

Algunas claves del castigo penal del deepfake de naturaleza sexual

[ibericonnect.blog /2023/07/algunas-claves-del-castigo-penal-del-deepfake-de-naturaleza-sexual/](https://ibericonnect.blog/2023/07/algunas-claves-del-castigo-penal-del-deepfake-de-naturaleza-sexual/)

Abraham Devís Matamoros

24 de julio de 2023



En esta sociedad que irónicamente podemos denominar Sociedad Digital de la Desinformación, la manipulación ha adquirido un merecido protagonismo en una era en la que el avance de las nuevas tecnologías ha permitido que esta sea más fácil, más accesible y más destacada. De hecho, ya no solo se alude a la información falsa (*fake news*), sino que también se habla de ultra falsos, también conocidos como *deepfakes*.

El contenido audiovisual, sobre el que hasta ahora existía un alto grado de credibilidad, queda en entredicho a propósito de la IA. Una herramienta tecnológica que se utiliza en muchas ocasiones con propósitos satíricos, pero también para, a partir de imágenes reales de la víctima, crear materiales audiovisuales falsos, sin consentimiento previo, en los que se hacen creer que un individuo realiza acciones de carácter sexual que nunca han ocurrido.

Un comportamiento que presenta la particularidad de no ser neutro desde el punto de vista del género puesto que afecta de manera más contundente a las mujeres, existiendo un potente binomio entre el *deepfake* sexual y la violencia de género. No es casualidad que prácticamente la totalidad de los materiales generados en este sentido sean las mujeres las que lo sufren como ha sido los casos conocidos de Taylor Swift, Scarlett Johansson o de las streamers estadounidenses Atrio y Sweet Anita, entre otros muchos.

En este sentido, ¿es necesaria una regulación penal expresa que aborde este problema? Bajo mi punto de vista no. No porque sea una conducta que no respete las garantías de ofensividad y de intervención mínima, sino porque ya es posible reconducir los tipos penales existentes. Aquí ocurre algo similar a lo que ya remarcó la Fiscalía General de Estado en 2020 sobre las *fake news*, y es que la propagación de *deepfakes* puede subsumirse en diferentes tipos penales como medio comisivo en función del bien jurídico que lesionen o pongan en peligro. En realidad, la pregunta precisa debería ser: ¿qué tipo se adapta mejor al desvalor que origina el *deepfake* pornográfico?

Desde luego, la propuesta presentada por el grupo parlamentario Plural en su enmienda a la LO 4/2023 de contemplarlos bajo los delitos contra la intimidad debe ser rechazada. certeza, nuestra posición como juristas frente a la tecnología debe ser la de exigencias penal vigente a las nuevas en la conformación de bienes jurídicos, una idea que implica reformular la interpretación de la intimidad, pero nunca hasta el punto de plantear nuevas configuraciones del derecho que se desvirtuarían por completo su contenido. Al fin y al cabo, el objeto material de la conducta se integra por acciones de carácter sexual que no realiza el propio sujeto pasivo, sino que son creaciones virtuales situadas completamente al margen de la esfera de la privacidad. La dimensión positiva o autodeterminación informativa (STS nº112/2023, 20 de febrero) y la facultad de exclusión de intromisión de aspectos de la vida (STC, Pleno, 29 de junio de 2022) del derecho consagrado en el art. 18.1 CE no se alteran mediante estas actuaciones.

No obstante, sí que pueden verse afectados derechos como el honor y la integridad moral. Dos valores jurídicos cuya conceptualización se encuentra íntimamente relacionada con la noción de dignidad. El primero, en la medida en que, por el mero hecho de ser personas, todos tenemos el mismo honor, independientemente de lo que cada uno logre (vertiente “igualitaria”). El segundo, tal y como sigue grabando el Tribunal Supremo (STS nº181/2023, 15 de marzo), la integridad moral debe entenderse como una manifestación directa de la dignidad humana y de la inviolabilidad de la persona. Lo que implica la proscripción de cualquier uso instrumental de un sujeto y el derecho a ser tratado como persona y no como un mero objeto.

Desde esta perspectiva, tanto el delito de injurias del artículo 208 CP como el delito de trato degradante del artículo 173.1 CP parecerían ser los más apropiados para abordar este fenómeno. Dos preceptos en los que destaca el requisito de la “gravedad” de la conducta. Elemento este último que responde al carácter fragmentario del Derecho penal sin el cual, los límites del *ius puniendi* en lo que respeta a estos tipos penales quedarían excesivamente amplios. No en vano los casos excepcionales sobre los que se castigan las injurias y vejaciones cuando son leves en el artículo 173.4 CP son críticos por la expansión punitiva que suponen.

Sin embargo, la vaguedad e imprecisión de este concepto jurídico indeterminado debilita considerablemente el mandato de taxatividad que exige el principio de legalidad. Y ello porque, al tratarse de un elemento eminentemente circunstancial, será el juez o tribunal quien, atendiendo a las circunstancias concurrentes en el hecho determinará en el caso concreto dicha gravedad. Pese a ello, en alguna ocasión (STS, nº 325/2013, de 2 de

abril) nuestro Tribunal Supremo ha acudido a la jurisprudencia del TEDH para extraer criterios que sopesen la gravedad acudiendo a argumentos como son la duración, la intensidad, los efectos físicos y mentales y las condiciones de la víctima (sexo, edad y estado de salud). Criterios que, aplicados al problema que nos ocupa, justifican la gravedad exigible en ambos tipos.

En primer lugar, porque la propia dinámica comisiva del *deepfake* pornográfico implica que el objeto material se encuentre en el ciberespacio, lo que provoca una permanencia con potencialidad *in aeternum*. En segundo lugar, se produce un aumento de la lesividad en relación con otras modalidades de ataque. Sobre todo, por la enorme dificultad de distinguir entre el contenido falso y el real debido a la precisión tecnológica de las GAN. Aspecto al que hay que añadir el mayor grado de veracidad que mantenemos respecto de materiales audiovisuales sobre materiales escritos. Razonamientos tecnológicos a los que hay que sumar la afectación a la heteroestima y la autoestima de la víctima y el propio contexto de subdiscriminación de las mujeres en el que las consideraciones sociales sobre la sexualidad femenina distan y mucho respecto de las masculinas.

Ahora bien, entre las dos alternativas plausibles expuestas, en mi opinión, la opción de contemplarlos bajo los delitos contra la integridad moral sería la opción más acertada. Desde una perspectiva formal porque operaría el principio de consunción en virtud del cual dicho precepto sería más amplio y abarcaría también los ataques contra el honor. No solo englobaría la afectación a la autoestima y a la heteroestima a la que se ha aludido, sino también a la cosificación e instrumentalización que se produce sobre el sujeto pasivo, generalmente la mujer, sobre los que se añade un cuerpo diferente como objeto de consumo. Una condición en la que el carácter degradante de la acción le es impuesta, violentando su voluntad. Razón que se refuerza en la propia regulación penal en la medida en que, cuando las vejaciones y las injurias son leves, el legislador ha decidido unificar ambos comportamientos en los delitos contra la integridad moral. Además, por cuestiones de aplicabilidad práctica ya que el *animus iniuriandi* exigible en las injurias no siempre concurre. De hecho, en muchas ocasiones se tratan de acciones realizadas con un ánimo distinto como la obtención de rendimientos económicos ya que se suben a webs pornográficas simplemente para conseguir más reproducciones y suscripciones en dichas páginas.

En cualquier caso, sería conveniente establecer una alternativa unificada al respecto sobre la subsunción en uno de ambos supuestos para evitar problemas relacionados con la seguridad jurídica.