EL USO DE LA INTELIGENCIA ARTIFICIAL EN SISTEMAS DE VIDEOVIGILANCIA Y RECONOCIMIENTO FACIAL EN ESPACIOS PRIVADOS Y SEMIPRIVADOS: TENSIONES ENTRE SEGURIDAD, INTIMIDAD, HABEAS DATA Y GARANTÍAS PROBATORIAS EN EL CONTEXTO COLOMBIANO

CHRISTIAN URANGO GÓMEZ

Director

NICOLÁS ORTEGA TAMAYO

Magíster en Derecho

Trabajo de grado presentado como requisito parcial para optar al título de abogado

Pregrado en Derecho

Escuela de Derecho y Ciencias Políticas

Universidad Pontificia Bolivariana

Medellín

(2026)

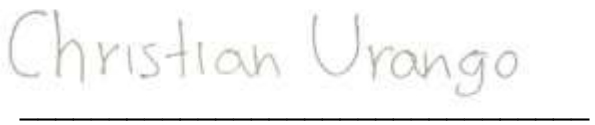
Declaración de originalidad

Fecha: 01/06/2026

Nombre del estudiante: Christian Urango Gómez

Declaro que este trabajo de grado no ha sido presentado con anterioridad para optar a un título, ya sea en igual forma o con variaciones, en esta o en cualquiera otra universidad.

Declaro, asimismo, que he respetado los derechos de autor y he hecho uso correcto de las normas de citación de fuentes, con base en lo dispuesto en las normas de publicación previstas en los reglamentos de la Universidad.



Estudiante: Christian Urango Gómez

C.C 1.003.308.581

ID 000423197

EL USO DE LA INTELIGENCIA ARTIFICIAL EN SISTEMAS DE VIDEOVIGILANCIA Y RECONOCIMIENTO FACIAL EN ESPACIOS CERRADOS, PRIVADOS Y SEMIPRIVADOS: TENSIONES ENTRE SEGURIDAD, INTIMIDAD, HABEAS DATA Y GARANTÍAS PROBATORIAS EN EL CONTEXTO COLOMBIANO.

THE USE OF ARTIFICIAL INTELLIGENCE IN VIDEO SURVEILLANCE AND FACIAL RECOGNITION SYSTEMS IN CLOSED, PRIVATE, AND SEMI-PRIVATE SPACES: TENSIONS BETWEEN SECURITY, PRIVACY, HABEAS DATA, AND EVIDENTIARY GUARANTEES IN THE COLOMBIAN CONTEXT.

RESUMEN

La presente investigación analiza las tensiones jurídicas entre la implementación de sistemas de videovigilancia asistidos por inteligencia artificial (IA) en espacios privados y semiprivados respecto de la garantía de los derechos fundamentales a la intimidad, el habeas data y el debido proceso en Colombia. Se parte de la hipótesis de que la transición de la videovigilancia pasiva a sistemas de monitoreo biométrico activo (reconocimiento facial 1:N y análisis comportamental) desborda el marco regulatorio actual centrado en la Ley 1581 de 2012 y genera riesgos significativos para la presunción de inocencia y la cadena de custodia probatoria. Metodológicamente, el estudio adopta un enfoque cualitativo y dogmático-jurídico, integrando una revisión crítica de la jurisprudencia constitucional reciente y el derecho comparado, específicamente el Reglamento de Inteligencia Artificial de la Unión Europea (“AI Act” Reglamento [UE] 2024/1689, 2024) y las normativas de la Administración del Ciberespacio de China (Cyberspace Administration of China, 2025). El trabajo concluye proponiendo criterios de delimitación técnica y jurídica para el uso de estas tecnologías, buscando equilibrar la seguridad preventiva con la protección de las libertades civiles en entornos que, aunque de acceso público, ostentan titularidad privada.

ABSTRACT

This research analyzes the legal tensions between the implementation of artificial intelligence (AI)-assisted video surveillance systems in private and semi-private spaces and the guarantee of fundamental rights to privacy, habeas data, and due process in Colombia. It starts from the hypothesis that the transition from passive video surveillance to active biometric monitoring systems (1:N facial recognition and behavioral analysis) goes beyond the current regulatory framework centered on Law 1581 of 2012 and generates significant risks for the presumption of innocence and the chain of custody of evidence. Methodologically, the study adopts a qualitative and dogmatic-legal approach, integrating a critical review of recent constitutional jurisprudence and comparative law, specifically the European Union's Artificial Intelligence Regulation ("AI Act" Regulation [EU] 2024/1689, 2024) and the regulations of the Cyberspace Administration of China (2025). The paper concludes by proposing technical and legal criteria for the use of these technologies, seeking to balance preventive security with the protection of civil liberties in environments that, although publicly accessible, are privately owned.

PALABRAS CLAVE:

Videovigilancia inteligente, Reconocimiento facial, Derecho a la intimidad, Habeas data, Principio de Proporcionalidad, Inteligencia Artificial, Sesgos algorítmicos.

KEYWORDS:

Intelligence video surveillance, Facial recognition, Right to privacy, Habeas data, Principle of proportionality, Artificial intelligence, Algorithmic biases.

INTRODUCCIÓN

La arquitectura de la vigilancia contemporánea ha superado los muros físicos del panóptico disciplinario descrito por Foucault (2008) para instalarse en una dimensión digital, algorítmica y omnipresente. En el contexto colombiano, la seguridad ciudadana y la protección de la propiedad privada han motivado una proliferación masiva de circuitos cerrados de televisión (CCTV) en espacios que, si bien son de titularidad privada, cumplen funciones sociales de alto tráfico, tales como centros comerciales, instituciones educativas y copropiedades residenciales. Sin embargo, la reciente incorporación de sistemas de inteligencia artificial (IA) en estos dispositivos ha transformado cualitativamente la naturaleza de la vigilancia: ya no se trata simplemente de ver o grabar hechos pasados, sino de *identificar*, *clasificar* y *predecir* conductas en tiempo real mediante el tratamiento automatizado de datos biométricos sensibles.

Este salto tecnológico plantea una tensión constitucional ineludible. Por un lado, existe un interés legítimo en la prevención del delito y la protección patrimonial; por otro, emerge la amenaza latente contra el núcleo esencial del derecho fundamental a la intimidad y la autodeterminación informática. A diferencia de la videovigilancia estatal en el espacio público, que cuenta con una regulación específica en el Código Nacional de Seguridad y Convivencia Ciudadana (Ley 1801 de 2016), la vigilancia en espacios privados y semiprivados opera en una zona gris, regida principalmente por el régimen general de protección de datos personales (Ley 1581 de 2012), el cual fue diseñado antes de la masificación de la biometría remota.

El problema jurídico central que aborda esta investigación radica en determinar de qué manera el uso de sistemas de videovigilancia con IA y reconocimiento facial en espacios cerrados, privados y semiprivados afecta las garantías de intimidad, habeas data y debido proceso probatorio en el ordenamiento jurídico colombiano. La relevancia de este cuestionamiento no es meramente teórica; tiene implicaciones prácticas directas en el proceso penal. La captura de imágenes biométricas sin el consentimiento explícito e informado del titular, o su tratamiento mediante algoritmos opacos ("cajas negras") propensos a sesgos discriminatorios, cuestiona la legalidad y autenticidad de la evidencia digital derivada de estos sistemas.

Autores como Byung-Chul Han (2013) advierten sobre el tránsito hacia una "sociedad de la transparencia" donde el individuo es despojado de sus secretos voluntariamente o por coacción sistémica. En el ámbito jurídico, esto se traduce en la erosión de la "expectativa razonable de privacidad", concepto desarrollado jurisprudencialmente que se ve desafiado cuando el ingreso a un centro comercial o a un conjunto residencial implica, *de facto*, la cesión de los rasgos faciales a una base de datos de cotejo algorítmico (Nissenbaum, 2010).

Para abordar esta problemática, el trabajo se estructura en tres capítulos que responden a objetivos específicos concatenados. Es imperativo, antes de entrar en el debate normativo, distinguir jurídicamente entre conceptos que a menudo se confunden: la videovigilancia pasiva frente a la analítica de video; la verificación biométrica (1:1) frente a la identificación remota (1:N) (ISO/IEC 19795-1, 2006); y los sistemas de seguridad perimetral frente a los de monitoreo comportamental.

La metodología empleada es de corte cualitativo y jurídico-dogmático, apoyada en el análisis de la normativa vigente, jurisprudencia colombiana y doctrina en derecho informático y penal. Se busca ofrecer criterios hermenéuticos que permitan al operador jurídico ponderar la eficacia probatoria de la tecnología frente a los derechos humanos en la era digital.

DELIMITACIÓN TÉCNICA Y OPERATIVA DE LA VIDEOVIGILANCIA CON INTELIGENCIA ARTIFICIAL

El propósito de este capítulo es construir la base conceptual y técnica necesaria para comprender el fenómeno de la vigilancia algorítmica. Para el derecho colombiano, se requiere una delimitación que distinga los sistemas de Circuito Cerrado de Televisión (CCTV) convencionales de aquellos dotados de redes neuronales, pues la intensidad de la injerencia en los derechos fundamentales varía según la capacidad del sistema para tratar datos biométricos de forma autónoma. Se examinarán las transiciones teóricas desde el panoptismo hasta el control digital, los riesgos operativos —falsos positivos, sesgos y opacidad algorítmica— que tensionan las garantías constitucionales en espacios semiprivados, y la diferencia técnica y jurídica entre verificación e identificación.

La videovigilancia ha transitado de una etapa analógica y pasiva a un ecosistema digital dinámico. En el modelo tradicional de CCTV, la cámara funcionaba como un "ojo extendido" cuya única función era el registro fotogramétrico para revisión humana posterior, afectando el derecho a la intimidad principalmente por la ubicación física de los dispositivos y la custodia de los soportes de almacenamiento. La integración de la Inteligencia Artificial ha transformado estas unidades en "sensores inteligentes" capaces de interpretar la realidad de forma autónoma. Un sistema moderno integra hardware de captura de alta resolución, un Sistema de Gestión de Video (VMS) y motores de inferencia algorítmica basados en Aprendizaje Profundo (Deep Learning), convirtiendo la imagen en una fuente de metadatos y vectores biométricos. Como advierte Deleuze (1992), este salto técnico representa el tránsito de una sociedad disciplinaria a una "sociedad de control", donde la vigilancia es fluida, continua y digitalmente omnipresente.

Esta evolución configura lo que denominamos "biovigilancia": la analítica de video inteligente permite detectar, seguir y clasificar sujetos de forma masiva, cotejando cientos de rostros por minuto contra listas de control (watchlists) en tiempo real. Cuando un sistema extrae patrones geométricos del rostro para crear una plantilla digital única, opera un tratamiento de datos biométricos catalogados como sensibles por el ordenamiento jurídico colombiano (Ley 1581 de 2012, art. 6; Circular Externa 002 de 2024, SIC). Estos datos se definen como aquellos cuyo uso indebido puede generar discriminación —origen racial, orientación política, datos biométricos— y su tratamiento está proscrito como regla general, dado el potencial riesgo de exclusión o perfilamiento indebido que conlleva (Corte Constitucional, Sentencia C-748 de 2011).

La sensibilidad del dato no reside únicamente en la imagen captada, sino en su transformación técnica: al extraer patrones geométricos del rostro se genera un vector inmutable que permite la identificación unívoca y remota del sujeto. Este vínculo inescindible entre el dato técnico y la individualidad biológica es lo que vincula el tratamiento de la IA con la dignidad humana, exigiendo protección reforzada y consentimiento cualificado (Corte Constitucional, Sentencia C-094 de 2020). En espacios semiprivados de acceso público —centros comerciales, universidades—, este monitoreo conductual evoca la "sociedad de la transparencia" descrita por Han (2013), donde el

ciudadano es despojado de su derecho al anonimato.

En consecuencia, la gestión de riesgos debe desplazarse hacia la responsabilidad demostrada: las organizaciones deben implementar un Estudio de Impacto de Privacidad (EIPD) que justifique, bajo el principio de proporcionalidad —entendido como el instrumento metodológico que pondera la idoneidad, necesidad y proporcionalidad en sentido estricto de toda medida restrictiva de derechos (Bernal Pulido, 2014)— por qué la seguridad no puede garantizarse con métodos menos invasivos. La captura de vectores biométricos sin consentimiento explícito y cualificado no solo vulnera el régimen de protección de datos, sino que convierte la seguridad en una intervención profunda en la esfera privada del individuo (Corte Constitucional, Sentencia C-094 de 2020).

Para efectos de determinar la proporcionalidad de la medida y la validez del consentimiento, es imperativo distinguir dos modos de operación biométrica, distinción recogida también en el Reglamento de Inteligencia Artificial de la Unión Europea (Reglamento [UE] 2024/1689):

A. Verificación o Autenticación (1:1): Proceso mediante el cual el sistema compara los datos biométricos de una persona con una única plantilla prealmacenada para confirmar su identidad. Es generalmente colaborativa y consensuada —por ejemplo, el acceso a una oficina mediante reconocimiento facial— y el consentimiento previo, expreso e informado es técnicamente viable y jurídicamente exigible (Ley 1581 de 2012, art. 9; ISO/IEC 19795-1, 2006).

B. Identificación Biométrica Remota (1:N): Proceso donde el sistema coteja los datos biométricos de un sujeto contra una base de datos masiva para determinar su identidad. Modalidad típica de cámaras en estadios o centros comerciales que escanean visitantes en tiempo real para verificar listas de restricción. Esta modalidad genera la mayor tensión con el Habeas Data —derecho fundamental cuyo contenido esencial fue fijado por la Corte Constitucional antes incluso de su desarrollo legal (Corte Constitucional, Sentencia T-729 de 2002)— y el derecho a la intimidad. En espacios semiprivados, obtener el consentimiento válido de cada transeúnte es fácticamente imposible o se obtiene mediante contratos de adhesión tácitos, cuya validez es cuestionable a la luz de la exigencia constitucional de consentimiento explícito y

facultativo para datos sensibles.

La delimitación operativa de estos sistemas exige abordar sus falencias técnicas intrínsecas, con repercusiones directas en el debido proceso y la calidad de la prueba. Las redes neuronales profundas utilizadas para el reconocimiento facial operan mediante capas de abstracción frecuentemente ininteligibles incluso para sus propios programadores (Goodfellow et al., 2016). Si un sistema emite una alerta señalando a un ciudadano como presunto hurtador con base en una coincidencia probabilística, surge un obstáculo para la defensa: ¿cómo se controvierte la decisión del algoritmo? Esta falta de explicabilidad afecta el derecho de contradicción, convirtiendo al algoritmo en un "testigo inimpugnable" cuya lógica interna es inaccesible para las partes.

La exactitud del sistema no es infalible y su margen de error tiene consecuencias directas sobre el principio de veracidad de la información (Ley 1581 de 2012). Dos métricas son determinantes:

- FPR (False Positive Rate): Identifica erróneamente a un inocente como sujeto de interés. Un FPR elevado en entornos semiprivados puede derivar en retenciones arbitrarias, vulnerando el derecho al buen nombre y la libertad de locomoción.
- FNR (False Negative Rate): No reconoce a un sujeto presente en la base de datos. Si la tasa de error es alta, el sistema no cumple su fin preventivo, privando de proporcionalidad la captura masiva de datos sensibles.

Estudios como el informe NISTIR 8280 (Grother et al., 2019) han demostrado que el desempeño algorítmico varía drásticamente según la demografía: algunos algoritmos presentan tasas de FPR hasta 100 veces mayores para personas de piel oscura, asiáticas o en grupos etarios extremos respecto a rostros caucásicos. Jurídicamente, esto configura un riesgo de discriminación algorítmica contrario al artículo 13 de la Constitución Política. Por ello, toda EIPD debe exigir que los responsables documenten estas métricas desagregadas por cohorte demográfica.

De esta forma, la introducción de metadatos y superposiciones analíticas (overlays) sobre el video crudo plantea desafíos probatorios bajo la Ley 906 de 2004: ¿el vídeo exportado contiene únicamente la imagen original o incluye las etiquetas generadas por

el algoritmo? Cualquier inferencia algorítmica superpuesta compromete la integridad de la prueba si no se garantiza una trazabilidad estricta mediante función hash inalterable desde el momento de la captura, separando evidencia visual de interpretaciones probabilísticas para evitar que el juez confunda una "inferencia algorítmica" con un "hecho demostrado".

La intensidad de la afectación a los derechos fundamentales se modula según el entorno físico donde se captura el dato. En el derecho colombiano, la licitud de la biovigilancia depende de la "expectativa razonable de privacidad", criterio jurisprudencial que evalúa si un ciudadano puede confiar legítimamente en que sus rasgos biométricos no serán procesados de forma automatizada en un lugar determinado (Corte Constitucional, Sentencia T-144 de 2024). Bajo este marco, se propone la siguiente categorización:

a) Espacios Privados: Domicilio y recintos de vida íntima. La expectativa de privacidad es absoluta; la instalación de sistemas de reconocimiento facial sin consentimiento explícito u orden judicial es inconstitucional, pues el dato biométrico goza de una "esfera de reserva" infranqueable vinculada directamente a la dignidad humana.
b) Espacios Semiprivados: Lugares cerrados donde un conjunto de personas comparte una actividad y el acceso al público es restringido, sin ser completamente privados porque las acciones de los individuos tienen "repercusiones sociales" (Corte Constitucional, Sentencia T-407 de 2012). Ejemplos: oficinas, salones de clase, áreas internas de conjuntos residenciales. La vigilancia permanente en estos espacios debe someterse a juicio de proporcionalidad estricta.
c) Espacios Abiertos de Propiedad Privada: Su naturaleza abierta al público genera una expectativa de anonimato relativo: el ciudadano asume que puede ser visto, pero no que su identidad será vectorizada y almacenada permanentemente.
d) Espacios Públicos: "Lugar de uso común en el que los ciudadanos ejercen numerosos derechos y libertades" (Corte Constitucional, Sentencia SU-360 de 1999). Aunque este trabajo no se centra en la vía pública, su definición es

necesaria para comprender el "Efecto Inhibitorio" (Chilling Effect): el temor a ser identificado y perfilado puede disuadir el ejercicio de libertades como la protesta o la libre asociación, transformando la conducta social por el solo hecho de sentirse "leído" por un algoritmo.

La videovigilancia con IA no constituye una evolución lineal del CCTV, sino una ruptura ontológica en la forma de monitorear el espacio. Al dotar a las cámaras de capacidades de identificación (1:N) y análisis predictivo, se transforma la naturaleza del espacio semiprivado y se reduce la expectativa de anonimato del ciudadano. Los mayores riesgos no residen en la verificación (1:1), sino en la identificación remota indiscriminada, que choca con los principios de necesidad, proporcionalidad y consentimiento cualificado del régimen colombiano de protección de datos sensibles. Esta delimitación permite avanzar hacia el análisis del marco normativo vigente y sus posibles lagunas.

SISTEMATIZACIÓN DEL MARCO NORMATIVO Y JURISPRUDENCIAL APLICABLE AL USO DE LA IA EN LA VIDEOVIGILANCIA Y RECONOCIMIENTO FACIAL EN EL CONTEXTO COLOMBIANO

La transición hacia arquitecturas de biovigilancia mediante Deep Learning exige una relectura del derecho fundamental a la intimidad y al habeas data, no solo como garantías de reserva, sino como facultades de autodeterminación frente al procesamiento automatizado de rasgos biométricos inmutables. Este capítulo sistematiza el marco normativo y la línea jurisprudencial aplicable, integrando disposiciones constitucionales, leyes estatutarias, lineamientos de la Superintendencia de Industria y Comercio (SIC) y precedentes de las altas cortes.

El núcleo de la protección frente a la vigilancia algorítmica se halla en el artículo 15 de la Constitución Política de 1991, que consagra el derecho fundamental a la intimidad personal y familiar, al buen nombre y al habeas data, derechos que la Corte Constitucional ha vinculado de manera inescindible a la dignidad humana (Corte Constitucional, Sentencia C-1011 de 2008) y cuyo contenido esencial comprende distintas esferas de reserva —íntima, privada y sociable— cuya intensidad de protección

varía según el contexto de exposición del dato (Corte Constitucional, Sentencia T-787 de 2004).

La jurisprudencia ha reconocido que la intimidad se proyecta más allá del domicilio privado a través del criterio de la "expectativa razonable de privacidad": herramienta hermenéutica que evalúa si, dadas las circunstancias de tiempo, modo y lugar, un ciudadano puede confiar legítimamente en que su actividad no será monitoreada, clasificada o identificada por sistemas automatizados (Corte Constitucional, Sentencias T-280 de 2022 y C-406 de 2022).

La Ley 1581 de 2012 operacionaliza el derecho al habeas data y somete el tratamiento de datos biométricos a protección reforzada, en tanto el rostro es un rasgo inmutable que permite la identificación unívoca del titular. Ante los algoritmos de IA, los principios rectores de esta ley adquieren una nueva dimensión:

El Principio de Finalidad trasciende la mera declaración de propósitos para exigir una justificación material de la necesidad de emplear analítica avanzada; la organización debe demostrar por qué la seguridad no puede alcanzarse mediante métodos convencionales menos invasivos (Ley 1581 de 2012, art. 4, lit. b; SIC, Circular Externa 002 de 2024, sección 3).

El Principio de Libertad impone que, ante la biometría remota, el consentimiento sea una manifestación previa, expresa y plenamente informada sobre la vectorización del dato sensible —no una simple adhesión tácita— (Ley 1581 de 2012, art. 4, lit. c; Corte Constitucional, Sentencia C-748 de 2011)

El Principio de Veracidad o Calidad se enfrenta directamente a los márgenes de error del aprendizaje profundo: las tasas de falsos positivos (FPR) y falsos negativos (FNR) son riesgos jurídicos que pueden derivar en perfilamientos erróneos o actos discriminatorios proscritos por el ordenamiento (Ley 1581 de 2012, art. 4, lit. d; SIC, Circular Externa 002 de 2024, sección V).

El Principio de Transparencia impone el deber de garantizar al titular información sobre la lógica del tratamiento, exigiendo niveles de explicabilidad técnica frente a la "caja negra" algorítmica (Ley 1581 de 2012, art. 4, lit. e; SIC, Circular Externa 002 de 2024,

sección IV)

Por último, el Principio de Seguridad instituye un estándar de debida diligencia que obliga a adoptar medidas técnicas, humanas y administrativas auditables para evitar la adulteración o acceso no autorizado a las plantillas biométricas (Ley 1581 de 2012, art. 4, lit. g; SIC, Circular Externa 002 de 2024, sección VIII)

Ante el vacío normativo específico sobre IA, la SIC emitió instrucciones sobre el tratamiento de datos personales en sistemas de inteligencia artificial, adoptando la definición de sistema de IA como aquel basado en máquinas que genera salidas con influencia en entornos reales o virtuales (SIC, Circular Externa 002 de 2024).

Bajo el modelo de "responsabilidad demostrada" que articula este instrumento, las organizaciones que implementen reconocimiento facial están obligadas a realizar un Estudio de Impacto en la Privacidad y Protección de Datos (EIPD) que incluya: descripción detallada de las operaciones de tratamiento, evaluación de necesidad y proporcionalidad, y gestión de riesgos técnicos como los sesgos demográficos. La circular precisa, además, que la información personal accesible en internet no tiene naturaleza pública, prohibiendo su uso para entrenar modelos sin autorización expresa del titular.

La Corte ha fijado un estándar de protección reforzada para los datos que, por su naturaleza, puedan propiciar escenarios de segregación, proscribiendo su tratamiento como regla general e imponiendo un escrutinio constitucional estricto para justificar cualquier recolección de rasgos biométricos (Corte Constitucional, Sentencia C-748 de 2011).

En materia de vigilancia estatal, la Sala Plena determinó que la fuerza pública no puede acceder de manera irrestricta a los circuitos cerrados de vigilancia privada bajo una noción ambigua de "prevención"; el acceso para fines de identificación o judicialización exige orden judicial previa y debe ser ejecutado exclusivamente por la Policía Judicial en el marco de una investigación penal (Corte Constitucional, Sentencia C-406 de 2022).

Respecto a los sujetos de especial protección, la Corte señaló que el monitoreo

permanente en espacios semiprivados donde se gestiona información de Niños, Niñas y Adolescentes (NNA) constituye una injerencia desproporcionada, derivando la obligación técnica de implementar mecanismos de desidentificación o anonimización de terceros cuando se ejerce el derecho de acceso a la propia información visual (Corte Constitucional, Sentencia T-144 de 2024.

La videovigilancia inteligente se erige también como elemento material probatorio cuya validez depende de la preservación de su autenticidad e integridad (Ley 906 de 2004, arts. 254–257; Fiscalía General de la Nación, 2016). En este campo, la Sala Penal de la Corte Suprema de Justicia ha establecido un hito jurisprudencial que moderniza la valoración de la evidencia digital.

En virtud del principio de libertad probatoria, la ausencia del código hash no genera la invalidez automática de la prueba: considerarlo un requisito sine qua non equivaldría a instaurar una "tarifa legal encubierta", proscrita en el sistema acusatorio colombiano (Corte Suprema de Justicia, Sentencia SP248-2025). En su lugar, la integridad se acredita mediante un análisis funcional que reconstruye la trazabilidad, evalúa la confiabilidad del método de obtención y se apoya en testigos técnicos o informes periciales que expliquen el funcionamiento del algoritmo, siempre que se garantice una contradicción eficaz y no se altere el contenido sustancial de la información captada (Corte Suprema de Justicia, Sentencia SP248-2025).

ARQUITECTURAS DE SALVAGUARDA Y GOBERNANZA ALGORÍTMICA: INSTRUMENTOS PARA LA OPERATIVIZACIÓN DEL DERECHO A LA INTIMIDAD Y EL HABEAS DATA FRENTE A LA VIDEOVIGILANCIA CON IA EN COLOMBIA

La implementación de sistemas de videovigilancia asistidos por IA exige transitar de una protección reactiva a una gobernanza preventiva y auditable.

Este capítulo operacionaliza los hallazgos dogmáticos previos mediante tres herramientas: el modelo CAP, las matrices EIPD y una propuesta de articulado regulatorio, con el objetivo de convertir la proporcionalidad en un procedimiento técnico verificable bajo el marco de responsabilidad demostrada exigido por la SIC.

El modelo CAP traduce los principios de libertad, finalidad y veracidad de la Ley

1581 de 2012 en especificaciones de diseño bajo el enfoque de Privacy by Design —que concibe la protección de datos como un requisito de arquitectura desde la fase de concepción del sistema y no como un ajuste correctivo posterior (Cavoukian, 2010)—, articulando tres componentes en el contexto colombiano:

- La evitabilidad: Se garantiza que el tratamiento de datos sensibles sea efectivamente facultativo (Ley 1581 de 2012, art. 6). Un sistema solo es constitucionalmente legítimo si ofrece un canal alternativo no biométrico —tarjeta RFID o código QR— cuya latencia y usabilidad no penalicen al usuario; la "fricción irrazonable" como mecanismo de coacción para la entrega de datos faciales invalida el consentimiento y vulnera la autonomía del titular.
- La cláusula de no-permanencia: Se exigen protocolos de supresión automática de vectores biométricos una vez cumplida la finalidad de la sesión de inferencia. En ausencia de coincidencia con listas de control autorizadas, la persistencia de plantillas matemáticas configura un riesgo de agregación y un uso secundario ilícito del dato, expresamente prohibido por la Circular 002 de 2024.
- La proporcionalidad del umbral: Se determina que la identificación 1:N sea excepcional y esté supeditada a umbrales de desempeño auditables. Con base en el estándar NISTIR 8280, ningún sistema debe operar en entornos masivos con un FPR superior a 1×10^{-5}, y toda decisión restrictiva de derechos debe estar sujeta a protocolo de doble revisión humana.

La obligatoriedad de la EIPD para tratamientos de alto riesgo no deriva únicamente de la Circular 002 de 2024 de la SIC, sino que tiene raíz reglamentaria en el Decreto Único Reglamentario del Sector Comercio, que establece las condiciones de licitud del tratamiento y las obligaciones del responsable frente a datos sensibles (Decreto 1074 de 2015, arts. 2.2.2.26.1 y ss.). Bajo este marco, las matrices que se presentan a continuación vinculan la taxonomía de riesgos de Solove (2006) con medidas de mitigación operativa para los tres escenarios objeto de estudio.

Copropiedades Residenciales y Propiedad Horizontal

Actividad de	Riesgo	Severidad /	Mitigación	Control

Tratamiento	(Solove)	Probabilidad	(CAP)	Probatorio
Captura en zonas de recreo común.	Vigilancia e Intrusión.	Muy Alta / Alta	Delimitación de campos visuales. Prohibición de IA en áreas íntimas.	Registro de ubicación de cámaras.
Control de acceso.	Identificación y Error.	Media / Media	Evitabilidad: Acceso por token físico obligatorio para residentes.	Log de acceso disociado de imagen.
Almacenamiento de registros.	Agregación	Alta / Baja	No-permanencia: Purga de vectores en 24 horas	Cadena de custodia (SP248-25).

En el ámbito de la propiedad horizontal, la protección del domicilio es prioritaria. La ubicación física de las cámaras determina la naturaleza de la información captada y exige una delimitación estricta de sus campos visuales (Corte Constitucional, Sentencia T-114 de 2018).

Instituciones de educación y campus universitarios

Actividad de Tratamiento	Riesgo Identificado (Solove)	Severidad / Probabilidad	Medida de Mitigación (CAP)	Control Probatorio y Operativo
Analítica de video para control perimetral del	Distorsión y Sesgo: Falsos positivos por variaciones en	Muy Alta / Media	Proporcionalidad: Auditoría de sesgos bajo estándar NISTIR	Declaración de confiabilidad del algoritmo por el

campus.	edad, tono de piel o vestimenta.		8280. Doble verificación humana de alertas de seguridad.	fabricante antes de su implementación en el campus.
Monitoreo de áreas comunes (bibliotecas, pasillos).	Exposición y Divulgación: Captación de conductas privadas de estudiantes o NNA.	Alta / Alta	Zonas Sensibles: Prohibición absoluta de IA en consultorios, baños, vestidores y áreas de bienestar.	Aplicación de algoritmos de desenfoque (blurring) automático para terceros en grabaciones generales.
Acceso a laboratorios mediante reconocimiento facial.	Exclusión e Inseguridad: Fallas técnicas que impiden el ejercicio de derechos académicos.24	Media / Baja	Evitabilidad: Carnet físico con chip como canal alternativo idéntico en tiempo y usabilidad al biométrico.	Bitácora inalterable de acceso con hash para auditoría de denegaciones de entrada injustificadas.

Los campus universitarios son espacios de libre desarrollo de la personalidad donde circulan NNA. La vigilancia permanente sobre menores es desproporcionada y obliga al uso de técnicas de anonimización facial en entornos de alta sensibilidad (Corte Constitucional, Sentencia T-144 de 2024).

Centros comerciales y retail inteligente

Actividad de	Riesgo	Severidad /	Medida de	Control

Tratamiento	Identificado (Solove)	Probabilidad	Mitigación (CAP)	Probatorio y Operativo
Detección de "sujetos de interés" (watchlists).	Identificación 1:N y Error: Perfilamiento masivo y retenciones arbitrarias de clientes.	Muy Alta / Alta	Proporcionalidad Reforzada: Las listas de control deben basarse en órdenes judiciales, no en criterios privados de "sospecha".	En caso de retención, el responsable debe presentar el análisis funcional del algoritmo ante la policía.
Rastreo de rutas de compra y aforo.	Agregación e Inseguridad: Creación de mapas de calor vinculados a identidades reales.	Alta / Media	No-permanencia: Supresión inmediata de vectores biométricos al finalizar la sesión de inferencia en ausencia de delito.	Los metadatos analíticos deben almacenarse en repositorios segregados de las imágenes originales.
Interacción de seguridad privada con fuerza pública.	Acceso no autorizado y Excesos: Entrega masiva de datos a la policía sin control judicial.	Muy Alta / Baja	Trazabilidad-Custodia: Registro de cada acceso estatal, supeditado a investigación penal en curso y orden previa.	El material entregado debe ser el "video crudo" para preservar la mejor evidencia bajo Ley 906 de 2004.

Este entorno presenta el mayor riesgo de "función deslizante" por el incentivo económico de monetizar datos de comportamiento. La Circular 002 de 2024 prohíbe expresamente el uso de información accesible al público para finalidades no autorizadas.

El video procesado por IA no es un hecho, sino una representación probabilística. Para cumplir con el debido proceso (Ley 906 de 2004), su incorporación como prueba debe someterse al análisis funcional establecido en la Sentencia SP248-2025 de la Corte Suprema de Justicia, que exige acreditar: i) la confiabilidad del algoritmo utilizado; ii) la trazabilidad total desde la captura; y iii) el suministro del "video crudo" (raw footage) sin las etiquetas analíticas del sistema, permitiendo que el juez valore la imagen sin la inducción al error que puede generar una etiqueta de sospecha algorítmica.

Como síntesis operativa de la investigación, se propone el siguiente proyecto de norma para cerrar la brecha regulatoria en Colombia:

Artículo 1. Objeto. La presente ley tiene por objeto establecer el marco técnico-jurídico para la implementación y operación de sistemas de videovigilancia asistidos por inteligencia artificial en espacios privados y semiprivados, garantizando la protección del núcleo esencial de la intimidad, el habeas data y el debido proceso probatorio.

Artículo 2. Ámbito de Aplicación. Esta normativa vincula a todas las personas naturales o jurídicas, de derecho privado o público, que operen sistemas de identificación biométrica remota o analítica de video inteligente en recintos cerrados de acceso restringido o con función social.

Artículo 3. Principio de Veracidad Algorítmica y Calidad del Dato. La información generada por sistemas de IA debe ser exacta, verificable y comprensible. Ningún sistema de identificación biométrica podrá operar si su Tasa de Falsos Positivos (FPR) es superior a $1x10^{-5}$ en las pruebas de desempeño por cohorte demográfica. El tratamiento de datos parciales o que induzcan a error algorítmico queda expresamente prohibido.

Artículo 4. Derecho a la Alternativa (Evitabilidad). El suministro de datos biométricos faciales para el acceso o permanencia en recintos semiprivados será siempre facultativo. Los responsables deben garantizar un medio de acceso no invasivo cuya

usabilidad, latencia y eficiencia no penalice al usuario que opta por el canal no biométrico. Se prohíbe la coacción técnica para la entrega de datos sensibles.

Artículo 5. Minimización y No-permanencia. Se prohíbe el almacenamiento masivo y preventivo de plantillas o vectores biométricos de transeúntes. Los metadatos resultantes de la analítica de video deberán ser eliminados de forma automática e inmediata una vez finalizada la sesión de inferencia, salvo en casos de detección de incidentes de seguridad documentados que activen protocolos de policía judicial.

Artículo 6. Trazabilidad y Explicabilidad. Todo sistema basado en IA debe generar una bitácora de eventos inalterable que registre el historial de versiones del modelo, los puntajes de similitud y la identificación del operador humano responsable de la validación final. Los titulares tendrán derecho a conocer la lógica del tratamiento y los parámetros que fundamentaron una decisión automatizada en su contra.

Artículo 7. Interés Superior del Menor. Queda proscrito el tratamiento de datos biométricos de Niños, Niñas y Adolescentes (NNA) mediante sistemas de identificación algorítmica, salvo autorización expresa de sus representantes legales para fines exclusivos de protección vital o salud. En las grabaciones generales, el sistema deberá aplicar algoritmos de anonimización facial (blurring) de oficio para proteger la identidad de los menores.

Artículo 8. Delimitación de Zonas Sensibles. Se prohíbe la instalación de cámaras dotadas de IA en áreas donde exista una expectativa máxima de privacidad, tales como consultorios de salud, áreas de atención psicológica o legal, vestidores, servicios sanitarios y habitaciones de hotel.

Artículo 9. Análisis Funcional de la Evidencia Digital. Para que el material fílmico procesado por IA tenga validez en procesos judiciales o administrativos, el responsable deberá acreditar la trazabilidad funcional del dato y la confiabilidad del método de obtención conforme a la jurisprudencia de la Sala Penal de la Corte Suprema de Justicia. La integridad se evaluará verificando que el contenido sustancial de la imagen no fue alterado por procesos de síntesis algorítmica.

Artículo 10. Límites al Acceso por Autoridades de Policía. El acceso remoto, en

tiempo real o masivo de la Fuerza Pública a circuitos de vigilancia privada dotados de IA queda supeditado a la existencia de una investigación penal en curso y a la orden previa de autoridad judicial competente. Se prohíbe el acceso bajo justificaciones genéricas de prevención conforme a la Sentencia C-406 de 2022 de la Corte Constitucional.

Artículo 11. Potestad de Vigilancia. La Superintendencia de Industria y Comercio tendrá la facultad de ordenar la suspensión inmediata de cualquier sistema que no cuente con una EIPD actualizada, que presente sesgos demográficos discriminatorios comprobados o que incumpla el principio de evitabilidad.

Artículo 12. Responsabilidad Algorítmica y Civil. El responsable del tratamiento responderá por los daños causados a la honra, el buen nombre y la libertad derivados de falsos positivos generados por sistemas que no cumplan con los umbrales de exactitud o los protocolos de revisión humana previstos en esta ley.

El modelo CAP y las matrices EIPD constituyen el andamiaje mínimo necesario para armonizar la pretensión legítima de seguridad con los derechos a la intimidad y el habeas data. La integración del estándar NISTIR 8280 eleva el nivel de protección frente a la discriminación algorítmica, mientras que el reconocimiento del análisis funcional de la Corte Suprema asegura que la tecnología sea una herramienta al servicio de la justicia y no un obstáculo para la verdad procesal. El articulado propuesto cierra la brecha regulatoria transformando los lineamientos administrativos de la SIC en obligaciones legales exigibles: la videovigilancia con IA en espacios semiprivados solo será constitucionalmente legítima bajo un régimen de transparencia absoluta, supervisión humana efectiva y respeto irrestricto a la autonomía del individuo sobre su información biométrica.

CONCLUSIONES

La videovigilancia inteligente no es una evolución lineal del CCTV, sino una ruptura ontológica que transmuta los espacios semiprivados en laboratorios de biovigilancia. Al dotar a las cámaras de capacidades de identificación 1:N, se elimina el anonimato dinámico, lo que genera un efecto inhibitorio. Este fenómeno disuade a los ciudadanos de ejercer libertades fundamentales, como el libre desarrollo de la personalidad o la

protesta, por el solo temor de ser leídos y clasificados permanentemente por un algoritmo.

El uso de redes neuronales profundas introduce una opacidad algorítmica que desafía el principio de contradicción en el proceso penal. El algoritmo, al operar como una caja negra ininteligible, se convierte en un testigo inimpugnable. Por ello, bajo el estándar de la Sentencia SP248-2025, la validez judicial de estos registros no depende de requisitos técnicos absolutos como el código hash, sino de un "Análisis Funcional" que acredite la trazabilidad absoluta, la confiabilidad del método y la preservación del video crudo para garantizar que el juez no confunda una inferencia probabilística con un hecho demostrado.

En el ordenamiento colombiano, el interés legítimo en la seguridad no es una carta blanca para el tratamiento masivo de datos biométricos. La licitud de estos sistemas en espacios como centros comerciales o copropiedades depende estrictamente de la evitabilidad. Basándose en la Resolución 52185 de 2025 de la SIC, es imperativo que los responsables ofrezcan alternativas no invasivas cuya usabilidad no penalice al usuario. Cualquier coacción utilizada para obtener datos faciales invalida el consentimiento y constituye una vulneración al derecho de autodeterminación informática.

Ante la capacidad de la IA para realizar perfilamientos profundos, la reserva judicial debe ser el único blindaje contra el acceso estatal a los datos privados captados en espacios semiprivados. La Sentencia C-406 de 2022 establece el estándar mínimo: la Policía Nacional no puede acceder a circuitos privados bajo justificaciones genéricas de prevención. El acceso debe estar supeditado a una investigación penal específica, ser ejecutado por la Policía Judicial y contar con una orden previa de un juez de control de garantías, evitando que el Estado consolide un sistema de vigilancia ubicuo y desproporcionado.

REFERENCIAS

Libros

1. Bernal Pulido, C. (2014). El principio de proporcionalidad y los derechos fundamentales (4.ª ed.). Universidad Externado de Colombia.
2. Foucault, M. (2008). Vigilar y castigar: Nacimiento de la prisión. (A. Garzón del Camino, Trad.) Madrid: Siglo XXI Editores.
3. Han, B.-C. (2013). La sociedad de la transparencia. (R. Gabás, Trad.) Barcelona: Herder Editorial.
4. Nissenbaum, H. (2010). Privacy in context: Technology, policy, and the integrity of social life. Stanford, CA: Stanford University Press.
5. Tamayo Jaramillo, J. (2007). Tratado de responsabilidad civil (2.ª ed., Tomo I). Legis.

Artículos

6. Cavoukian, A. (2010). Privacy by Design: The 7 Foundational Principles. Information and Privacy Commissioner of Ontario.
7. Deleuze, G. (1992). Postscript on the societies of control. October (59), 3 - 7.
8. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
9. Grother, P., Ngan, M., & Hanaoka, K. (2019). Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects (NISTIR 8280). National Institute of Standards and Technology.
10. Solove, D. J. (2006). A taxonomy of privacy. University of Pennsylvania Law Review, 154 (3), 477 - 560.

Normas

11. China. Cyberspace Administration of China. (2025). Measures for the Security Management of Facial Recognition Technology Applications. Beijing.
12. Colombia. Asamblea Nacional Constituyente. (1991). Constitución Política de Colombia.

13. Colombia. Congreso de la República. Ley 906. (2004). Por la cual se expide el Código de Procedimiento Penal.
14. Colombia. Congreso de la República. Ley 1581. (2012). Estatutaria de Protección de Datos Personales.
15. Colombia. Fiscalía General de la Nación. (2016). Manual de procedimientos del sistema de cadena de custodia.
16. Colombia. Superintendencia de Industria y Comercio. Circular Externa 002. (2024).
17. Colombia. Superintendencia de Industria y Comercio. Resolución 52185 de 2025. (2025).
18. ISO/IEC 19795-1. (2006). Information technology — Biometric performance testing and reporting — Part 1: Principles and framework.
19. ISO/IEC 29134. (2017). Information technology — Security techniques — Guidelines for privacy impact assessment
20. Unión Europea. Parlamento Europeo y del Consejo. Reglamento (UE) 2024/1689. (2024).

Sentencias

21. Corte Constitucional. (2003). Bogotá D.C. Sentencia T-729. M.P. Eduardo Montealegre Lynett.
22. Corte Constitucional. (2004). Bogotá D.C. Sentencia T-787. M.P. Rodrigo Escobar Gil.
23. Corte Constitucional. (2008). Bogotá D.C. Sentencia C-1011. M.P. Jaime Córdoba Traviño.
24. Corte Constitucional. (2011). Bogotá D.C. Sentencia C-748. M.P. Jorge Ignacio Pretelt Chaljub.
25. Corte Constitucional. (2012). Bogotá D.C. Sentencia T-407. M.P. Mauricio González Cuervo.
26. Corte Constitucional. (2020). Bogotá D.C. Sentencia C-094. M.P. José Fernando Reyes Cuartas.
27. Corte Constitucional. (2022). Bogotá D.C. Sentencia C-406. M.P. Cristina Pardo

Schlesinger.
28. Corte Constitucional. (2022). Bogotá D.C. Sentencia T-280. M.P. José Fernando Reyes Cuartas.
29. Corte Constitucional. (2024). Bogotá D.C. Sentencia T-144. M.P. Juan Carlos Cortés González.
30. Corte Suprema de Justicia de Colombia. Sala de Casación Penal. (2025). Bogotá D.C. Sentencia SP-248. M.P. Gerson Chaverra Castro.