
ARTÍCULO PRODUCTO DE INVESTIGACIÓN

**El consentimiento informado en la era de la inteligencia artificial:
retos éticos y derecho a la información del paciente**

Informed consent in the age of artificial intelligence:
Ethical challenges and the patient's right to information

Laura López-Almansa Beaus¹
Universidad Europea de Valencia

Recibido: 10.10.2025
Aceptado: 10.12.2025

Resumen

Introducción: La incorporación de sistemas de inteligencia artificial en el ámbito sanitario está transformando los procesos de toma de decisiones clínicas, generando nuevos desafíos para los derechos fundamentales de los pacientes. En particular, el consentimiento informado, como garantía de la autonomía del paciente, se ve tensionado por la complejidad técnica, la opacidad algorítmica y la automatización de decisiones, lo que plantea interrogantes sobre la efectividad del derecho a la información en este nuevo contexto. **Métodos:** El trabajo se desarrolla mediante una investigación jurídico-doctrinal de carácter cualitativo, basada en el análisis crítico del marco normativo vigente, especialmente del Reglamento General de Protección de Datos y del Reglamento de Inteligencia Artificial de la Unión Europea, así como en la revisión de doctrina científica e informes institucionales relevantes en materia de inteligencia artificial y salud. El análisis se centra en la relación entre consentimiento

¹laura.lopezalmansa@universidadeuropea.es
<https://orcid.org/0009-0000-7604-0576>

informado, derecho a la información y sistemas de decisión automatizados. **Resultados:** El estudio pone de relieve las limitaciones del modelo clásico de consentimiento informado en el contexto de la incorporación de sistemas de inteligencia artificial. En este sentido, se identifican fenómenos como la opacidad algorítmica, el aumento de la asimetría informativa y las dificultades para asegurar una comprensión efectiva por parte del paciente. Asimismo, se evidencia un desajuste entre las exigencias normativas actuales y la realidad tecnológica, que conduce a la configuración de un consentimiento meramente formal o aparente. **Conclusiones:** Los resultados ponen de manifiesto la necesidad de una reinterpretación del derecho a la información del paciente que incorpore criterios de comprensión efectiva, explicabilidad operativa y adaptación a contextos de alta complejidad tecnológica. El consentimiento informado debe evolucionar hacia un modelo reforzado que garantice la autonomía real del paciente en entornos clínicos digitalizados.

Palabras clave: consentimiento informado, inteligencia artificial, derecho a la información, autonomía del paciente, bioética

Abstract

Background: The integration of artificial intelligence systems in the healthcare sector is transforming clinical decision-making processes, generating new challenges for patients' fundamental rights. Informed consent, as a safeguard of patient autonomy, is being strained by technical complexity, algorithmic opacity, and automated decision-making, raising concerns about the effectiveness of the right to information in this new context. **Methods:** This study is conducted through a qualitative legal-doctrinal approach, based on a critical analysis of the current regulatory framework, particularly the General Data Protection Regulation and the European Union Artificial Intelligence Act, as well as a review of relevant academic literature and institutional reports on artificial intelligence and healthcare. The analysis focuses on the relationship between informed consent, the right to information, and automated decision-making systems. **Results:** The study highlights the limitations of the traditional model of informed consent in the context of the integration of artificial intelligence

systems. It identifies phenomena such as algorithmic opacity, increased informational asymmetry, and difficulties in ensuring patients' effective understanding. Furthermore, a mismatch is observed between current regulatory requirements and technological reality, which leads to the emergence of a merely formal or apparent form of consent. **Conclusions:** The findings demonstrate the need to reinterpret the patient's right to information by incorporating criteria of effective understanding, operational explainability, and adaptation to contexts of high technological complexity. Informed consent must evolve toward a reinforced model capable of ensuring genuine patient autonomy in digitally mediated clinical environments.

Keywords: informed consent, artificial intelligence, right to information, patient autonomy, bioethics

Introducción

La incorporación progresiva de sistemas de inteligencia artificial en el ámbito sanitario está generando una transformación estructural en los procesos de diagnóstico, tratamiento y toma de decisiones clínicas. Estas tecnologías, basadas en el análisis automatizado de grandes volúmenes de datos, permiten mejorar la precisión diagnóstica y optimizar tratamientos, pero al mismo tiempo introducen nuevos riesgos para los derechos fundamentales de los pacientes, particularmente en lo relativo al derecho a la información.

El consentimiento informado se ha configurado tradicionalmente como uno de los pilares fundamentales de la autonomía del paciente, sustentado en la obligación de proporcionar información suficiente, comprensible y adecuada que permita una toma de decisiones libre y consciente. En este sentido, la doctrina clásica ha señalado que el consentimiento no constituye un mero requisito formal, sino un proceso orientado a garantizar la autonomía del paciente a través de la comprensión efectiva de la información (Beauchamp & Childress, 2019; Faden & Beauchamp, 1986; Simón Lorda, 2000).

Asimismo, el Convenio de Oviedo reconoce el consentimiento libre e informado como condición esencial de cualquier intervención en el ámbito de la salud. En el ámbito jurídico español, esta exigencia se recoge expresamente en la Ley 41/2002 de autonomía del paciente, que establece el derecho del paciente a recibir información adecuada y comprensible como presupuesto del consentimiento informado.

Sin embargo, la creciente utilización de sistemas de inteligencia artificial en la práctica clínica plantea interrogantes relevantes sobre la capacidad real de los pacientes para comprender los procesos que subyacen a dichas decisiones, especialmente cuando estos dependen de modelos algorítmicos complejos.

Diversos organismos internacionales han advertido que los sistemas de inteligencia artificial pueden operar mediante modelos de alta complejidad cuya lógica resulta difícilmente accesible, incluso para los propios profesionales sanitarios. Esta característica, comúnmente descrita como opacidad algorítmica, incrementa la asimetría informativa en la relación médico-paciente y dificulta la transmisión de información en términos comprensibles, afectando directamente al contenido material del derecho a la información (World Health Organization, 2021; Council of Europe, s. f.; Floridi et al., 2018).

Desde el punto de vista jurídico, esta situación entra en tensión con los marcos normativos vigentes. En el ámbito europeo, el Reglamento General de Protección de Datos reconoce el derecho del interesado a recibir información significativa sobre la lógica aplicada en decisiones automatizadas, así como garantías frente a este tipo de decisiones (European Parliament and Council of the European Union, 2016). Asimismo, el Reglamento de Inteligencia Artificial de la Unión Europea introduce obligaciones de transparencia y control para sistemas considerados de alto riesgo, entre los que se incluyen determinados usos sanitarios (European Parliament and Council of the European Union, 2024).

No obstante, la aplicación efectiva de estas garantías presenta importantes limitaciones, especialmente en lo relativo a la traducción de la complejidad técnica de los sistemas en

información comprensible para el paciente. En consecuencia, surge la cuestión de si el consentimiento informado, en su configuración actual, continúa siendo un instrumento válido para garantizar el derecho a la información en entornos mediados por inteligencia artificial.

En consecuencia, resulta necesario analizar el impacto de la inteligencia artificial en el derecho a la información del paciente, identificando los principales riesgos derivados de la opacidad algorítmica, la complejidad técnica y la automatización de decisiones, así como examinar las limitaciones del marco normativo vigente. A partir de este análisis, el trabajo identifica una serie de riesgos estructurales que afectan a la efectividad del consentimiento informado en entornos mediados por inteligencia artificial, así como limitaciones concretas del marco normativo actual. Estos hallazgos permiten plantear la necesidad de una reinterpretación del derecho a la información del paciente, adaptada a los sistemas automatizados de decisión.

1. Metodología

El presente trabajo se enmarca en una investigación de carácter jurídico-doctrinal, basada en el análisis cualitativo de fuentes normativas, doctrinales e institucionales relacionadas con el consentimiento informado, el derecho a la información del paciente y la regulación de la inteligencia artificial en el ámbito sanitario.

Desde una perspectiva metodológica, se ha adoptado un enfoque analítico-crítico orientado a examinar las implicaciones jurídicas derivadas de la incorporación de sistemas de inteligencia artificial en la práctica clínica. Para ello, se ha llevado a cabo una revisión sistemática del marco normativo vigente, con especial atención al Reglamento General de Protección de Datos y al Reglamento de Inteligencia Artificial de la Unión Europea (European Parliament and Council of the European Union, 2016, 2024), así como a informes elaborados por organismos internacionales en materia de ética y gobernanza de la inteligencia artificial en salud (World Health Organization, 2021).

El desarrollo del análisis se ha estructurado en tres fases. En primer lugar, se ha realizado una delimitación conceptual del consentimiento informado como institución jurídica, identificando sus elementos esenciales y su vinculación con el derecho a la información del paciente. En segundo lugar, se ha examinado el impacto de los sistemas de inteligencia artificial en las condiciones materiales que permiten el ejercicio efectivo de dicho derecho, prestando especial atención a factores como la opacidad algorítmica, la complejidad técnica y la automatización de decisiones. En tercer lugar, se ha efectuado un análisis crítico del marco normativo actual, con el objetivo de identificar tensiones, lagunas regulatorias y limitaciones en su aplicación práctica.

Este enfoque metodológico permite no solo describir el estado actual de la cuestión, sino también identificar patrones de riesgo jurídicamente relevantes y formular una interpretación crítica orientada a la adaptación del consentimiento informado a los entornos mediados por inteligencia artificial.

2. Resultados

2.1. Inadecuación del modelo clásico de consentimiento informado

El análisis del consentimiento informado como institución jurídica permite afirmar que su validez se fundamenta en tres elementos esenciales: información suficiente, comprensión real y voluntariedad de la decisión. Este esquema, consolidado tanto en la práctica clínica como en el derecho sanitario, presupone que la información proporcionada puede ser entendida por el paciente en términos que le permitan valorar las consecuencias de la intervención. Este planteamiento coincide con la concepción doctrinal del consentimiento informado como un proceso orientado a la comprensión sustantiva y no meramente formal (Faden & Beauchamp, 1986).

Sin embargo, la incorporación de sistemas de inteligencia artificial altera este presupuesto básico. En particular, cuando la decisión clínica depende de modelos algorítmicos complejos,

la información relevante deja de ser plenamente accesible en términos comprensibles para el paciente. En este contexto, aunque formalmente se proporcione información, no puede garantizarse el requisito de comprensión, lo que resulta contrario a las exigencias de información adecuada y comprensible previstas en la Ley 41/2002.

Desde esta perspectiva, la aplicación del modelo clásico conduce a una conclusión relevante: el consentimiento informado puede considerarse formalmente válido, pero materialmente defectuoso, al no cumplirse el estándar de comprensión exigido por la propia institución.

Esta situación evidencia una incompatibilidad estructural entre el modelo jurídico vigente y los sistemas de toma de decisiones basados en inteligencia artificial. No obstante, esta problemática no es exclusiva de los entornos tecnológicos, sino que ha sido identificada previamente en el ámbito del consentimiento informado tradicional.

Diversos estudios empíricos han puesto de manifiesto que una parte significativa de los pacientes no comprende adecuadamente la información recibida, especialmente cuando esta se presenta en formatos escritos o con elevado contenido técnico (López-Almansa Beaus, 2015; Barrio Cantalejo et al., 2008). Asimismo, se ha observado que variables como el nivel educativo, la forma de transmisión de la información o su grado de legibilidad influyen directamente en la comprensión del paciente, mejorando esta cuando la información se adapta de forma individualizada y se transmite mediante procesos comunicativos efectivos.

Estos resultados refuerzan la idea de que la validez del consentimiento informado depende de factores materiales vinculados a la comprensión real del paciente y no únicamente del cumplimiento formal del deber de información, lo que evidencia un desequilibrio informativo relevante en la relación clínica entre médico y paciente (Katz, 2002).

2.2. Alcance limitado del derecho a la información en decisiones automatizadas

El Reglamento General de Protección de Datos reconoce el derecho del interesado a recibir “información significativa sobre la lógica aplicada” en decisiones automatizadas (arts. 13, 15

y 22). Este precepto constituye, en principio, un intento de adaptar el derecho a la información a los entornos tecnológicos, incorporando garantías específicas frente al uso de sistemas automatizados en la toma de decisiones (European Parliament and Council of the European Union, 2016).

No obstante, el análisis de este mandato revela una limitación esencial. El Reglamento exige la comunicación de la lógica del sistema, pero no establece criterios claros sobre el nivel de inteligibilidad que debe alcanzar dicha información. En el contexto de sistemas de aprendizaje automático, donde la lógica no se expresa en reglas comprensibles sino en patrones estadísticos complejos, esta exigencia resulta difícilmente operativa. Como ha señalado la doctrina, el denominado “derecho a explicación” presenta importantes dificultades cuando se aplica a modelos algorítmicos complejos, ya que la información proporcionada puede resultar técnicamente correcta pero materialmente incomprensible para el destinatario (Goodman & Flaxman, 2017; Mittelstadt et al., 2016).

Asimismo, esta problemática se ve acentuada por la naturaleza probabilística de muchos sistemas de inteligencia artificial, en los que las decisiones no responden a una lógica determinista fácilmente explicable, sino a correlaciones estadísticas que dificultan su traducción en términos accesibles para el paciente. En consecuencia, incluso cuando se proporciona información sobre el funcionamiento del sistema, esta puede no ser suficiente para garantizar un entendimiento real.

De este modo, se produce una disociación entre el cumplimiento meramente formal de las obligaciones informativas y la capacidad real de comprensión por parte del paciente. Este desajuste permite concluir que el derecho a la información, tal como está configurado en el marco normativo actual, no garantiza por sí mismo una capacidad de entendimiento en entornos mediados por inteligencia artificial, lo que limita su eficacia como instrumento de protección de la autonomía del paciente.

2.3. Configuración de un “consentimiento aparente”

A partir de los elementos anteriores, el análisis permite identificar la aparición de un fenómeno jurídicamente relevante que puede denominarse consentimiento aparente. Este concepto describe aquellas situaciones en las que se cumple formalmente con el deber de información, pero el paciente carece de los elementos necesarios para comprender el alcance de la decisión adoptada.

Desde una perspectiva doctrinal, este fenómeno no es completamente nuevo, sino que se inserta en la distinción entre consentimiento formal y consentimiento materialmente válido. Como ha señalado la literatura clásica, la mera transmisión de información no garantiza por sí misma la autonomía del paciente, en la medida en que esta requiere condiciones reales de comprensión y confianza en la relación clínica (O'Neill, 2002; Katz, 2002).

En el contexto de la inteligencia artificial, este fenómeno se ve intensificado por dos factores principales: en primer lugar, la opacidad técnica de los sistemas y, en segundo lugar, la limitada capacidad explicativa por parte de los profesionales sanitarios. Diversos estudios han señalado que los modelos de inteligencia artificial pueden operar como sistemas no interpretables, lo que dificulta trasladar al paciente una explicación comprensible del proceso decisional (Cohen & Slottje, 2023).

Asimismo, la complejidad técnica de estos sistemas introduce una nueva dimensión de asimetría informativa, en la que el conocimiento relevante sobre la decisión clínica se desplaza desde el profesional hacia el propio sistema tecnológico, reduciendo la capacidad del paciente para participar de manera informada en la toma de decisiones.

El resultado es que el consentimiento informado pierde su función como instrumento de garantía de la autonomía, convirtiéndose en un mecanismo carente de eficacia material que no refleja una comprensión real por parte del paciente. En este sentido, el consentimiento aparente constituye una manifestación específica de las limitaciones del modelo tradicional

en entornos mediados por inteligencia artificial, en los que el cumplimiento formal de la obligación de informar no se traduce necesariamente en una decisión verdaderamente informada.

2.4. Impacto de los sesgos algorítmicos en la calidad de la información

El análisis de la literatura sobre inteligencia artificial en salud pone de relieve que los sistemas algorítmicos pueden incorporar sesgos derivados de los datos de entrenamiento, así como de las decisiones de diseño y de las condiciones en las que se desarrollan y aplican. Estos sesgos pueden afectar directamente a la equidad y fiabilidad de las decisiones clínicas, generando resultados diferenciados en función de variables como el sexo, la edad o el origen étnico del paciente (Centers for Disease Control and Prevention, 2024; Floridi et al., 2018; Mittelstadt et al., 2016).

Desde la perspectiva del consentimiento informado, este elemento introduce una exigencia adicional: no basta con informar sobre el funcionamiento general del sistema, sino que resulta necesario incorporar información relativa a sus limitaciones, márgenes de error y posibles sesgos. Sin embargo, en la práctica, este tipo de información rara vez se traslada al paciente, lo que implica que este consiente decisiones cuyo grado de fiabilidad, precisión y equidad desconoce.

Esta situación refuerza la conclusión anterior: la información proporcionada no satisface los estándares sustantivos exigibles al consentimiento informado, en la medida en que omite elementos esenciales para una valoración real de los riesgos y beneficios asociados al uso de sistemas de inteligencia artificial en la práctica clínica.

2.5. Desajuste estructural del marco normativo

El Reglamento de Inteligencia Artificial de la Unión Europea introduce obligaciones de transparencia, supervisión y gestión de riesgos para sistemas considerados de alto riesgo,

incluidos aquellos utilizados en el ámbito sanitario. Este marco normativo supone un avance relevante en la regulación de los sistemas de inteligencia artificial, al establecer requisitos orientados a garantizar su seguridad, fiabilidad y control a lo largo de su ciclo de vida (European Parliament and Council of the European Union, 2024).

No obstante, el análisis de estas obligaciones revela que están orientadas principalmente a garantizar la seguridad del sistema y a facilitar su supervisión por parte de autoridades y profesionales, más que a asegurar la efectiva comprensión por parte del paciente. En este sentido, el modelo regulatorio europeo se articula fundamentalmente en torno a obligaciones dirigidas a los proveedores y usuarios de los sistemas, sin configurar de manera clara un derecho sustantivo del paciente a comprender el funcionamiento de las decisiones automatizadas.

Esta limitación se conecta con un problema previamente identificado en el marco del Reglamento General de Protección de Datos. Aunque este reconoce el derecho a recibir “información significativa sobre la lógica aplicada” en decisiones automatizadas, diversos autores han cuestionado la existencia y efectividad de un verdadero “derecho a explicación”, señalando que el contenido de esta obligación es limitado y ambiguo (Goodman & Flaxman, 2017; Wachter, Mittelstadt, & Floridi, 2017). En particular, se ha destacado que el Reglamento no exige una explicación individualizada plenamente comprensible, sino únicamente una información general sobre el funcionamiento del sistema, lo que reduce significativamente su utilidad práctica.

Asimismo, desde una perspectiva técnica y jurídica, se ha puesto de manifiesto que los sistemas de inteligencia artificial basados en aprendizaje automático presentan dificultades estructurales para cumplir con estas exigencias de transparencia. La denominada “problemática de la caja negra” implica que incluso cuando existe voluntad de explicar el funcionamiento del sistema, dicha explicación puede no ser inteligible ni operativa para el destinatario final (Mittelstadt et al., 2016).

En consecuencia, se produce una disociación entre los objetivos normativos declarados (transparencia, explicabilidad y control), y su efectividad real en términos de comprensión. El marco regulatorio vigente garantiza, en gran medida, la trazabilidad, la supervisión y la responsabilidad técnica de los sistemas, pero no asegura que la información proporcionada sea accesible, comprensible y útil para el paciente en el proceso de toma de decisiones clínicas.

Este desajuste permite identificar una limitación estructural del modelo normativo actual: mientras que el derecho se orienta a regular el funcionamiento de los sistemas y a mitigar riesgos desde una perspectiva técnica, no logra trasladar estas garantías al plano de la comprensión efectiva del paciente. En consecuencia, la protección del derecho a la información queda incompleta, en la medida en que el cumplimiento nominal de las obligaciones de transparencia no se traduce necesariamente en una capacidad real del paciente para entender y valorar las decisiones adoptadas en entornos mediados por inteligencia artificial.

2.6. Comprensión efectiva, legibilidad y consentimiento informado en entornos tecnológicos

Estas limitaciones normativas adquieren especial relevancia cuando se analizan desde la perspectiva de la comprensión efectiva del paciente.

El análisis del consentimiento informado desde una perspectiva clásica ha puesto de manifiesto que su validez no depende únicamente de la transmisión formal de información, sino de la capacidad real de entendimiento por parte del paciente. En este sentido, la doctrina ha señalado que no puede hablarse de consentimiento informado si la información proporcionada no es inteligible ni se encuentra adaptada a las características del destinatario (Barrio Cantalejo et al., 2008).

Asimismo, se ha evidenciado que la legibilidad y la comprensibilidad de la información constituyen elementos esenciales para garantizar la validez del consentimiento, especialmente

en contextos donde el lenguaje técnico o la complejidad de la información pueden dificultar su asimilación. La información debe ser adecuada, suficiente y adaptada al nivel cultural y cognitivo del paciente, evitando tecnicismos innecesarios y favoreciendo un proceso de comunicación real y efectivo.

Sin embargo, en entornos mediados por sistemas de inteligencia artificial, estas exigencias se ven particularmente comprometidas. La complejidad técnica de los sistemas, su carácter opaco y la dificultad de traducir su funcionamiento a un lenguaje accesible dificultan la comprensión por parte del paciente. En este contexto, la literatura ha señalado que el uso de sistemas algorítmicos en la toma de decisiones clínicas puede generar incertidumbre sobre el alcance del deber de información y dificultar la explicación del proceso decisional, incluso para los propios profesionales sanitarios (Cohen & Slottje, 2023).

A ello se suma que los sistemas de inteligencia artificial pueden incorporar sesgos derivados de los datos de entrenamiento o del diseño del sistema, lo que puede afectar a la calidad, fiabilidad y equidad de la información utilizada en la toma de decisiones clínicas (Centers for Disease Control and Prevention, 2024). Estas limitaciones no siempre son trasladadas al paciente, lo que implica que la información proporcionada puede resultar incompleta o insuficiente para una valoración real de los riesgos y beneficios asociados al uso de estos sistemas.

Desde el punto de vista normativo, tanto el Reglamento General de Protección de Datos como el Reglamento de Inteligencia Artificial de la Unión Europea introducen obligaciones de transparencia y acceso a la información en relación con sistemas automatizados (European Parliament and Council of the European Union, 2016, 2024). No obstante, estas obligaciones se orientan fundamentalmente a garantizar el acceso a la información, sin asegurar necesariamente su inteligibilidad efectiva ni su adecuación a las capacidades del paciente.

En consecuencia, se observa una continuidad entre los problemas clásicos de legibilidad y comprensión del consentimiento informado y los nuevos desafíos derivados del uso de inteligencia artificial, que no solo reproducen estas dificultades, sino que las intensifican.

Un ejemplo ilustrativo de esta problemática puede observarse en el uso de sistemas de inteligencia artificial para la generación de documentos de consentimiento informado en procedimientos clínicos. En un análisis comparativo previo de formularios de consentimiento para cirugía de implantes dentales generados mediante modelos de inteligencia artificial, se constató que, si bien estos documentos presentaban una mejora relativa en términos de legibilidad y claridad del lenguaje, dicha mejora formal no se traducía necesariamente en una mayor calidad jurídica o informativa del contenido.

En particular, el caso de estudio puso de manifiesto la existencia de omisiones relevantes en aspectos esenciales del consentimiento informado, como la falta de referencia expresa a la posibilidad de revocación del consentimiento, la ausencia de una descripción completa de los riesgos asociados al procedimiento o la insuficiente contextualización de las alternativas terapéuticas. Estas carencias afectan directamente a la capacidad del paciente para valorar de forma adecuada las implicaciones de la intervención. Asimismo, se detectaron errores en la interpretación y presentación de la información clínica, derivados de la utilización de modelos generativos que, en algunos casos, incorporaban información genérica o descontextualizada respecto al caso concreto. Este tipo de errores, aunque puedan pasar desapercibidos en una lectura superficial, resultan especialmente relevantes desde el punto de vista jurídico, en la medida en que pueden distorsionar la información proporcionada al paciente y afectar a la validez del consentimiento prestado.

El análisis evidenció, además, que los sistemas de inteligencia artificial tienden a priorizar la claridad expositiva sobre la exhaustividad del contenido, lo que puede conducir a simplificaciones excesivas o a la omisión de información jurídicamente relevante. Este fenómeno pone de manifiesto una tensión entre legibilidad y completitud que resulta

especialmente problemática en el ámbito del consentimiento informado, donde la precisión y la integridad de la información constituyen elementos esenciales.

En este contexto, el caso analizado permite concluir que la utilización de sistemas de inteligencia artificial en la elaboración de documentos de consentimiento informado, aunque puede contribuir a mejorar ciertos aspectos formales, no garantiza por sí misma la adecuación del contenido ni la comprensión efectiva por parte del paciente. Por el contrario, puede generar nuevas fuentes de riesgo vinculadas a la calidad, fiabilidad y contextualización de la información proporcionada (López-Almansa Beaus, 2024).

Esta situación evidencia que una mayor claridad formal del documento no garantiza necesariamente una mejor calidad del consentimiento informado, ya que la comprensión real del paciente depende no solo de la forma en que se presenta la información, sino también de su exactitud, integridad y adecuación al contexto clínico.

3. Discusión

Los resultados obtenidos evidencian que el impacto de la inteligencia artificial en el ámbito sanitario no constituye únicamente un desafío técnico o regulatorio, sino que plantea una reconfiguración profunda del consentimiento informado como institución jurídica. En particular, la dificultad para garantizar una comprensión efectiva de los procesos automatizados pone en cuestión uno de los presupuestos esenciales de la autonomía del paciente.

Desde esta perspectiva, los problemas identificados en relación con la inteligencia artificial no son completamente nuevos, sino que amplifican limitaciones ya existentes en el modelo tradicional de consentimiento informado. La doctrina ha puesto de manifiesto que la existencia de un documento firmado no garantiza por sí misma la validez del consentimiento, especialmente cuando el paciente no comprende el contenido de la información proporcionada.

En este sentido, la exigencia de una comprensión suficiente, así como la necesidad de adaptar la información a las características individuales del paciente, constituyen elementos esenciales del consentimiento informado que resultan difíciles de garantizar en entornos tecnológicos avanzados. La estandarización de la información y la automatización de los procesos decisionales pueden dificultar la personalización de la comunicación, incrementando el riesgo de que el consentimiento se reduzca a un cumplimiento meramente formal. Este doble efecto —mejora formal del lenguaje y persistencia de problemas de comprensión— pone de relieve la necesidad de no identificar automáticamente claridad textual con validez del consentimiento informado (López-Almansa Beaus, 2024).

Desde una perspectiva doctrinal, el consentimiento informado no puede entenderse como un mero cumplimiento formal del deber de información, sino como un proceso orientado a asegurar la comprensión sustantiva del paciente. Este fenómeno, ya identificado en contextos tradicionales del consentimiento informado, se intensifica significativamente en entornos mediados por inteligencia artificial.

En esta línea, la doctrina reciente ha puesto de manifiesto que el consentimiento informado no puede considerarse válido cuando la información proporcionada al paciente no es suficientemente clara, completa y específica. En particular, el análisis de la sentencia del Tribunal Europeo de Derechos Humanos de 26 de junio de 2025 evidencia que la falta de información adecuada sobre un riesgo previsible impide al paciente adoptar una decisión verdaderamente informada, incluso cuando existe un documento de consentimiento firmado. Este pronunciamiento refuerza la idea de que el consentimiento informado no puede reducirse a un requisito formal, sino que exige una comprensión real por parte del paciente (Hernández Herrero, 2026).

Este planteamiento se ve reforzado al atender directamente al contenido de la resolución del Tribunal, que distingue claramente entre la adecuación del marco normativo y su aplicación práctica. En el caso analizado, el Tribunal no aprecia deficiencias en la normativa española

sobre consentimiento informado, pero sí en su implementación efectiva, al considerar que la información proporcionada al paciente no fue suficientemente clara, específica ni comprensible. En particular, se destaca que la utilización de formularios genéricos o ambiguos, así como la ausencia de documentación sobre la información verbal facilitada, impiden acreditar que el paciente haya comprendido realmente el alcance de la intervención. Asimismo, el Tribunal insiste en que la obligación de información exige comunicar de forma expresa los riesgos previsibles concretos de la intervención, no siendo suficiente una referencia genérica a posibles modificaciones de la técnica quirúrgica (Hernández Herrero, 2026).

En este escenario, la inteligencia artificial introduce una dimensión adicional de complejidad, al generar nuevas formas de opacidad en los procesos de toma de decisiones clínicas. Tal como advierte la literatura especializada, la dificultad para explicar el funcionamiento de determinados sistemas algorítmicos limita la capacidad del paciente para comprender los fundamentos de las decisiones que afectan a su salud, lo que refuerza la asimetría informativa en la relación clínica (Cohen & Slottje, 2023).

Asimismo, autores como Luciano Floridi han señalado que la inteligencia artificial introduce una nueva forma de asimetría informativa, en la que la capacidad de comprender los procesos decisionales se desplaza desde los individuos hacia los sistemas tecnológicos. En la misma línea, Brent Mittelstadt advierte que la opacidad de los algoritmos limita la posibilidad de rendición de cuentas y dificulta el control por parte de los afectados, lo que tiene implicaciones directas en términos de derechos fundamentales.

Bajo estas circunstancias, se observa una tensión entre el marco normativo nacional y el europeo. Mientras la Ley 41/2002 de autonomía del paciente exige que la información proporcionada al paciente sea adecuada y comprensible como presupuesto del consentimiento informado, las normas europeas, como el Reglamento General de Protección de Datos y el Reglamento de Inteligencia Artificial, se centran en la transparencia y el acceso a la información, sin garantizar necesariamente su inteligibilidad efectiva. Esta diferencia pone de

manifiesto una brecha entre el cumplimiento formal de las obligaciones informativas y la comprensión real por parte del paciente en entornos mediados por inteligencia artificial.

Todo ello obliga a replantear el alcance del derecho a la información. El Reglamento General de Protección de Datos reconoce el derecho a recibir información sobre la lógica de las decisiones automatizadas, pero no garantiza que dicha información sea comprensible en la práctica. La doctrina ha señalado que el denominado “derecho a explicación” presenta importantes limitaciones cuando se aplica a sistemas complejos de aprendizaje automático, ya que la lógica de estos sistemas no siempre puede traducirse en términos accesibles para el usuario (Goodman & Flaxman, 2017).

Por su parte, el Reglamento de Inteligencia Artificial de la Unión Europea introduce obligaciones de transparencia y supervisión para sistemas de alto riesgo, pero su enfoque se centra principalmente en la seguridad y la gobernanza del sistema, no en la adecuada comprensión por parte del paciente. Como ha señalado la Organización Mundial de la Salud, la implementación de inteligencia artificial en salud debe garantizar no solo la eficacia técnica, sino también la protección de los derechos de los pacientes, incluyendo su capacidad de tomar decisiones informadas.

En este contexto, resulta especialmente relevante la reciente propuesta de Espacio Europeo de Datos Sanitarios (European Health Data Space), que refuerza el uso secundario de datos de salud y amplía los escenarios de tratamiento automatizado de información clínica. Este desarrollo normativo introduce nuevas complejidades en el ejercicio del derecho a la información, al incrementar la distancia entre el origen de los datos, su tratamiento y su utilización final, lo que puede dificultar aún más la capacidad de entendimiento por parte del paciente.

Esta problemática no solo afecta a la autonomía individual del paciente, sino que tiene implicaciones directas en el desarrollo humano, al incidir en la equidad en el acceso a una información comprensible en contextos tecnológicos avanzados. En particular, la creciente

complejidad de los sistemas de inteligencia artificial puede generar nuevas formas de desigualdad entre aquellos pacientes con mayor capacidad de comprensión tecnológica y aquellos en situación de vulnerabilidad, profundizando brechas en el ejercicio efectivo de derechos fundamentales en el ámbito sanitario.

A partir de estos elementos, puede sostenerse que el problema no reside únicamente en la falta de información, sino en la inadecuación del modelo jurídico actual para entornos de alta complejidad tecnológica. El consentimiento informado, tal como está configurado, se basa en una concepción lineal de la información (transmisión → comprensión → decisión), que resulta difícilmente aplicable cuando la información no es plenamente inteligible.

En este sentido, se plantea la necesidad de una evolución del modelo hacia un consentimiento informado reforzado, articulado en torno a tres elementos esenciales:

- . Explicabilidad operativa: no basta con proporcionar información técnica, sino que esta debe traducirse en términos comprensibles para el paciente, adaptados a su nivel de conocimiento.
- . Información sobre incertidumbre y sesgo: el paciente debe ser informado no solo del funcionamiento del sistema, sino también de sus limitaciones, incluyendo posibles sesgos y márgenes de error.
- . Refuerzo del papel del profesional sanitario: el profesional debe asumir una función de mediación interpretativa, garantizando que la información derivada del sistema sea comprensible y contextualizada.

Este enfoque se alinea con propuestas recientes en el ámbito de la ética de la inteligencia artificial, que abogan por un modelo centrado en el usuario y en la protección efectiva de sus derechos, más allá del cumplimiento formal de obligaciones de transparencia.

En consecuencia, la discusión permite concluir que la adaptación del consentimiento informado a los entornos de inteligencia artificial no puede limitarse a ajustes normativos puntuales, sino que requiere una reinterpretación profunda del derecho a la información,

orientada a garantizar la comprensión real del paciente en contextos de creciente complejidad tecnológica.

Conclusiones

El análisis realizado pone de manifiesto que la incorporación de sistemas de inteligencia artificial en el ámbito sanitario no constituye únicamente un avance tecnológico, sino que implica una transformación estructural de las condiciones que hacen posible el ejercicio efectivo del consentimiento informado.

En particular, se ha evidenciado que el modelo clásico de consentimiento, basado en la transmisión de información comprensible al paciente, resulta insuficiente en entornos mediados por sistemas algorítmicos complejos. La opacidad técnica, la automatización de decisiones y la dificultad para traducir la lógica de los sistemas en términos accesibles generan una ruptura entre el cumplimiento de las obligaciones informativas y la comprensión real por parte del paciente.

Este desajuste permite afirmar que el consentimiento informado, en su configuración actual, puede devenir en muchos casos un consentimiento meramente aparente, incapaz de garantizar de forma efectiva la autonomía del paciente.

Asimismo, se ha constatado que el marco normativo vigente introduce avances relevantes en materia de transparencia, supervisión y control, pero no resuelve plenamente el problema de la inteligibilidad efectiva de la información en contextos de alta complejidad tecnológica.

En consecuencia, la principal aportación de este trabajo radica en la identificación de una insuficiencia estructural del modelo jurídico actual, que no puede ser corregida mediante ajustes normativos puntuales. Por el contrario, resulta necesario avanzar hacia una reinterpretación del derecho a la información del paciente que incorpore criterios de comprensión efectiva, explicabilidad operativa y comunicación adaptada a las características

y capacidades del destinatario. En este sentido, la problemática analizada no se limita al ámbito del derecho sanitario, sino que incide directamente en la protección de derechos fundamentales, en particular en la integridad personal y el derecho al respeto de la vida privada.

Este replanteamiento no implica necesariamente la creación de nuevos derechos, sino una reinterpretación de los existentes orientada a garantizar su efectividad en contextos tecnológicos complejos.

Desde esta perspectiva, el consentimiento informado debe evolucionar desde un modelo centrado en la mera transmisión de información hacia un modelo orientado a garantizar la capacidad real del paciente para entender, valorar y decidir en entornos mediados por inteligencia artificial. Esta evolución constituye un desafío tanto para el derecho sanitario como para la regulación de la inteligencia artificial, y abre nuevas líneas de investigación en torno a la protección efectiva de los derechos fundamentales en contextos digitales.

En ausencia de esta adaptación, el consentimiento informado corre el riesgo de convertirse en una garantía meramente formal en entornos clínicos digitalizados, lo que compromete su función como instrumento de protección de la autonomía del paciente. Este replanteamiento resulta especialmente relevante en un contexto de creciente digitalización de la práctica clínica, en el que la efectividad de los derechos no depende únicamente del acceso a la información, sino de la capacidad real de comprenderla y utilizarla en la toma de decisiones.

Referencias

Barrio Cantalejo, I. M., Simón Lorda, P., Melguizo, M., & Escalona, I. (2008). Validación de la escala INFLESZ para evaluar la legibilidad de los textos dirigidos a pacientes. *Anales del Sistema Sanitario de Navarra*, 31(2), 135–152. https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1137-66272008000300004

Beauchamp, T. L., & Childress, J. F. (2019). *Principles of biomedical ethics* (8th ed.). Oxford University Press.

Centers for Disease Control and Prevention. (2024). Addressing bias in artificial intelligence in healthcare. https://www.cdc.gov/pcd/issues/2024/24_0245.htm

Cohen, I. G., & Slottje, P. (2023). Artificial intelligence and the law of informed consent. En *StatPearls*. StatPearls Publishing. <https://www.ncbi.nlm.nih.gov/books/NBK613199/>

Council of Europe. (1997). *Convention on Human Rights and Biomedicine (Oviedo Convention)*. <https://www.coe.int/en/web/conventions/full-list?module=treaty-detail&treaty-num=164>

Council of Europe. (s. f.). Artificial intelligence and human rights. <https://www.coe.int/en/web/artificial-intelligence>

European Commission. (2022). *Proposal for a regulation on the European Health Data Space*. https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en

European Parliament and Council of the European Union. (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation)*. <https://eur-lex.europa.eu/eli/reg/2016/679/oj>

European Parliament and Council of the European Union. (2024). *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689>

Faden, R. R., & Beauchamp, T. L. (1986). *A history and theory of informed consent*. Oxford University Press.

Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and

recommendations. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>

Goodman, B., & Flaxman, S. (2017). European Union regulations on algorithmic decision-making and a “right to explanation”. *AI Magazine*, 38(3), 50–57.

Hernández Herrero, M. (2026). Información clínica y consentimiento informado (a propósito de la STEDH de 26 de junio de 2025). *Diario La Ley*, (10907).

Katz, J. (2002). *The silent world of doctor and patient*. Johns Hopkins University Press.

Ley 41/2002, de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica. (2002). <https://www.boe.es/buscar/act.php?id=BOE-A-2002-22188>

López-Almansa Beaus, L. (2015). *Finalidad, alcance y legibilidad de la información en el consentimiento informado* (Tesis doctoral, Universidad Católica de Valencia).

López-Almansa Beaus, L. (2024). El impacto de la inteligencia artificial en la mejora del consentimiento informado: análisis de la evolución de la legibilidad y la comprensión del paciente. *Derecho y Salud*, 34(1), 71–92.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2). <https://doi.org/10.1177/2053951716679679>

O'Neill, O. (2002). *Autonomy and trust in bioethics*. Cambridge University Press.

Simón Lorda, P. (2000). *El consentimiento informado: historia, teoría y práctica*. Triacastela.

Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. *International Data Privacy Law*, 7(2), 76–99. <https://doi.org/10.1093/idpl/ix005>

World Health Organization. (2021). *Ethics and governance of artificial intelligence for health*. <https://www.who.int/publications/i/item/9789240029200>