [S.d.]. Disponível em: <https://cloud.google.com/learn/what-is-machine-learning?hl=en>. Acesso em: 03 dez. 2025). Embora a terminologia tecnológica esteja em constante evolução, são habitualmente assinalados quatro a cinco tipos de machine learning: Supervised learning, unsupervised learning, reinforcement learning, self-supervised machine learning, e semi-supervised learning. Para mais desenvolvimentos, consultar CHINA, Chrystal R. Five machine learning types to know. In: Cole STRYKER et al. (eds.). **The 2025 Guide to Machine Learning**. [S.l.]: IBM, 2025. Disponível em: <https://www.ibm.com/think/topics/machine-learning-types>. Acesso em: 01 nov. 2025 e GOOGLE. What is machine learning (ML)? De interesse, DURÃES, Dalila; FREITAS, Pedro Miguel e NOVAIS, Paulo. The Relevance of Deepfakes in the Administration of Criminal Justice. In: ANTUNES, Henrique Sousa et al. (eds.). **Multidisciplinary Perspectives on Artificial Intelligence and the Law**. Cham: Springer, 2024, p. 354 e ss.

<sup>8</sup> STRYKER, Cole et al. What is deep learning? In: Cole STRYKER et al. (eds.). **The 2025 Guide to Machine Learning**. [S.l.]: IBM, 2025. Disponível em: <https://www.ibm.com/think/topics/deep-learning>. Acesso em: 30 nov. 2025.

<sup>9</sup> LEE, Fangfang. What is a neural network? In: Cole STRYKER et al. (eds.). **The 2025 Guide to Machine Learning**. [S.l.]: IBM, 2025. Disponível em: <https://www.ibm.com/think/topics/neural-networks>. Acesso em: 30 nov. 2025.

– pelo menos quatro camadas<sup>10</sup> – que são treinados com enormes conjuntos de dados, procurando identificar entidades, padrões<sup>11</sup> e relações, para depois serem capazes de produzir conteúdo a partir de um *input* ou *prompt*<sup>12</sup>.

Dito de uma forma mais simples, a inteligência artificial generativa emprega algoritmos complexos para gerar texto, código, áudio, imagens ou vídeo em resposta a um *prompt*. Este aspecto característico, assente no *output* mais do que na exata técnica que permite gerar esse mesmo *output* (v.g. *variational autoencoders*, *generative adversarial networks*<sup>13</sup>, *diffusion models* ou *transformers*), está habitualmente presente na definição de inteligência artificial generativa. A título de exemplo, os considerandos 99 e 105 do Regulamento Europeu sobre Inteligência Artificial (EU AI Act)<sup>14</sup> aludem à inteligência artificial generativa como modelos que permitem a geração de conteúdos, nomeadamente texto, áudio, imagens ou vídeo<sup>15</sup>.

Um modo de exemplificação dos diferentes modelos de inteligência artificial generativa e da sua versatilidade, baseia-se na combinação “tipo de *input*” - “tipo de *output*”<sup>16</sup>. De acordo com este critério, existem, entre outros, modelos text-to-text, text-to-image, text-to-video, text-to-audio, text-to-code, audio-to-audio, audio-to-text, brain-to-text ou image-to-text. As versões iniciais do ChatGPT por serem capazes apenas de gerar texto a partir de um *input* de texto eram classificadas então como *text-to-text*.

Embora o ritmo de desenvolvimento de soluções de inteligência artificial imponha um constante processo de atualização sobre potenciais aplicações, benefícios e riscos, pode afirmar-se que, no estado atual das coisas, os benefícios comumente apontados centram-se em ganhos de eficiência e produtividade, automatização de tarefas repetitivas e intensas

---

<sup>10</sup> GOOGLE. What is machine learning (ML)?.

<sup>11</sup> EUROPOL. **Facing reality?** Law enforcement and the challenge of deepfakes. Luxemburgo: Publications Office of the European Union, 2022, p. 7.

<sup>12</sup> De acordo com o Collins Dictionary, uma das aceções de *prompt* corresponde a uma instrução oral ou escrita dada por um utilizador, usando linguagem natural, a um sistema de inteligência artificial. Cf. COLLINS DICTIONARY. **Collins Online Dictionary**: definitions, thesaurus and translations. Disponível em: <https://www.collinsdictionary.com>. Acesso em: 03 dez. 2025.

<sup>13</sup> Acerca da importância das *generative adversarial networks* (redes adversárias generativas) no aumento da qualidade e acessibilidade dos *deepfakes*, para o qual contribuiu o trabalho seminal de Ian Goodfellow em 2014, EUROPOL. **Facing reality?** Law enforcement and the challenge of deepfakes. p. 5. Ver igualmente sobre GANs, DURÃES, Dalila; FREITAS, Pedro Miguel; NOVAIS, Paulo. The Relevance of Deepfakes in the Administration of Criminal Justice. p. 362.

<sup>14</sup> Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024, que cria regras harmonizadas em matéria de inteligência artificial.

<sup>15</sup> Veja-se também UNIVERSIDADE DE HARVARD. The Benefits and Limitations of Generative AI: Harvard Experts Answer Your Questions. **Harvard**, 2023. Disponível em: <https://harvardonline.harvard.edu/blog/benefits-limitations-generative-ai>. Acesso em: 27 nov. 2025.

<sup>16</sup> GOOGLE. What is Machine Learning? In: GOOGLE. **Introduction to Machine Learning**. [S.l.], 2025. Disponível em: <https://developers.google.com/machine-learning/intro-to-ml/what-is-ml>. Acesso em: 27 nov. 2025.

do ponto de vista de mão-de-obra, ou na sua capacidade de personalização elevada de conteúdos<sup>17</sup>. Mas são vastos os riscos e os desafios colocados pela inteligência artificial em geral, e a generativa em particular, que têm merecido especial atenção de juristas, eticistas, filósofos e cientistas da computação.

Sem pretensão alguma de sermos exaustivos, podemos referenciar como preocupações ético-jurídicas associadas ao uso da inteligência artificial (generativa) os vieses e discriminação, a desinformação, manipulação e *fake news*, a tutela da privacidade, a proteção de direitos autorais, a opacidade ou falta de transparência, a concentração de poder, as desigualdades socioeconómicas, questões ambientais e os riscos existenciais<sup>18</sup>. Além destas preocupações, ou de entre delas, emerge o fenómeno dos *deepfakes*, como exemplo de uma aplicação prática da inteligência artificial que, dependendo do motivo da sua criação numa situação em concreto, constitui um instrumento ao serviço da desinformação, da manipulação, da criminalidade (ciber ou tradicional), quando não mesmo da erosão dos pilares da democracia.

O termo *deepfake* resulta da aglutinação das palavras *deep learning* e *fake media*<sup>19</sup> e pode ser definido como “conteúdos de imagem, áudio ou vídeo gerados ou manipulados por IA, que sejam semelhantes a pessoas, objetos, locais, entidades ou acontecimentos reais, e que possam levar uma pessoa a crer, erroneamente, que são autênticos ou verdadeiros”<sup>20</sup>. Constitui uma aplicação específica de inteligência artificial generativa que se concretiza na criação de conteúdos audiovisuais hiper-realistas, seja através da troca de faces numa foto ou vídeo (*face replacement*), da manipulação das características faciais de um indivíduo para que, num vídeo, pareça estar a dizer ou fazer algo que não

---

<sup>17</sup> Cf. STRYKER, Cole; SCAPICCHIO, Mark. What is generative AI? In: Cole STRYKER et al. (eds.). **The 2025 Guide to Machine Learning**. [S.l.]: IBM, 2025. Disponível em: <https://www.ibm.com/think/topics/generative-ai>. Acesso em: 27 nov. 2025 e UNIVERSIDADE DE LEEDS. What is Generative AI? **University of Leeds**, Leeds, [S.d.]. Disponível em: <https://generative-ai.leeds.ac.uk/intro-gen-ai/strengths-and-weaknesses/>. Acesso em: 27 nov. 2025.

<sup>18</sup> STRYKER, Cole; SCAPICCHIO, Mark. What is generative AI? e UNIVERSIDADE DE LEEDS. What is Generative AI?

<sup>19</sup> TREND MICRO RESEARCH. **Malicious Uses and Abuses of Artificial Intelligence**. Europol's European Cybercrime Centre, 2020. p. 52. Para TWOMEY, John et al. What Is So Deep About Deepfakes? A Multi-Disciplinary Thematic Analysis of Academic Narratives About Deepfake Technology. **IEEE Transactions on Technology and Society**, v. 6, n. 1, p. 64-79, 2025, a palavra *deepfake* resulta somente da junção de *deep learning* e *fake*. Em WHITTAKER, Lucas et al. Mapping the deepfake landscape for innovation: A multidisciplinary systematic review and future research agenda. **Technovation**, v. 125, p. 01-17, 2023, encontra-se uma revisão sistemática sobre o conceito de *deepfake* de que resultou a seguinte síntese: “Deepfakes are synthetic media generated using artificial intelligence and deep learning technology which produce realistic yet fake representations of people undertaking actions or saying words in the form of video, image, or audio content”.

<sup>20</sup> Artigo 3.º, n.º 60 do EU AI Act. Cf. MARAS, Marie-Helen; ALEXANDROU, Alex. Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake vídeos. **The International Journal of Evidence & Proof**, v. 23, n. 3, p. 255-262, 2018 e TWOMEY, John et al. What Is So Deep About Deepfakes? A Multi-Disciplinary Thematic Analysis of Academic Narratives About Deepfake Technology. p. 64.

aconteceu (*face re-enactment*), da geração de rostos (*face generation*), corpos ou partes do corpo em imagens ou vídeos, ou da clonagem de voz (*speech synthesis*)<sup>21</sup>.

Ainda que seja admissível a criação de *deepfakes* para fins lícitos, v.g. cópia autorizada da voz de uma pessoa para fins publicitários, o sentido original ou estrito de *deepfake*<sup>22</sup> remete-nos para a produção e divulgação de dados sintéticos – conteúdos total ou parcialmente gerados ou manipulados por inteligência artificial – para fins ilícitos, de entre os quais a geração de material pornográfico não consensual<sup>23</sup>.

Perscrutando a origem dos *deepfakes* constata-se que o momento central relativamente à sua proliferação deu-se em 2017, com um membro do fórum *Reddit*, de alcunha *deepfakes*, que usando inteligência artificial criou vídeos pornográficos realistas em que as caras de atrizes de Hollywood eram colocadas no corpo de atrizes pornográficas<sup>24</sup>. Indubitavelmente, a pornografia não consensual continua a ser o tipo de *deepfake* mais divulgado, mas o seu potencial lesivo não se fica por aqui.

De acordo com Robert Chesney e Danielle K. Citron <sup>25</sup>, os usos negativos dos *deepfakes* podem ser agrupados em a) danos contra indivíduos e organizações (exploração [roubo de identidade, burla, extorsão, ameaça, coação, gratificação sexual] e sabotagem [danos reputacionais]); e b) danos contra a sociedade (distorção do discurso democrático, manipulação de eleições, erosão da confiança nas instituições, exacerbamento de divisões sociais, enfraquecimento da segurança pública, disrupção da diplomacia, comprometimento da segurança nacional, comprometimento do jornalismo, a descredibilização da verdade [*liar's dividend*]). Já John Twomey *et al.*<sup>26</sup> optam por agrupar os riscos dos *deepfakes* em três domínios: a) danos pessoais resultantes de *deepfakes* pornográficos difamatórios; b) danos societários de *deepfakes* políticos difamatórios; e c) danos epistémicos de *deepfakes* em formato vídeo (danos à verdade e prova).

<sup>21</sup> Soluções tecnológicas recentes permitem a criação de vídeos com áudio (sincronizado) – p. ex. a Veo 3 da Google – ou a substituição facial em tempo real.

<sup>22</sup> EUROPOL. **Facing reality?** Law enforcement and the challenge of deepfakes. p. 5.

<sup>23</sup> Segundo Marília Papaléo Gagliardi, 96% dos vídeos deepfake têm uma natureza pornográfica e as mulheres como alvo (GAGLIARDI, Marília Papaléo. Artificial Intelligence and Women's Rights. In: QUINTAVALLA, Alberto e TEMPERMAN, Jeroen (eds.). **Artificial Intelligence and Human Rights**. Oxford: OUP Oxford, 2023. p. 237).

<sup>24</sup> Por exemplo, Taylor Swift, Scarlett Johansson, Aubrey Plaza, Gal Gadot, e Maisie Williams. Cf. TREND MICRO RESEARCH. **Malicious Uses and Abuses of Artificial Intelligence**. p. 53; WAGNER, Travis L.; BLEWER, Ashley. The Word Real Is No Longer Real: Deepfakes, Gender, and the Challenges of AI-Altered Video. **Open Information Science**, v. 3, n. 1, 2019, p.33 e TWOMEY, John *et al.* What Is So Deep About Deepfakes? A Multi-Disciplinary Thematic Analysis of Academic Narratives About Deepfake Technology. p. 65.

<sup>25</sup> CHESNEY, Robert; CITRON, Danielle K. Deep Fakes: A looming challenge for privacy, democracy, and national security. **California Law Review**, v. 107, n. 6, p. 1771-1786, 2019.

<sup>26</sup> TWOMEY, John *et al.* What Is So Deep About Deepfakes? A Multi-Disciplinary Thematic Analysis of Academic Narratives About Deepfake Technology. p. 70 e ss.

É precisamente o perigo epistémico causado pela existência de *deepfakes* e o seu potencial impacto no sistema de justiça penal, no que à autenticidade e valor da prova diz respeito, que fundamenta o nosso estudo. Embora John Twomey *et al.* se centrem nos perigos associados a *deepfakes* em formato vídeo, não apenas este género de *deepfakes* é susceptível de ser submetido como prova de existência de um crime ou da identidade do seu autor.

Dado que, na nossa “vida diária, as pessoas confiam na sua percepção para guiá-las e dizer-lhes o que é real ou não (...) [e que fotografias e vídeos são informações importantes para o trabalho policial e prova em tribunal]”<sup>27</sup>, há um risco claro associado aos *deepfakes*, mais concretamente o de falsificação ou manipulação de prova digital como áudio, imagem ou vídeo. A prevalência da prova digital na investigação criminal nos dias de hoje<sup>28</sup> e a maior confiança depositada naquilo que é visualizado ou escutado em detrimento do que é expresso verbalmente<sup>29</sup>, *v.g.* prova testemunhal, constitui um quadro facilitador para a permeabilidade do sistema de justiça penal a representações sintéticas dos factos e da realidade operadas por *deepfakes*.

Embora a falsificação de provas não seja um fenómeno novo – pense-se na alteração de fotografias com software como o Photoshop –, a democratização do acesso a modelos de inteligência artificial avançados, capazes de fabricar, em poucos segundos e com relativa facilidade, imagens, vídeos ou áudios falsos, mas hiper-realistas, distancia os *deepfakes* de adventos tecnológicos anteriores ou mesmo contemporâneos. Se para alguns Autores, os *deepfakes* são “more frightening than Photoshop on steroids”<sup>30</sup>, a razão desta caracterização prende-se, desde logo, com a sua credibilidade (*believability*), que os tornam praticamente indistinguíveis de um vídeo, imagem ou áudio autênticos. Com isto, o *seeing is believing*<sup>31</sup> fica inevitavelmente em xeque. Especialmente se se tomar em consideração o avanço galopante da tecnologia e o risco de, num futuro próximo, o “testimony of digital media forensic experts may not be enough to authenticate evidence, because even expert witnesses may not be able to discern the modifications made to digital videos”<sup>32</sup>.

Deste ponto de vista, os *deepfakes* são um relevante poluidor do ambiente

---

<sup>27</sup> EUROPOL. **Facing reality?** Law enforcement and the challenge of deepfakes. p. 5.

<sup>28</sup> De acordo com estimativas da Comissão Europeia, a prova digital está presente em 85% das investigações criminais atuais (Cf. COMISSÃO EUROPEIA. **Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões:** Roteiro para o acesso lícito e efetivo aos dados para efeitos da aplicação da lei. Bruxelas, 2025, p. 1).

<sup>29</sup> MARAS, Marie-Helen; ALEXANDROU, Alex. Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake vídeos. p. 257.

<sup>30</sup> DIXON JR., Herbert B. DeepFakes: More frightening than photoshop on steroids. **Judges' Journal**, v. 58, n. 3, p. 35–37, 2019.

<sup>31</sup> DIXON JR., Herbert B. DeepFakes: More frightening than photoshop on steroids. p. 35.

<sup>32</sup> MARAS, Marie-Helen; ALEXANDROU, Alex. Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake vídeos. p. 259.

epistémico<sup>33</sup>, com a capacidade de manipulação da realidade e erosão da confiança nas informações extraídas, direta ou indiretamente<sup>34</sup>, da prova digital. Em que medida são os atores do sistema de justiça penal suscetíveis de serem convencidos pelos *deepfakes* em contextos de escassa ou difícil prova, como os casos de violência doméstica, é a questão a que tentamos responder com o estudo que se segue.

## 2. CONSIDERAÇÕES METODOLÓGICAS

O presente estudo teve como objetivo geral conhecer as percepções, experiências, práticas e necessidades dos/as magistrados relativamente à admissibilidade das provas digitais. De forma mais concreta, pretendeu-se: i) identificar a frequência com que os magistrados se deparam com dúvidas sobre a autenticidade digital; ii) identificar as estratégias ou critérios utilizados na avaliação da admissibilidade das provas digitais; iii) compreender até que ponto consideram necessário adaptar os procedimentos legais ou receber formação específica sobre as provas digitais e a sua fragilidade.

De modo a responder aos objetivos propostos, foram delineadas as seguintes questões de investigação que se pretendem ver respondidas com os resultados alcançados:

- 1) Que tipos de meios de prova digital, em particular audiovisuais, são mais frequentemente apresentados em processos de violência doméstica e como são valorados pelos/as magistrados/as?
- 2) Qual o valor probatório dado aos meios de prova audiovisual arrolados aos processos de violência doméstica?
- 3) Que estratégias ou critérios são adotados pelos/as magistrados/as perante a possibilidade de manipulação ou adulteração de meios de prova audiovisuais digitais?
- 4) Que desafios identificam os/as magistrados/as na avaliação da prova digital e que necessidades de adaptação consideram relevantes face à evolução tecnológica no sistema de justiça?

A realização de entrevistas qualitativa semiestruturadas constituiu o principal método de recolha de dados, sendo particularmente adequado para a recolha de dados

---

<sup>33</sup> LEVY, Neil. **Bad Beliefs: Why They Happen to Good People**. Oxford: Oxford University Press, 2022, p. 110 e ss.

<sup>34</sup> Pense-se no caso de uma testemunha que afirma ter visto um vídeo ou imagem que corrobora um determinado facto.

empíricos aprofundados sobre as experiências, práticas e percepções dos participantes<sup>35 - 36</sup>. O seu carácter flexível e interativo favorece a produção de conhecimento a partir do diálogo, possibilitando uma compreensão detalhada das perspetivas individuais dos/as participantes e dos contextos sociais em que estas se constroem<sup>37</sup>.

A análise centrou-se na forma como os/as magistrados/as interpretam e operacionalizam a prova no processo de tomada de decisão, atendendo às especificidades do crime de violência doméstica, frequentemente praticado em contexto privado e marcado pela, não raras vezes, ausência de testemunhas diretas. O estudo explorou ainda o impacto da crescente incorporação de meios digitais nos processos judiciais, nomeadamente no que respeita às práticas de valoração da prova audiovisual e às estratégias adotadas face a questões de autenticidade, fiabilidade e eventual manipulação, num contexto de rápidos avanços tecnológicos e de desenvolvimento da inteligência artificial.

Para a realização do presente estudo, selecionaram-se 8 magistrados/as (Cf. Tabela 1), dos/as quais quatro pertencem ao Ministério Público e quatro exercem funções como Juizes de Instrução Criminal (JIC), todos/as a exercer atividade profissional na zona do Porto e com experiência no domínio da violência doméstica. Foram definidos como critérios de inclusão: i) exercício de funções como magistrado/a judicial ou do Ministério Público; ii) atuação profissional no âmbito da violência doméstica; iii) disponibilidade para participar numa entrevista individual de natureza qualitativa.

A seleção dos/as participantes foi realizada de modo intencional e teve como objetivo assegurar a diversidade funcional dentro do sistema de justiça, permitindo captar perspetivas complementares sobre a valoração da prova.

**Tabela 1 – Caracterização dos/as participantes**

| Entrevistado/a | Género    | Categoria Profissional           |
|----------------|-----------|----------------------------------|
| E1             | Masculino | Juiz de Instrução Criminal       |
| E2             | Masculino | Juiz de Instrução Criminal       |
| E3             | Masculino | Juiz de Instrução Criminal       |
| E4             | Masculino | Juiz de Instrução Criminal       |
| E5             | Masculino | Procurador do Ministério Público |

<sup>35</sup> BRINKMANN, Steinar. The interview. In: DENZIN, Norman K.; LINCOLN, Yvonna S. (Org.). **The SAGE handbook of qualitative research**. 5. ed. Thousand Oaks: SAGE Publications, 2018.

<sup>36</sup> RENNISON, Callie Marie; HART, Thomas C. **Research methods in criminal justice and criminology**. Thousand Oaks: SAGE Publications, 2019.

<sup>37</sup> SUN, Dawei; BENSON, Michael L. Mixed methods: A justification, explication, and example. In: FARIA, Rita; DODGE, Mary. (Org.). **Qualitative research in criminology: Cutting-edge methods**. Cham: Springer, 2022. p. 37–49.

|    |           |                                  |
|----|-----------|----------------------------------|
| E6 | Feminino  | Procurador do Ministério Público |
| E7 | Feminino  | Procurador do Ministério Público |
| E8 | Masculino | Procurador do Ministério Público |

A constituição da amostra seguiu o princípio da saturação empírica, tendo sido interrompida quando se verificou a repetição de ideias e a ausência de novas informações relevantes<sup>38</sup>. Antes de cada entrevista, foi obtido o consentimento informado dos/as participantes, após explicação dos objetivos do estudo, dos procedimentos de recolha de dados e do tratamento da informação, em conformidade com as boas práticas éticas em investigação qualitativa<sup>39</sup>. Foi garantida a participação voluntária, bem como o anonimato e a confidencialidade dos dados. Os dados foram recolhidos entre novembro e dezembro de 2025, presencialmente, em local e horário previamente acordados com cada participante.

No decurso das entrevistas, foram igualmente apresentados aos/às participantes estímulos visuais de natureza fotográfica, incluindo imagens criadas através de ferramentas de inteligência artificial e fotografias reais, todas representando pessoas com sinais visíveis de agressão física e concebidas para representar situações de violência doméstica. Verificou-se que a larga maioria dos/as participantes não conseguiu identificar de forma segura as imagens geradas por IA, evidenciando dificuldades generalizadas na distinção entre conteúdos manipulados e não manipulados em contexto audiovisual.

As entrevistas foram gravadas e armazenadas em suporte digital protegido por palavra-passe, acessível apenas à equipa de investigação e autores/as deste artigo. Para efeitos de anonimização, os/as participantes foram identificados/as através de um código alfanumérico “E”, seguido de numeração sequencial.

Após a recolha dos dados através de entrevistas semiestruturadas, recorreu-se à análise temática de conteúdo<sup>40</sup> como estratégia analítica, tendo-se recorrido ao software MAXQDA (versão 24) para a organização, codificação e gestão sistemática dos dados qualitativos.

A Tabela 2 sintetiza os principais temas e subtemas emergentes da análise temática das entrevistas, evidenciando a articulação entre trajetórias profissionais, modelos

---

<sup>38</sup> SCHREIER, Margrit. Sampling and generalization. In: FLICK, Uwe. (Org.). **The SAGE handbook of qualitative data collection**. London: SAGE Publications, 2018, p. 84-98.

<sup>39</sup> WILES, Rose. **What are qualitative research ethics?** London: Bloomsbury Academic, 2013.

<sup>40</sup> BRAUN, Virginia; CLARKE, Victoria. Thematic analysis. In: COOPER, Harris *et al.* (Org.). **APA handbook of research methods in psychology**. v. 2. Washington, DC: American Psychological Association, 2012. p. 57–71; BRAUN, Virginia; CLARKE, Victoria. Using thematic analysis in psychology. **Qualitative Research in Psychology**, London, v. 3, n. 2, p. 77–101, 2006.

de valoração probatória e desafios associados à prova digital e audiovisual.

**Tabela 2 – Análise Temática das Entrevistas**

| <b>Temas</b>                                                           | <b>Subtemas</b>                                                       |
|------------------------------------------------------------------------|-----------------------------------------------------------------------|
| Trajetória profissional consolidada no domínio da Violência Doméstica  | Reconhecimento da violência doméstica como fenómeno criminal singular |
| Centralidade do depoimento da vítima na formação da convicção judicial | Avaliação integrada dos diferentes meios de prova                     |
|                                                                        | Relevância e valoração dos meios de prova audiovisuais                |
|                                                                        | Dependência narrativa e técnica da prova audiovisual                  |
| Presunção da autenticidade da prova audiovisual                        | ---                                                                   |
| Limites profissionais e institucionais na avaliação da prova digital   | ---                                                                   |

### 3. RESULTADOS

#### 3.1. Trajetória profissional consolidada no domínio da Violência Doméstica

A análise das 8 entrevistas revela que os/as magistrados/as apresentam um perfil profissional caracterizado por uma experiência prolongada na magistratura e por um contacto especializado e continuado com processos de violência doméstica. A maioria dos/as entrevistados/as refere uma carreira superior a dez anos, o que reflete percursos profissionais consolidados: “Sou magistrado há 15 anos” (E6), “Vai fazer 16 anos em janeiro” (E7); ou ainda superior a 20 anos “sou magistrado há quase 30 anos” (E3). Paralelamente, todos/as os/as magistrados/as referem possuir experiência especializada em violência doméstica. Alguns/as entrevistados/as sublinham mesmo a exclusividade ou a longevidade desse contacto, como ilustra a afirmação: “Trabalho em exclusivo na violência doméstica” (E6), ou ainda “Praticamente desde o início” (E8).

##### 3.1.1. Reconhecimento da violência doméstica como fenómeno criminal singular

Os discursos revelam também um entendimento consensual da violência doméstica como um crime com características próprias, distinto de outros crimes violentos. Os/as magistrados/as destacam o seu carácter intrafamiliar e a circunstância de ocorrer

maioritariamente em contexto privado, como expressa um/a entrevistado/a ao afirmar: “Normalmente os crimes de violência doméstica ocorrem dentro de casa” (E7). Outro/a magistrado/a reforça esta ideia ao referir que “(...) essencialmente, é praticado num ambiente recatado, do lar, num ambiente familiar, com quatro paredes” (E3), sublinhando a necessidade de uma abordagem diferenciada na apreciação da prova e na condução do processo. Neste enquadramento, todos/as os/as participantes reconhecem a acentuada complexidade probatória associada aos crimes de violência doméstica. A ausência frequente de testemunhas diretas surge como um dos principais obstáculos à prova, sendo reiterada em vários discursos: “Normalmente não há testemunhas” (E7).

### **3.2 Centralidade do depoimento da vítima na formação da convicção judicial**

No que respeita aos meios de prova mais frequentemente mobilizados nos processos de violência doméstica, bem como quanto à forma como estes são valorados pelos/as magistrados/as, parece haver uma consensualidade em todos/as os/as entrevistados/as. Os discursos revelam uma abordagem probatória marcada pela centralidade do depoimento da vítima, pela valorização de prova testemunhal indireta e periciais, e por uma crescente presença de prova documental e digital, sendo consensual a ideia de que nenhum meio de prova é, por si só, suficiente para sustentar uma decisão judicial.

O depoimento da vítima surge, de forma unânime, como o elemento probatório central nestes processos. Todos/as os/as magistrados/as entrevistados/as referem que as declarações da ofendida constituem a base da investigação e da formação da convicção judicial, sendo frequentemente descritas como o ponto de partida de todo o processo probatório. Esta centralidade é expressa de forma clara em afirmações como “a prova rainha é o depoimento da ofendida” (E5) ou “e até se calhar acima do que acontece noutro tipo de crime, pelas características do crime.” (E8). Ainda que reconhecendo a sua importância, os/as magistrados/as sublinham implicitamente a necessidade de proceder a uma análise crítica e contextualizada deste meio de prova, dada a natureza relacional e privada do crime.

#### **3.2.1 Avaliação integrada dos diferentes meios de prova**

É unânime entre os/as magistrados/as o entendimento de que nenhum meio de prova isolado é suficiente para fundamentar uma decisão. A prova é descrita como necessariamente cumulativa e interdependente, exigindo uma análise global e articulada dos diferentes elementos disponíveis. Esta ideia é expressa de forma clara nos excertos “Depois temos todo o resto da prova que vai ser produzida” (E3) e “Tem de ser concatenada com outro tipo de prova.” (E8), evidenciando uma lógica de valoração probatória baseada na coerência interna do conjunto da prova e na sua articulação contextual.

Assim, associada ao depoimento da vítima, a prova testemunhal indireta assume

igualmente um papel relevante em parte significativa dos processos, nomeadamente o recurso a testemunhos de terceiros que relatam confidências ou desabafos previamente feitos pela vítima, frequentemente junto de familiares ou pessoas próximas. Estes testemunhos são descritos como “desabafos à mãe, à irmã, à amiga” (E5), sendo reconhecidos como meios complementares que contribuem para a credibilização do relato da vítima e para a contextualização temporal dos factos, apesar das limitações inerentes à sua natureza indireta.

A prova pericial é também identificada, de forma consensual, como um meio recorrente e particularmente relevante na apreciação dos factos, sempre que existam alegações de agressões físicas, sendo este meio considerado essencial para a avaliação objetiva das lesões. A importância atribuída a este tipo de prova é evidenciada em expressões como “encaminhamos sempre para o Gabinete Médico-Legal” (E6) e “Há um recurso muito forte a esse tipo de prova pericial.” (E4), sublinhando o seu contributo para a corroboração do relato da vítima e para a delimitação da gravidade dos factos.

Paralelamente, os discursos revelam uma crescente valorização da prova documental e digital, nomeadamente mensagens escritas, registos telefónicos, fotografias e vídeos. Todos/as os/as entrevistados/as referem a presença frequente deste tipo de prova nos processos de violência doméstica, associando-a à evolução dos meios de comunicação e à generalização do uso de tecnologias digitais nas relações interpessoais: “hoje em dia é muito documental, mas de mensagens” (E6), bem como pela referência a “fotografias acima de tudo, porque normalmente as vítimas acabam por fazer o registo das situações mais graves, agora com os telemóveis não é difícil.” (E8) ou ainda “É comum, é comum. É comum. Temos muito fotografias de lesões, sobretudo, e temos muito vídeos e áudios de episódios de violência” (E5), evidenciando a centralidade crescente destes elementos na reconstrução dos factos e na análise das dinâmicas relacionais entre vítima e arguido.

### **3.2.2 Relevância e valoração dos meios de prova audiovisuais**

A prova digital e audiovisual assume um papel cada vez mais central nos processos de violência doméstica, sendo reconhecida pelos/as magistrados/as como um complemento relevante e, em alguns casos, determinante, na reconstrução dos factos e na formação da convicção judicial, evidência bem comprovada pelo excerto “Esta coisa das redes sociais facilita-nos imensa vida” (E3). Os discursos permitem identificar diferentes tipologias de prova digital, cuja frequência e valoração variam em função da sua natureza, contexto de obtenção e articulação com outros meios probatórios.

As mensagens escritas surgem de forma unânime como a principal tipologia de prova digital apresentada nos processos. Os/As magistrados/as referem a utilização recorrente de comunicações trocadas entre vítima e arguido, através de aplicações de

mensagens instantâneas, redes sociais, SMS e até plataformas associadas a transações financeiras (e.g., Mbwat). Esta centralidade é expressa em excertos como “hoje em dia é muito documental, mas de mensagens” (E6) e “utilizam WhatsApp, redes sociais... até o MB Way” (E7). As mensagens são descritas como particularmente relevantes para evidenciar padrões de controlo, ameaças, humilhações ou persistência de comportamentos abusivos. As fotografias de lesões tiradas pela própria vítima constituem igualmente um meio de prova digital amplamente referido. Estas imagens são frequentemente apresentadas pela ofendida como forma de documentar agressões físicas, sobretudo quando não existe intervenção imediata de terceiros ou quando o exame médico-legal ocorre dias depois dos factos. A recorrência deste meio é ilustrada por afirmações como “temos muitas fotografias de lesões” (E5) e “é muito comum as vítimas juntarem fotografias” (E6), sendo estas imagens entendidas como elementos de suporte ao relato da vítima, ainda que carecendo de contextualização e articulação com outros meios de prova, nomeadamente relatórios médicos “Sem relatórios médicos, pouco valor lhes atribuiria só de olhar” (E4). Por outro lado, os vídeos e gravações áudio, de episódios de violência surgem com menor frequência: “São mais raros, mas aparecem... Aparecem.”(3).

### 3.2.3 Dependência narrativa e técnica da prova audiovisual

Apesar da sua potencial força probatória, os/as magistrados/as encaram estes meios de prova com particular cautela, reconhecendo a necessidade de avaliar cuidadosamente o contexto em que a gravação foi obtida, a sua integridade e o respeito pelos limites legais. De forma consistente, a análise das entrevistas evidencia uma posição amplamente consensual relativamente à valoração das fotografias de lesões em processos de violência doméstica, as quais são reconhecidas como um meio de prova relevante, mas desprovido de valor probatório autónomo. Os/As entrevistados/as afirmam que este tipo de prova, quando analisado de forma autónoma, é insuficiente para sustentar uma decisão judicial, posição expressa de forma clara em excertos como “(...) Não podemos é fazer uma coisa sem a outra. Nem só com as fotografias, nem só com o relato.” (E1); “A fotografia por si só não documenta a autoria de um facto” (E2), ou ainda “Eu percebo que tipo de lesões são, mas se ela não colaborar ou se não houver outro tipo de prova testemunhal, documental, eu não consigo ter garantia de que estas lesões foram provocadas pelo agressor e quando” (E6). Esta rejeição assenta na impossibilidade de, a partir da imagem isolada, determinar com segurança a origem, o momento, o contexto ou a autoria das lesões retratadas. Neste sentido, os discursos sublinham de forma consistente a dependência das fotografias em relação ao depoimento da vítima, entendido como elemento essencial de contextualização narrativa. Todos/as os/as magistrados/as referem a necessidade de a ofendida explicar as circunstâncias em que as imagens foram captadas, o momento temporal a que dizem respeito e a sua ligação aos factos imputados ao arguido: “por isso é que digo que o contributo da ofendida é

fundamental, sem o contributo da ofendida nós ficamos sem saber o que é que concretamente provocou aquelas lesões” (E5) evidenciando que a valoração das fotografias está intrinsecamente ligada à credibilidade e coerência do relato da vítima.

Os/As magistrados/as destacam ainda a importância do reforço das fotografias através de prova pericial ou de registos clínicos, nomeadamente exames médico-legais. Quando acompanhadas por este tipo de prova, as fotografias ganham consistência e força probatória, funcionando como elementos complementares que corroboram objetivamente o relato da vítima. Esta articulação é expressa em excertos como “quando temos agressões físicas, temos elementos probatórios que acompanham e que sustentam estas declarações, nomeadamente os relatórios de registos clínicos, perícias médico-legais” (E3). Apesar da rejeição do seu valor autónomo, alguns/as magistrados/as reconhecem o impacto visual significativo das fotografias numa fase inicial do processo, nomeadamente no momento da avaliação indiciária e da aplicação de medidas de coação: “um processo de violência doméstica com fotografias é meio caminho dado para uma prisão preventiva.” (E5), sugerindo que, embora juridicamente insuficientes por si só, as fotografias podem contribuir para a perceção da gravidade dos factos numa fase preliminar.

### 3.3 Presunção da autenticidade da prova audiovisual

A análise das entrevistas revela que a verificação da autenticidade da prova audiovisual não surge, de forma geral, como uma preocupação inicial dos/as magistrados/as no momento da apreciação da prova. Os discursos evidenciam uma abordagem predominantemente presuntiva e reativa, em que a autenticidade é assumida à partida, sendo apenas questionada ou aprofundada em contextos específicos.

A maioria dos/as magistrados/as refere a ausência de suspeita inicial de manipulação dos conteúdos audiovisuais apresentados, assumindo a sua autenticidade como ponto de partida da análise probatória: “por regra nós não suspeitamos da autenticidade da prova.” (E5) ou “Eu não tenho por hábito pedir autenticidade” (E7). Tal presunção parece assentar na confiança no contexto processual, na forma como a prova é introduzida e na inexistência, na maioria dos casos, de indícios concretos de adulteração. Ainda assim, em alguns casos, os/as magistrados/as referem a prática de solicitar o ficheiro digital original, sobretudo quando pretendem assegurar uma análise mais rigorosa do conteúdo apresentado: “Pedimos o ficheiro digital, sim” (E8). Além disso, foi ainda referida a análise dos metadados como critério suficiente para aferir a autenticidade do ficheiro, salientando a relevância de informações como a data, a hora e a origem do ficheiro, conforme expresso em afirmações como “nós temos que ver aí os metadados do próprio ficheiro, da própria foto.” (E2) ou “daí a essencialidade de verificarmos os metadados das fotografias...Muitas vezes pode ser muito impressionante, mas não corresponder àquilo que realmente aconteceu.” (E8). No entanto,

esta estratégia não surge como uma prática generalizada, mas antes como um recurso pontual, dependente das características do caso concreto.

Um aspeto particularmente consensual prende-se com o carácter reativo da ativação de perícias técnicas, normalmente pelo/a arguido/a ou pela defesa: “Claro que o arguido pode sempre, a todo o tempo, quando for confrontado, dizer ‘olhe, isto não é possível, eu quero ver o ficheiro’” (E5) ou “Portanto, se o arguido coloca em causa (...) então nós temos que ver aí os metadados do próprio ficheiro...” (E2), evidenciando que a iniciativa de aprofundar tecnicamente a autenticidade da prova depende, em grande medida, do contraditório processual e não de uma avaliação oficiosa sistemática.

Por fim, é unânime entre os/as magistrados/as a ausência de experiências de perícia motivadas por suspeitas de manipulação com recurso a inteligência artificial. Todos/as os/as entrevistados/as referem nunca ter ordenado ou considerado uma perícia com esse fundamento, como ilustram os excertos “Não, não. Não pedi.” (E6). Este dado sugere que, apesar da crescente discussão pública sobre o impacto da inteligência artificial na fiabilidade da prova digital, estas preocupações ainda não se materializaram de forma concreta na prática judicial quotidiana.

### **3.4 Limites profissionais e institucionais na avaliação da prova digital**

De forma unânime, os/as magistrados/as assumem a sua incapacidade técnica para detetar autonomamente manipulações digitais, reconhecendo explicitamente que não dispõem dos conhecimentos necessários para identificar alterações sofisticadas em conteúdos audiovisuais. Esta limitação é expressa de forma direta em afirmações como “não tenho conhecimentos para isso” (E2) e “não estamos preparados” (E7), revelando uma perceção clara da distância entre a função jurisdicional e as competências técnicas exigidas por este tipo de análise. Além disso afirmam evitar qualquer juízo baseado em impressões visuais ou perceções subjetivas, considerando essa prática inadequada e potencialmente perigosa: “isso é um jogo perigoso para os juízes” (E4), evidenciando uma preocupação com a objetividade. Perante estes limites, os/as magistrados/as manifestam uma dependência clara da perícia especializada, deslocando a responsabilidade da verificação técnica da autenticidade para peritos com formação adequada: “temos peritos para isso” (E4); “não é função do juiz” (E1), reforçando uma divisão funcional entre a apreciação jurídica da prova e a sua validação técnica.

Em paralelo, é igualmente unânime o reconhecimento da necessidade de formação específica em matéria de prova digital e inteligência artificial, entendida como uma resposta prioritária aos desafios emergentes. Todos/as os/as magistrados/as referem a formação como essencial para uma compreensão mínima dos riscos, limites e potencialidades da prova digital: “é absolutamente essencial” (E7), considerando que a evolução digital

dificultará o trabalho dos/as julgadores/as “Para quem decida vai ser muito mais complicado. Muito mais complicado.” (E3). Esta formação é implicitamente concebida não como substituta da perícia, mas como um instrumento de literacia técnica que permita uma melhor avaliação crítica dos pareceres periciais e das alegações das partes.

Por fim, parte dos/as entrevistados/as destaca a falta de meios humanos e técnicos como um constrangimento estrutural adicional. Esta fragilidade é expressa em afirmações como “falta de meios humanos” (E5) e “nem os OPCs estão preparados” (E7), sugerindo que os desafios associados à prova digital ultrapassam o plano individual e se estendem à organização e capacidade do sistema de justiça como um todo.

## CONSIDERAÇÕES FINAIS

O presente estudo procurou analisar as percepções, práticas e necessidades dos/as magistrados/as relativamente à admissibilidade e valoração da prova digital, com especial enfoque nos meios de prova audiovisual em processos de violência doméstica. Os resultados evidenciam que a prática judicial continua a assentar numa lógica de valoração probatória cumulativa e articulada – não sendo suficiente a utilização de um único tipo de forma isolada, havendo necessidade de uma apreciação global e contextual –, em que o depoimento da vítima assume um papel central, sendo complementado por outros meios de prova como a prova testemunhal, pericial e, de forma crescente, a prova digital.

A prova audiovisual, designadamente em formato de mensagens escritas e fotografias de lesões, surge como um instrumento amplamente utilizado e relevante na reconstrução dos factos, muito devido à evolução tecnológica a que se assistiu nas últimas décadas. Ainda assim, esta prova está desprovida de valor probatório autónomo dependendo, a sua valoração, da contextualização através do relato da vítima e da sua corroboração por outros meios de prova, em particular a perícia médico-legal. Apesar do impacto visual que este tipo de prova pode assumir em fases iniciais do processo, os/as magistrados/as revelam uma postura prudente quanto à sua suficiência para fundamentar as suas decisões. Esta prudência revela-se particularmente significativa à luz dos contributos teóricos que identificam os *deepfakes* como um fator de erosão do princípio do *seeing is believing*, sublinhando que a força persuasiva das imagens pode coexistir com elevados níveis de falsificação ou manipulação impercetível.

No que respeita à autenticidade da prova audiovisual, os resultados apontam para uma abordagem predominantemente presuntiva e reativa. A autenticidade é, em regra, assumida como ponto de partida, sendo a ativação de mecanismos de verificação técnica, nomeadamente perícias, maioritariamente desencadeada por reação de outros sujeitos

processuais e não por iniciativa oficiosa sistemática. Importa salientar que nenhum/a dos/as magistrados/as entrevistados/as reportou experiências concretas de suspeita ou detecção de manipulação de prova audiovisual com recurso a inteligência artificial, o que sugere um desfasamento entre a crescente discussão teórica e mediática sobre os *deepfakes* e a sua materialização prática no quotidiano judicial. Este desfasamento pode ser interpretado não como ausência de risco, mas como manifestação de uma assimetria temporal entre o avanço acelerado das tecnologias de inteligência artificial generativa e a sua incorporação efetiva nas práticas judiciais, reforçando a natureza latente do perigo epistémico associado aos *deepfakes*.

Paralelamente, é unânime o reconhecimento, por parte dos/as participantes, das suas limitações técnicas na detecção autónoma de manipulações digitais sofisticadas, bem como a rejeição de qualquer juízo baseado em impressões subjetivas ou perceções visuais. Esta consciência dos limites da função jurisdicional reforça a dependência da perícia especializada e evidencia a importância de uma clara divisão funcional entre a avaliação jurídica da prova e a sua validação técnica. Ainda assim, os/as magistrados/as sublinham a necessidade de formação específica em matéria de prova digital e inteligência artificial, entendida como condição essencial para uma apreciação crítica informada dos pareceres periciais e das alegações dos sujeitos processuais.

Em termos globais, os resultados apontam para a existência de desafios que extravasam o plano individual e se estendem ao nível estrutural do sistema de justiça, nomeadamente no que respeita à escassez de meios humanos e técnicos adequados à crescente complexidade da prova digital. Neste sentido, o estudo sugere que a resposta aos riscos colocados pela inteligência artificial generativa não se esgota na adaptação normativa, exigindo igualmente investimento em formação, recursos especializados e estratégias institucionais que reforcem a resiliência epistémica do processo penal. Embora exploratório e limitado pelo número de participantes e pelo seu enquadramento territorial, o presente estudo contribui para uma compreensão empiricamente informada dos desafios emergentes da prova digital em contextos de elevada vulnerabilidade probatória, abrindo caminho a futuras investigações e à reflexão crítica sobre a adequação do sistema de justiça penal face à transformação tecnológica em curso.

## REFERÊNCIAS DAS FONTES CITADAS

BELCIC, Ivan; STRYKER, Cole. **What is GPT** (generative pretrained transformer)?. IBM, [S.l.], [S.d.]. Disponível em: <https://www.ibm.com/think/topics/gpt>. Acesso em: 03 dez. 2025.

BRAUN, Virginia; CLARKE, Victoria. Thematic analysis. *In*: COOPER, Harris *et al.* (Org.). **APA handbook of research methods in psychology**. v. 2. Washington, DC: American Psychological Association, 2012. p. 57–71.

BRAUN, Virginia; CLARKE, Victoria. Using thematic analysis in psychology. **Qualitative Research in Psychology**, London, v. 3, n. 2, p. 77–101, 2006.

BRINKMANN, Steinar. The interview. *In*: DENZIN, Norman K.; LINCOLN, Yvonna S. (Org.). **The SAGE handbook of qualitative research**. 5. ed. Thousand Oaks: SAGE Publications, 2018. p. 997-1038.

CHESNEY, Robert; CITRON, Danielle K. Deep Fakes: A looming challenge for privacy, democracy, and national security. **California Law Review**, v. 107, n. 6, p. 1771-1786, 2019.

CHINA, Chrystal R. Five machine learning types to know. *In*: Cole STRYKER *et al.* (eds.). **The 2025 Guide to Machine Learning**. [S.l.]: IBM, 2025. Disponível em: <https://www.ibm.com/think/topics/machine-learning-types>. Acesso em: 01 nov. 2025.

COLLINS DICTIONARY. **Collins Online Dictionary**: definitions, thesaurus and translations. Disponível em: <https://www.collinsdictionary.com>. Acesso em: 03 dez. 2025.

COMISSÃO EUROPEIA. **Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões**: Roteiro para o acesso lícito e efetivo aos dados para efeitos da aplicação da lei. Bruxelas, 2025.

DIXON JR., Herbert B. DeepFakes: More frightening than photoshop on steroids. **Judges' Journal**, v. 58, n. 3, p. 35-37, 2019.

DURÃES, Dalila; FREITAS, Pedro Miguel; NOVAIS, Paulo. The Relevance of Deepfakes in the Administration of Criminal Justice. *In*: ANTUNES, Henrique Sousa *et al.* (eds.). **Multidisciplinary Perspectives on Artificial Intelligence and the Law**. Cham: Springer, 2024.

EDMONDS, Lauren. 3 years old, 800 million users, 29,000 prompts a second: ChatGPT's meteoric rise, by the numbers". **Business Insider**, [S.l.], 2025. Disponível em: <https://www.businessinsider.com/chatgpt-by-the-numbers-2025-11#chat-gpt-has-800-million-weekly-active-users-2>. Acesso em: 03 dez. 2025.

EUROPOL. **Facing reality?** Law enforcement and the challenge of deepfakes. Luxemburgo: Publications Office of the European Union, 2022.

GAGLIARDI, Marília Papaléo. Artificial Intelligence and Women's Rights. *In*: QUINTAVALLA, Alberto e TEMPERMAN, Jeroen (eds.). **Artificial Intelligence and Human Rights**. Oxford: OUP Oxford, 2023.

GOOGLE. What is machine learning (ML)?. **Google**, [S.l.], [S.d.]. Disponível em: <https://cloud.google.com/learn/what-is-machine-learning?hl=en>. Acesso em: 03 dez. 2025.

GOOGLE. What is Machine Learning? *In: GOOGLE. **Introduction to Machine Learning***. [S.l.], 2025. Disponível em: <https://developers.google.com/machine-learning/intro-to-ml/what-is-ml>. Acesso em: 27 nov. 2025.

LEE, Fangfang. What is a neural network? *In: Cole STRYKER et al. (eds.). **The 2025 Guide to Machine Learning***. [S.l.]: IBM, 2025. Disponível em: <https://www.ibm.com/think/topics/neural-networks>. Acesso em: 30 nov. 2025.

LEVY, Neil. **Bad Beliefs: Why They Happen to Good People**. Oxford: Oxford University Press, 2022.

MARAS, Marie-Helen; ALEXANDROU, Alex. Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake vídeos. **The International Journal of Evidence & Proof**, v. 23, n. 3, p. 255-262, 2018

OPENAI. Fine-tuning GPT-2 from human preferences. **OpenAI**, [S.l.], 2019. Disponível em: <https://openai.com/index/fine-tuning-gpt-2/>. Acesso em: 03 dez. 2025.

OPENAI. Introducing ChatGPT. **OpenAI**, [S.l.], 2022. Disponível em: <https://openai.com/index/chatgpt/>. Acesso em: 03 dez. 2025.

OPENAI. OpenAI API. **OpenAI**, [S.l.], 2020. Disponível em: <https://openai.com/index/openai-api/>. Acesso em: 03 dez. 2025.

RENNISON, Callie Marie; HART, Thomas C. **Research methods in criminal justice and criminology**. Thousand Oaks: SAGE Publications, 2019.

SCHREIER, Margrit. Sampling and generalization. *In: FLICK, Uwe. (Org.). **The SAGE handbook of qualitative data collection***. London: SAGE Publications, 2018, p. 84-98.

SOR, Jennifer. Sam Altman touts ChatGPT's 800 million weekly users, double all its main competitors combined. **Business Insider**, [S.l.], 2025. Disponível em: <https://www.businessinsider.com/chatgpt-users-openai-sam-altman-devday-llm-artificial-intelligence-2025-10>. Acesso em: 03 dez. 2025.

STRYKER, Cole *et al.* What is deep learning? *In: Cole STRYKER et al. (eds.). **The 2025 Guide to Machine Learning***. [S.l.]: IBM, 2025. Disponível em: <https://www.ibm.com/think/topics/deep-learning>. Acesso em: 30 nov. 2025.

STRYKER, Cole; SCAPICCHIO, Mark. What is generative AI? *In: Cole STRYKER et al. (eds.). **The 2025 Guide to Machine Learning***. [S.l.]: IBM, 2025. Disponível em: <https://www.ibm.com/think/topics/generative-ai>. Acesso em: 27 nov. 2025.

SUN, Dawei; BENSON, Michael L. Mixed methods: A justification, explication, and example. *In: FARIA, Rita; DODGE, Mary. (Org.). **Qualitative research in criminology: Cutting-edge methods***. Cham: Springer, 2022. p. 37-49.

TREND MICRO RESEARCH. **Malicious Uses and Abuses of Artificial Intelligence**. Europol's European Cybercrime Centre, 2020.

TWOMEY, John *et al.* What Is So Deep About Deepfakes? A Multi-Disciplinary Thematic Analysis of Academic Narratives About Deepfake Technology. **IEEE Transactions on Technology and Society**, v. 6, n. 1, p. 64-79, 2025.

UNIVERSIDADE DE HARVARD. The Benefits and Limitations of Generative AI: Harvard Experts Answer Your Questions. **Harvard**, 2023. Disponível em: <https://harvardonline.harvard.edu/blog/benefits-limitations-generative-ai>. Acesso em: 27 nov. 2025.

UNIVERSIDADE DE LEEDS. What is Generative AI? **University of Leeds**, Leeds, [S.d.]. Disponível em: <https://generative-ai.leeds.ac.uk/intro-gen-ai/strengths-and-weaknesses/>. Acesso em: 27 nov. 2025.

WAGNER, Travis L.; BLEWER, Ashley. The Word Real Is No Longer Real: Deepfakes, Gender, and the Challenges of AI-Altered Video. **Open Information Science**, v. 3, n. 1, 2019.

WHITTAKER, Lucas *et al.* Mapping the deepfake landscape for innovation: A multidisciplinary systematic review and future research agenda. **Technovation**, v. 125, p. 01-17. 2023.

WILES, Rose. **What are qualitative research ethics?** London: Bloomsbury Academic, 2013.

---

## INFORMAÇÕES DOS AUTORES

---

### Pedro Miguel Freitas

Investigador do Católica Research Centre for the Future of Law (CEID) - Portugal. Doutor pela Escola de Direito da Universidade do Minho - Portugal. Mestre pela Escola de Direito da Universidade do Minho - Portugal. Endereço eletrônico: [pfreitas@ucp.pt](mailto:pfreitas@ucp.pt). ORCID: 0000-0003-4516-0588.

### Ana Guerreiro

PhD in Criminology from the Faculty of Law, University of Porto. Assistant Professor at the University of Maia and Invited Assistant Professor at the Faculty of Law, School of Criminology, University of Porto (Portugal). Researcher at both the Interdisciplinary Centre for Gender Studies, Institute of Social and Political Sciences, University of Lisbon (CIEG-ISCSP, UL), and the CIJ – Centre for Interdisciplinary Research on Justice, University of Porto (Portugal). Endereço eletrônico: [aguerreiro@umaia.pt](mailto:aguerreiro@umaia.pt). ORCID: 0000-0003-2312-6266.

---

## COMO CITAR

---

FREITAS, Pedro Miguel; GUERREIRO, Ana. Deepfakes e prova digital na violência doméstica: os riscos da Inteligência Artificial. **Novos Estudos Jurídicos**, Itajaí (SC), v. 30, n. 3, p. 398-419, 2025. DOI: 10.14210/nej.v30n3.p398-419.

---

## INFORMAÇÕES ADICIONAIS

---

Este trabalho contou com o apoio financeiro da Fundación Manuel Serra Domínguez (Espanha).

Recebido em: 05 de dez. de 2025.

Aprovado em: 18 dez. de 2025.