

¿Es posible sancionar penalmente en España la difusión de las deepfakes pornográficas?

A propósito de la Sentencia del Juzgado de Menores de Badajoz 86/2024, de 20 de junio

Dr. CRISTIAN SÁNCHEZ BENÍTEZ

UNIVERSIDAD DE JAÉN - ESPAÑA



ORCID: 0000-0001-6725-1320



Resumen. En jurisdicción ciber, se analiza cuáles son las posibilidades reales del castigo del Deepfake en España en la actualidad.

Palabras clave. Deepfake, Condena menores, LO protección integral de los menores

Abstract. In cyber jurisdiction, this paper examines the real possibilities for punishing deepfakes in Spain today.

Keywords. Deepfake, Juvenile Sentencing, Comprehensive Protection of Minors Act.

Cómo citar: Sánchez Benítez, Cristian (2025). "¿Es posible sancionar penalmente en España la difusión de las deepfakes pornográficas? A propósito de la Sentencia del Juzgado de Menores de Badajoz 86/2024, de 20 de junio". En CiberData, núm. 1, 2025, pp. 35-38.

En una fecha indeterminada entre los meses de agosto y septiembre de 2023, varios adolescentes –todos menores de edad– de la localidad de Almendralejo (Extremadura) generaron mediante aplicaciones de inteligencia artificial imágenes manipuladas de veinte vecinas de la localidad, también menores de edad, en concreto, de edades comprendidas entre los once y quince años. Para ello, los menores emplearon al menos dos aplicaciones que habían obtenido desde enlaces subidos a la aplicación de mensajería Telegram: clothoff.io y CLF 2 bot. Con estas herramientas, y a partir de fotografías reales de las menores obtenidas en sus perfiles en redes sociales, crearon nuevas imágenes de sus rostros a las que se les superponían imágenes de otros cuerpos femeninos desnudos, “en poses que enfatizaban los órganos sexuales, aparentando de modo realista corresponder a las propias menores” (según se indica en el apartado de Hechos Probados de la resolución objeto de análisis, la cual me facilitó Omar Abdedaim-Herrera). De hecho, según informaciones aparecidas en prensa, cuando las imágenes se difundieron tras compartirlas los menores en dos grupos de WhatsApp, una de las afectadas llegó a jurar a sus padres que no había hecho nada.

A este tipo de imágenes -y vídeos e incluso textos y audios de voz- generadas con herramientas de inteligencia artificial y que aparentan representar a personas, objetos, lugares o sucesos reales, se las conoce como deepfakes. Se trata de un término creado a partir de la unión de los vocablos deep (profundo) y fake (falsificación), y que en castellano puede traducirse como ultrafalsificaciones o ultrasuplantaciones. El prelegislador español de 2025, en la Exposición de Motivos del Proyecto de Ley Orgánica de protección integral de los menores de edad en los entornos digitales, define las ultrafalsificaciones como "imágenes o voces manipuladas tecnológicamente y extremadamente realistas". La Unión Europea por su parte define el término ultrasuplantación como "un contenido de imagen, audio o vídeo generado o manipulado por una IA que se asemeja a personas, objetos, lugares, entidades o sucesos reales y que puede inducir a una persona a pensar erróneamente que son auténticos o verídicos" (artículo 3 del Reglamento 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024).

No ha de pasarse por alto que la totalidad de los autores de las ultrafalsificaciones de Almendralejo eran varones, que todas las víctimas eran mujeres y que las imágenes generadas eran de naturaleza sexual. Ello concuerda con lo revelado por diversos estudios sobre la temática, por cuanto encuentran que la inmensa mayoría de los autores de estos contenidos, casi siempre de naturaleza sexual (asociando cuerpos ficticios, desnudos y/o en poses o contextos de naturaleza sexual a rostros reales), así como de su posterior difusión, son hombres; mientras que más del 90% de las víctimas son mujeres. Estos datos evidencian que este fenómeno es una manifestación clara de violencia digital y sexual de género, sin que ello suponga negar la posibilidad de que entre los autores de estos comportamientos en

algunos casos pueda haber mujeres, de que entre las víctimas pueda haber a veces hombres y de que la naturaleza de los deepfakes no siempre sea sexual (en ocasiones se generan y difunden contenidos de carácter vejatorio); ni se niega la necesidad político-crímenal de sancionar estas conductas con independencia del sexo de víctimas y victimarios y del carácter sexual o no de los contenidos.

Retomando el caso, el 20 de junio de 2024, en la Sentencia -de conformidad- 86/2024, el Juzgado de Menores de Badajoz declaró a los menores de Almendralejo (quince) responsables de veinte delitos de trato degradante del artículo 173.1 del Código penal, pues la difusión ocasionó "humillación y desasosiego en las afectadas"; y de veinte delitos de difusión de pornografía infantil del artículo 189 del mismo texto, ya que la totalidad de las víctimas, como se indicó, eran menores de edad. Recuérdese que este último precepto incluye en el concepto de pornografía infantil la conocida entre los especialistas como la pornografía infantil virtual: "imágenes realistas [falsas, en tanto que creadas artificialmente, pero que aparentan ser reales] de un menor participando en una conducta sexualmente explícita o imágenes realistas [ídem] de los órganos sexuales de un menor, con fines principalmente sexuales".

La magistrada impuso a los menores la medida de libertad vigilada durante un año, "con contenido especialmente orientado a recibir formación afectivo sexual, sobre uso responsable de las tecnologías de la información y la comunicación, sensibilización en materia de igualdad y género, así como intervención en los aspectos deficitarios que presenten los menores, de acuerdo con los respectivos informes del Equipo Técnico". Asimismo, se decretó en el fallo "la devolución, a cada uno de los menores, de sus terminales móviles previamente formateados o restaurados a

valores de fábrica, quedando así eliminados cualesquiera contenidos o imágenes relativas a las menores perjudicadas, que obtuvieron accediendo de manera irregular, a las aplicaciones referidas en los hechos probados". También fueron condenados al pago de las costas, pero sorprende sin embargo que no se contenga en la resolución referencia alguna al pago de eventuales responsabilidades civiles derivadas del delito. Desconozco si es que los representantes legales de las perjudicadas declinaron la vía penal para solicitar la reparación civil o si es que directamente renunciaron a dicha reparación.

Por otra parte, en septiembre de 2025 se conoció que la Agencia Española de Protección de Datos había sancionado administrativamente (EXP202503445) a uno de esos menores con una multa de 2.000 euros (que se redujo a 1.200 euros por reconocimiento de responsabilidad y pago voluntario). Y ello porque la difusión por su parte de las imágenes manipuladas con inteligencia artificial, asociando "rostros reales a cuerpos desnudos que no se corresponden", supuso un tratamiento ilícito de datos personales, "por cuanto la imagen es un dato personal" y porque dicho tratamiento "se efectuó sin que concurriera ninguna de las causas de legitimación enumeradas en el artículo 6.1 del Reglamento General de Protección de Datos". Quizá con esta segunda sanción a este menor ya condenado con anterioridad por los mismos hechos se infringe el principio non bis in ídem, que prohíbe la doble sanción (incluso cuando una sea de naturaleza penal y otra administrativa) si existe identidad de sujetos, hechos y fundamento sancionador. Un mes después, la Agencia apercibió (EXP202506010) a un niño de trece años que había difundido una deepfake sexual de una menor (aparecía su rostro real asociado a un cuerpo desnudo ajeno), generada a partir de una fotografía real que aquella había publicado en su perfil de Instagram.

De este modo, a día de hoy es posible sancionar penal y administrativamente en España la difusión de deepfakes pornográficas o de contenido vejatorio sin necesidad de emprender modificaciones normativas. Además de las mencionadas sanciones administrativas impuestas por la Agencia Española de Protección de Datos y de los precitados delitos de trato degradante y de pornografía infantil (para cuando las imágenes generadas con inteligencia artificial representen a menores de edad) aplicados por el Juzgado de Menores extremeño, existen otros tipos penales que la doctrina ha llegado a proponer para el castigo de estos comportamientos. Así, se ha planteado la posibilidad de emplear los delitos contra el honor (injurias) y contra la intimidad. Respecto de estos últimos, no resulta adecuada su invocación por la sencilla razón de que no hay imágenes ni intimidad real que proteger en los casos de deepfakes. No hay atentado a la intimidad en la difusión de una ultrafalsificación porque los contenidos difundidos son falsos. En mi opinión, en la medida en que la difusión de deepfakes, por su naturaleza cosificadora, atenta principalmente contra la integridad moral, bien jurídico más amplio y que podría integrar también los ataques contra el honor, es el delito de trato degradante el tipo penal más idóneo para reprimir tales conductas.

Sin embargo, desde marzo de 2025 se tramita en el Congreso de los Diputados el precitado Proyecto de Ley Orgánica para la protección integral de los menores de edad en los entornos digitales, que incluye entre sus novedades la incriminación expresa de la difusión, exhibición o cesión de la imagen corporal o la voz de una persona generados con inteligencia artificial u otras tecnologías, "de modo que parezca real, simulando situaciones de contenido sexual o gravemente vejatorios". La creación o la mera posesión de estos contenidos serían penalmente atípicas, salvo que pudieran encuadrarse en el amplísimo delito de pornografía infantil,

- siempre y cuando las víctimas de estas ultrafalsificaciones generadas pero no compartidas fuesen menores de edad.

Para tal fin se pretende introducir un nuevo artículo (173 bis), ubicado entre los delitos contra la integridad moral –de hecho, se exige en el delito proyectado que concurra en el autor el ánimo de menoscabar la integridad moral de la víctima–, y que castigaría con la pena de prisión de entre uno y dos años la difusión de estas deepfakes sexuales o gravemente vejatorias; y con la pena en su mitad superior –de un año y medio a dos años– si se hiciera a través de “un medio de comunicación social, por medio de internet o mediante el uso de tecnologías” –lo que suele ocurrir en la mayor parte de los casos–, “de modo que aquel se hiciera accesible a un elevado número de personas en el espacio virtual”.

Como se ha podido ver en los párrafos anteriores, el ordenamiento jurídico español ya cuenta con varios instrumentos sancionadores (penales y administrativos) perfectamente aplicables a este tipo de supuestos. Así, la concreta tipificación penal de la difusión de las ultrasuplantaciones sexuales o gravemente vejatorias no resultaría necesaria, en la medida en que se cuenta ya con varios tipos penales que pueden utilizarse y que de hecho se han utilizado para su represión. No es que piense que no deban castigarse penalmente comportamientos como los acaecidos en Almendralejo. Son lo suficientemente graves como para invocar al Derecho penal, cuya intervención en tales casos resultará indispensable tanto para prevenir la comisión de hechos análogos como para sancionar contundentemente los ya perpetrados. De lo que se duda aquí es de la conveniencia de crear un nuevo delito expreso de difusión de deepfakes, por resultar innecesario, puesto que puede aplicarse, entre otros, el delito de trato degradante. Un nuevo delito que, adviértase,

se situaría, como aquél, entre los delitos contra la integridad moral y que tendría idéntico límite superior de la pena (dos años). En cualquier caso, habrá que esperar un tiempo –quizá años– para poder constatar si la reforma proyectada, de aprobarse, aunque prescindible por los motivos expuestos, resulta positiva al menos en un sentido preventivo general: en términos de reforzamiento simbólico de un mensaje de intolerancia social-penal frente a hechos como el comentado en la presente contribución.

Créditos:

Este trabajo se ha desarrollado en el marco del Proyecto PROPIMAGENXXI “Protección penal de la imagen frente al chantaje de los nuevos desafíos del siglo XXI: venganzas sexuales, fake news, trampantojos y otros enredos”, financiado por la Agencia Estatal de la Investigación del Ministerio de Ciencia e Innovación del Gobierno de España dentro del Programa de Generación de Conocimiento 2024 (Ref.: PID2024-15934OOB-IOO).