RED IBEROAMERICANA DE PROTECCION DE DATOS

Red Iberoamericana
Rede Ibero-Americana

Estándares Iberoamericanos sobre el uso de Datos Personales en Sistemas de Inteligencia Artificial

Publicada en noviembre 2026.



Preámbulo

La inteligencia artificial está transformando profundamente la forma en que se recolectan, utilizan y reutilizan los datos personales. Su desarrollo ha incrementado la capacidad para generar inferencias, automatizar decisiones, reutilizar información para nuevas finalidades y combinar grandes volúmenes de datos en contextos cada vez más complejos. Estos avances no alteran la vigencia de los principios de protección de datos personales, pero sí exigen precisar cómo deben aplicarse cuando los tratamientos incorporan sistemas de inteligencia artificial.

Con esta finalidad, y como resultado de un proceso de elaboración, consulta y deliberación técnica en el que han participado sus autoridades miembros, la Red Iberoamericana de Protección de Datos aprueba los presentes Estándares iberoamericanos sobre el uso de datos personales en sistemas de inteligencia artificial (IA) (en adelante, «el Estándar»), concebidos como un instrumento orientador para favorecer una interpretación común del régimen de protección de datos en el contexto de la inteligencia artificial en el ámbito iberoamericano y contribuir al desarrollo de una inteligencia artificial respetuosa con los derechos y libertades de las personas.

El Estándar constituye una iniciativa pionera al desarrollar de forma sistemática las especialidades que introduce la inteligencia artificial desde la perspectiva de la protección de datos personales. No crea un régimen ético- jurídico paralelo para la inteligencia artificial ni pretende sustituir las legislaciones nacionales o los Estándares de Protección de Datos Personales para los Estados Iberoamericanos. Su finalidad es complementar ese marco común allí donde la utilización de sistemas de inteligencia artificial plantea cuestiones específicas que requieren criterios adicionales para la aplicación de los principios, derechos y obligaciones propios de la protección de datos, sin perjuicio de las exigencias adicionales que establezcan las legislaciones nacionales. El Estándar no reduce ni relaja el nivel de protección reconocido por ese marco común: las especialidades que desarrolla operan como criterios adicionales para su aplicación y no como excepciones a él.

La protección de los datos personales constituye un derecho fundamental de las personas y se asienta en la autonomía informativa, entendida como la facultad de decidir sobre la utilización de los propios datos personales y de conservar el control sobre su tratamiento; su reconocimiento en los ordenamientos de la región responde a la necesidad de preservar la dignidad humana, la autonomía personal, la igualdad y las demás libertades que el tratamiento de la información puede comprometer. La intensificación de la capacidad de inferir, predecir y decidir que caracteriza a los sistemas de inteligencia artificial hace especialmente exigible preservar esa facultad, pues los tratamientos que los incorporan pueden producir datos y decisiones sobre la persona sin su intervención directa y pueden incidir en las decisiones que le afectan. Esa condición de derecho fundamental constituye el parámetro desde el que han de leerse e interpretarse las disposiciones del Estándar. El Estándar no crea ni sustituye los derechos reconocidos por el régimen general de protección de datos y por las legislaciones nacionales: precisa el modo en que su ejercicio se proyecta cuando el tratamiento incorpora sistemas de inteligencia artificial.

El Estándar se apoya en un cuerpo normativo de referencia preexistente, al que alude mediante expresiones equivalentes cuyo alcance queda fijado en las disposiciones generales del Estándar relativas a su objeto. El Estándar no aspira a sustituir ese cuerpo normativo ni a anticipar las decisiones que a cada ordenamiento corresponden; su función es precisar cómo han de aplicarse las obligaciones y

garantías ya existentes cuando el tratamiento incorpora sistemas de inteligencia artificial. Entre los responsables se encuentran los organismos del sector público, cuya actuación se rige además por reglas propias del ordenamiento nacional en materia de habilitación del tratamiento, de comunicación de datos entre organismos o hacia particulares, de publicidad de la información y de acceso a la información pública, así como por los deberes de transparencia que ese ordenamiento les imponga respecto del uso de sistemas de inteligencia artificial; conforme a su naturaleza orientadora y complementaria, el Estándar se dirige también a esos responsables sin perjuicio de tales reglas, sin desplazarlas ni condicionarlas y atendiendo a las particularidades que la actuación pública imprime a los tratamientos que incorporan sistemas de inteligencia artificial.

El Estándar parte de una idea esencial: su objeto no son los sistemas de inteligencia artificial considerados de forma abstracta, sino los tratamientos de datos personales que los incorporan. De esa premisa deriva el enfoque de ciclo de vida que vertebra el instrumento y que el Estándar desarrolla al ocuparse de los tratamientos y de su ciclo de vida, proyectando las obligaciones sobre cada tratamiento y sobre las finalidades para las que se realiza. El principio de lealtad adquiere en este ámbito una función de contraste especialmente relevante, pues la inteligencia artificial puede acentuar la distancia entre el momento en que las personas facilitan sus datos personales y el modo en que estos se emplean después; el criterio conforme al cual esa lealtad se aprecia se establece en el Estándar al regular la lealtad, la finalidad y la compatibilidad.

La metodología seguida responde a una decisión deliberada de especialización normativa: el Estándar desarrolla únicamente aquellas cuestiones en las que la inteligencia artificial exige una modulación de las reglas generales de protección de datos y aporta criterios para resolver los problemas específicos que esta plantea, evitando duplicidades innecesarias.

Asimismo, el Estándar adopta un enfoque basado en el riesgo y en la responsabilidad proactiva. Las medidas previstas deberán aplicarse de forma proporcionada, atendiendo a la naturaleza del tratamiento, a los datos utilizados, a las finalidades perseguidas, al contexto de utilización y al impacto que el tratamiento pueda producir sobre los derechos y libertades de las personas. Este enfoque pretende reforzar la protección efectiva sin imponer cargas innecesarias ni obstaculizar el desarrollo y la adopción responsable de la inteligencia artificial.

Ese enfoque se proyecta sobre operaciones de valoración de naturaleza e intensidad distintas: la determinación de la base jurídica y, cuando proceda, la correspondiente ponderación; el análisis de la necesidad y proporcionalidad; la identificación y valoración de los riesgos para los derechos y libertades de las personas y la determinación de las medidas adecuadas para gestionarlos; y, cuando resulte exigible, la evaluación de impacto. Estas operaciones responden a finalidades propias y son complementarias entre sí: ninguna sustituye a las demás, aunque los elementos documentados en una puedan aprovecharse en otra cuando su objeto coincida, y la exigibilidad de cada una se rige por el ordenamiento que resulte aplicable al tratamiento, sin que el Estándar altere el momento en que nace ni las condiciones a que se sujeta. El recurso al interés legítimo como base jurídica plantea, en este ámbito, dificultades propias: la escala con que pueden reutilizarse los datos personales, la distancia entre el momento de su obtención y el de su empleo posterior, y el carácter poco perceptible que estos tratamientos suelen presentar para las personas. Estas circunstancias deberán tenerse en cuenta en la ponderación, que habrá de ser especialmente rigurosa cuando, por la naturaleza de los datos, la escala del tratamiento, sus finalidades o sus posibles efectos, puedan producirse mayores riesgos para los derechos y libertades de las personas. La graduación de las medidas atiende asimismo a los efectos del

tratamiento sobre las personas, a cuyo efecto el Estándar emplea de manera deliberada dos formulaciones distintas: cuando exige efectos jurídicos o una afectación significativa, delimita el supuesto conforme a la noción que el régimen general de protección de datos atribuye a esas expresiones, sin que la afectación significativa quede circunscrita a consecuencias de naturaleza jurídica, pues comprende también las de orden material, social o económico que la utilización de sistemas de inteligencia artificial es susceptible de producir; cuando se refiere a los efectos sin calificarlos, alude a toda consecuencia que el tratamiento pueda producir sobre la persona, sin que ello suponga uniformar la intensidad de las medidas exigibles, que se determina conforme al criterio de proporcionalidad a los riesgos y no mediante una calificación previa de los efectos.

Con esa misma finalidad, el texto emplea formulaciones tecnológicamente neutrales — esto es, independientes de productos, marcas o proveedores determinados— y evita imponer soluciones técnicas concretas cuando basta con definir el objetivo jurídico que debe alcanzarse, sin que ello comporte atenuación alguna de las garantías reconocidas a las personas. Su vocación es ofrecer un marco estable, capaz de acompañar la evolución de la inteligencia artificial sin generar barreras innecesarias a la innovación ni perder utilidad como referencia para autoridades de protección de datos, legisladores, responsables y encargados del tratamiento, proveedores de sistemas de inteligencia artificial y demás organizaciones de los sectores público y privado.

El Estándar se interpretará de forma coherente con el régimen general de protección de datos, constituido por los Estándares de Protección de Datos Personales para los Estados Iberoamericanos, con los demás documentos y criterios adoptados por la Red Iberoamericana de Protección de Datos, con las legislaciones nacionales aplicables y con los instrumentos internacionales que resulten pertinentes. Aspira a fortalecer la convergencia regulatoria en la región, facilitar la aplicación consistente de la protección de datos en el contexto de la inteligencia artificial y promover un modelo de innovación responsable que sitúe la confianza, la protección de las personas y el respeto de los derechos fundamentales como condiciones esenciales para el desarrollo y la adopción de la inteligencia artificial.

Algunas manifestaciones de la inteligencia artificial suscitan una preocupación específica y han sido objeto de trabajos propios en el seno de la Red Iberoamericana de Protección de Datos y de los organismos internacionales especializados: la protección de los datos personales de personas o grupos en situación de vulnerabilidad , que ocupa un lugar propio en la tradición iberoamericana de protección de datos; el tratamiento de datos relativos a la actividad neuronal y la generación de inferencias sobre estados cognitivos, emocionales o conductuales, por su aptitud para revelar aspectos especialmente reservados de la persona y por su incidencia sobre la autonomía individual; y el tratamiento de datos biométricos y el reconocimiento o la inferencia de emociones. El Estándar no sustituye ese acervo ni el régimen que los ordenamientos nacionales dispensan a estas materias: recoge únicamente las especialidades que su tratamiento presenta, desde la perspectiva de la protección de datos personales, cuando se realiza mediante sistemas de inteligencia artificial. Sus criterios sobre estas materias se interpretarán en coherencia con dichos trabajos y con la evolución del conocimiento científico disponible sobre la fiabilidad de tales sistemas.

Capítulo I. Disposiciones generales, ámbito y criterios generales de aplicación

1.1. Objeto

1.1.1. El Estándar establece criterios orientadores para la protección de datos personales en los tratamientos que incorporen sistemas de inteligencia artificial o sean necesarios para su desarrollo, con el fin de reforzar la protección de las personas y favorecer una aplicación armonizada en la región iberoamericana.

1.1.2. El Estándar tiene naturaleza orientadora y carácter complementario del régimen general de protección de datos. Sus disposiciones se aplican sin perjuicio de lo establecido en los ordenamientos aplicables y expresan el modo en que, conforme al Estándar, deben aplicarse las obligaciones y garantías de protección de datos cuando el tratamiento incorpora sistemas de inteligencia artificial. Las referencias que el Estándar hace al régimen general de protección de datos tienen carácter aclaratorio: no condicionan ni delimitan la aplicación de dicho régimen, que rige con independencia del Estándar. A los efectos del Estándar, se entenderá que las expresiones «normativa de protección de datos», «normativa aplicable» y «ordenamiento aplicable» designan ese mismo régimen general; las referencias a la normativa específica designan las disposiciones sectoriales o de alcance particular que resulten adicionalmente aplicables conforme a cada ordenamiento.

1.1.3. Las referencias del Estándar al régimen general de Estándar al régimen general de protección de datos se entenderán hechas al ordenamiento aplicable en cada caso y, como referencia común de la región iberoamericana, a los Estándares de Protección de Datos de los Estados Iberoamericanos.

1.2. Ámbito y condiciones de aplicación

1.2.1. El Estándar se aplicará a los tratamientos de datos personales que incorporen sistemas de inteligencia artificial o sean necesarios para su desarrollo, comprendidos los datos personales inferidos o derivados que estos generen, con independencia de que tales sistemas intervengan en la totalidad o en parte de las operaciones que integran el tratamiento. La aplicación del Estándar deberá ser proporcional a los riesgos derivados del tratamiento, de modo que no comportará obligaciones innecesarias cuando dichos riesgos sean limitados ni atenuará las garantías exigibles cuando el tratamiento comporte riesgos elevados para los derechos y libertades de las personas.

1.2.2. Los ámbitos objetivo, subjetivo y territorial se regirán por el régimen general de protección de datos, sin que el Estándar altere la determinación de la ley aplicable que corresponda conforme a cada ordenamiento.

1.2.3. El Estándar no resultará aplicable cuando no se traten datos personales, en particular en los supuestos de datos anonimizados en los términos exigidos por el Estándar, datos sintéticos que no permitan identificar ni reidentificar a las personas o modelos de inteligencia artificial que no traten ni permitan obtener datos personales.

1.2.4. Quien invoque la inaplicación del Estándar deberá poder acreditarla, cualquiera que sea el supuesto invocado, y disponer, a tal efecto, del análisis en que la funde. Si posteriormente existe un riesgo razonable de identificación o reidentificación, en los términos previstos en el Estándar, el tratamiento quedará sujeto a sus disposiciones.

1.2.5. Cuando de un modelo, comprendidos sus parámetros y los resultados que produce, puedan extraerse, revelarse o atribuirse razonablemente datos personales relativos a una persona identificada o identificable, el tratamiento quedará sujeto al Estándar.

1.3. Tratamientos y ciclo de vida

1.3.1. La aplicación del Estándar alcanzará a los tratamientos de datos personales y a las finalidades para las que se lleven a cabo, y no a los sistemas o modelos de inteligencia artificial, ni a las tecnologías en que se sustenten, considerados de forma abstracta.

1.3.2. El análisis de los tratamientos que incorporen sistemas de inteligencia artificial o sean necesarios para su desarrollo, cualquiera que sea la función que el sistema desempeñe en ellos, deberá tener en cuenta las distintas fases del ciclo de vida del tratamiento, a fin de identificar los cambios que puedan producirse en las finalidades, los datos tratados, los actores intervinientes o los riesgos para las personas y determinar, en cada caso, las obligaciones y garantías aplicables.

1.3.3. A los efectos previstos en el numeral anterior, las referencias del Estándar al desarrollo, el entrenamiento, la evaluación o la mejora designan fases diferenciadas del ciclo de vida del tratamiento. Las obligaciones y garantías aplicables, en particular las relativas a la licitud, la minimización y la calidad de los datos, deberán determinarse de forma separada para cada una de ellas.

1.4. Documentación del tratamiento

1.4.1. Para acreditar el cumplimiento de la normativa de protección de datos y del Estándar, la documentación del tratamiento de datos personales deberá ser suficiente, teniendo en cuenta la naturaleza, la complejidad y los riesgos del tratamiento.

1.4.2. Cuando resulte relevante para acreditar dicho cumplimiento, la documentación podrá comprender las decisiones adoptadas sobre el diseño, la configuración o la utilización del sistema de inteligencia artificial.

1.4.3. Cuando intervengan encargados del tratamiento o proveedores de sistemas o modelos de inteligencia artificial, el responsable deberá asegurarse, mediante las relaciones jurídicas que correspondan, de disponer de la información y de la documentación necesarias sobre el sistema y, cuando resulte relevante, sobre los tratamientos previos realizados por el proveedor, para cumplir sus obligaciones, incluidas las de transparencia y explicabilidad, conforme a la normativa

de protección de datos y al Estándar. Esa información comprenderá, en lo necesario, la relativa a la procedencia, la representatividad y las condiciones de obtención de los datos utilizados, a las limitaciones conocidas del sistema, a los sesgos identificados y a las medidas de mitigación adoptadas. Cuando existan tratamientos de datos personales, proporcionados por el responsable o generados o inferidos durante la prestación del servicio, que excedan del objeto del encargo, deberá documentarse al menos el cumplimiento de la normativa de protección de datos respecto de ellos. Las condiciones de utilización de los datos por los proveedores durante la relación se rigen por lo dispuesto en el Estándar sobre actores, cadena de valor y transferencias internacionales.

1.5. Finalidades y reutilización de datos

1.5.1. Deberán identificarse de forma diferenciada, cuando impliquen tratamientos distintos o una reutilización de los datos para fines diferentes, las finalidades de operación del tratamiento de datos personales, las de desarrollo, entrenamiento, evaluación o mejora y las finalidades instrumentales, como las de seguridad o prevención del fraude.

1.5.2. Cuando, en el marco de un tratamiento que incorpore un sistema de inteligencia artificial, los datos personales se utilicen para una finalidad distinta de aquella para la que fueron obtenidos, deberá verificarse que dicha utilización cuenta con una base jurídica adecuada y, cuando proceda, realizarse el correspondiente análisis de compatibilidad conforme al régimen general de protección de datos.

1.5.3. El entrenamiento, ajuste o reentrenamiento de modelos constituirán un tratamiento diferenciado cuando respondan a una finalidad propia o comporten cambios relevantes en la finalidad, los datos utilizados, los actores intervinientes o los riesgos del tratamiento.

1.6. Definiciones

1.6.1. A los efectos del Estándar en su conjunto, y sin perjuicio del régimen general de protección de datos, del que se toman las nociones no definidas en este apartado —entre ellas las de responsable, encargado y corresponsable del tratamiento, cuya determinación en los tratamientos que incorporen sistemas de inteligencia artificial se rige por lo dispuesto en el Estándar sobre actores, cadena de valor y transferencias internacionales—, se entenderá por:

a) Sistema de inteligencia artificial: conjunto integrado de componentes que, para objetivos explícitos o implícitos y a partir de datos de entrada, genera predicciones, inferencias, recomendaciones, contenidos o decisiones mediante aprendizaje automático, procesamiento estadístico o reglas basadas en conocimiento, con distintos grados de autonomía.

b) Modelo: artefacto resultante del entrenamiento o ajuste que incorpora en sus parámetros patrones derivados de los datos utilizados y que, aplicado a nuevos datos, produce inferencias o predicciones. El componente característico que define a un sistema de inteligencia artificial es la inclusión de uno o más modelos.

c) Inferencia: resultado que el sistema deriva, predice o estima a partir de datos de entrada, incluidos atributos, comportamientos, preferencias o categorías no suministrados por la persona titular.

d) Entrenamiento y ajuste: operaciones que construyen o modifican un modelo para que cumpla propósitos predefinidos con niveles adecuados de calidad; comprenden el entrenamiento inicial y las operaciones posteriores de ajuste o especialización.

e) Desarrollo: actividades preparatorias del entrenamiento, comprendidos el diseño del sistema, la selección de la arquitectura, la preparación y exploración de los conjuntos de datos y las pruebas experimentales, que no constituyen por sí mismas entrenamiento o ajuste

f) Dato inferido y dato derivado: datos personales producidos por el sistema. El dato inferido resulta de una predicción, estimación o hipótesis; el dato derivado, de la transformación o combinación de datos preexistentes.

g) Datos sintéticos: datos generados artificialmente para reproducir las propiedades estadísticas de un conjunto de referencia, sin corresponder a personas concretas. Su carácter sintético no implica por sí mismo la anonimización de los datos de referencia, si los hubiera.

h) Proveedor de inteligencia artificial: quien desarrolla, entrena, suministra u opera modelos o servicios de inteligencia artificial para su incorporación por otros en tratamientos de datos personales.

i) Neurodatos: datos personales obtenidos a partir de un tratamiento técnico específico (de manera directa, indirecta o inferida), en relación con las características anatómicas, fisiológicas, funcionales o conductuales de una persona física relativos a la actividad cerebral y de su sistema nervioso

j) Personas o grupos en situación de vulnerabilidad: aquellos que, de forma permanente o circunstancial y por razón de su edad, discapacidad o de otras circunstancias personales, sociales, económicas o contextuales, incluidas las propias del entorno digital, presentan un mayor riesgo de sufrir daños o mayores dificultades para comprender, controlar o ejercer sus derechos en relación con el tratamiento de datos personales o el uso de sistemas de inteligencia artificial.

Capítulo II. Principios y obligaciones generales de protección de datos en los tratamientos con inteligencia artificial

2.1. Licitud y bases jurídicas

2.1.1. La licitud del tratamiento y las bases jurídicas que puedan fundamentarla se rigen por el régimen general de protección de datos, que determina las bases jurídicas admisibles y las condiciones de su utilización; este apartado no las enumera ni las jerarquiza, y se limita a precisar su aplicación cuando el tratamiento incorpora sistemas de inteligencia artificial.

La necesidad de una nueva base jurídica o de un análisis de compatibilidad dependerá de que cambien la finalidad del tratamiento o las condiciones que justificaron inicialmente su licitud. En consecuencia, las distintas fases del ciclo de vida de un sistema de inteligencia artificial no constituyen, por sí mismas, tratamientos independientes ni exigen necesariamente una base jurídica distinta.

2.1.2. Cuando el ordenamiento aplicable lo permita y en los términos en que lo haga, el interés legítimo podrá constituir una base jurídica adecuada para determinados tratamientos que respondan a las finalidades de desarrollo, entrenamiento, evaluación o mejora de sistemas de inteligencia artificial, siempre que se realice la correspondiente ponderación y se adopten las medidas técnicas y organizativas necesarias para proteger los derechos y libertades de las personas. La idoneidad del interés legítimo como base jurídica no podrá presumirse. En dicha ponderación podrán resultar especialmente relevantes, entre otros factores, la naturaleza de los datos tratados y el contexto del tratamiento, la aplicación de medidas de minimización o seudonimización, las expectativas razonables de las personas respecto de esa utilización, el impacto sobre ellas y la gravedad de los riesgos para sus derechos y libertades, la inexistencia de decisiones con efectos jurídicos o que afecten significativamente de modo similar a las personas durante esas fases y la posibilidad de ejercer efectivamente el derecho de oposición. La concurrencia de esta base jurídica no dispensa de los principios de limitación de la finalidad y de minimización ni de las técnicas de reducción del riesgo de identificación previstas en el Estándar, aplicables a cada una de las finalidades enunciadas.

La ponderación deberá ser especialmente rigurosa cuando el tratamiento implique la reutilización masiva de datos personales, categorías especiales de datos, datos de personas o grupos en situación de vulnerabilidad o actividades de perfilado que puedan producir efectos jurídicos o que afecten significativamente de modo similar a las personas. Dicha ponderación no dispensa del régimen específicamente aplicable a las categorías especiales de datos, cuyas condiciones de licitud y garantías reforzadas deben concurrir de manera adicional; tampoco sustituye a la evaluación de impacto ni queda sustituida por ella, al tratarse de exigencias distintas.

2.1.3. Tratándose de tratamientos realizados por entidades del sector público en el ejercicio de funciones públicas o de potestades de autoridad, la admisibilidad del interés legítimo y la determinación de las bases jurídicas que puedan fundamentar el tratamiento corresponden a las reglas de habilitación establecidas por el ordenamiento aplicable

2.2. Lealtad, finalidad y compatibilidad

2.2.1. El principio de lealtad adquiere especial relevancia en los tratamientos de datos personales destinados al desarrollo, entrenamiento, evaluación o mejora de sistemas de inteligencia artificial, así como en aquellos que generen perfiles, clasificaciones o priorizaciones de personas o adopten decisiones sobre ellas a partir de tales operaciones, especialmente cuando impliquen reutilizar datos personales para finalidades distintas de aquellas para las que fueron obtenidos. En estos supuestos, la apreciación de la lealtad atenderá a la correspondencia entre el tratamiento efectivamente realizado y aquello que las personas pueden razonablemente esperar cuando facilitan sus datos personales.

2.2.2. La utilización, por proveedores o encargados, de datos personales tratados por cuenta del responsable para finalidades propias, como el desarrollo, entrenamiento, evaluación o mejora de modelos o servicios de inteligencia artificial, se considerará, con carácter general, una finalidad distinta y no se entenderá comprendida en el tratamiento encomendado, salvo que exista una habilitación expresa y una base jurídica que la ampare. A estos efectos, la anonimización o transformación de dichos datos realizada para esas finalidades propias constituye una utilización sujeta a lo dispuesto en este numeral, sin que el resultado anonimizado pueda entenderse amparado por el encargo cuando la operación que lo produjo no lo estuviera.

2.2.3. El carácter accesible al público de los datos personales o dato público no legitima por sí solo su utilización para el desarrollo, entrenamiento o mejora de modelos o sistemas de inteligencia artificial.

2.2.4. Son incompatibles con los principios de protección de datos los tratamientos mediante sistemas de inteligencia artificial destinados a manipular de forma engañosa la conducta de las personas, explotar indebidamente situaciones de vulnerabilidad —entre otras, las derivadas de la edad, la discapacidad o la situación socioeconómica— o establecer mecanismos de puntuación social que produzcan efectos discriminatorios o restricciones injustificadas de derechos. Lo dispuesto en este numeral se entiende sin perjuicio de las garantías reforzadas que el Estándar establece para los tratamientos que incorporen técnicas biométricas y para la obtención de datos personales accesibles al público, aplicables a la constitución o el enriquecimiento de bases de datos biométricas mediante recopilaciones masivas, sistemáticas e indiscriminadas de imágenes o grabaciones, así como de las prohibiciones, limitaciones o requisitos adicionales establecidos por la normativa aplicable.

2.3. Minimización y proporcionalidad

2.3.1. En los tratamientos con inteligencia artificial, el principio de minimización se proyecta también sobre la capacidad del sistema para identificar, singularizar o generar inferencias sobre las personas, evitando dichos resultados cuando no sean necesarios para la finalidad del tratamiento.

2.3.2. La minimización deberá tener especialmente en cuenta la selección de los datos y la generación de inferencias durante el tratamiento.

2.3.3. En los tratamientos que utilicen sistemas de inteligencia artificial para identificar o verificar la identidad de las personas deberá distinguirse claramente entre la identificación y la verificación de la identidad, atendiendo a sus diferentes riesgos y garantías. Esta distinción no prejuzga la licitud de tales tratamientos, que se determinará conforme al régimen general de protección de datos y a la normativa específica aplicable, ni excluye la calificación de los datos utilizados como categorías especiales cuando así proceda conforme a dicho régimen, con las garantías reforzadas que de ello se sigan.

2.4. Calidad y exactitud

2.4.1. La calidad y la exactitud de los datos tratados mediante sistemas de inteligencia artificial, así como de las inferencias y los resultados obtenidos de dicho tratamiento, deberán valorarse en función de su idoneidad para la finalidad del tratamiento de datos personales y del contexto de uso. La exactitud exige el grado de adecuación a la realidad necesario para cumplir esa finalidad y proteger los derechos de las personas, sin que resulte exigible en todo caso una correspondencia absoluta. Cuando hayan de fundamentar decisiones que produzcan efectos jurídicos sobre una persona o la afecten significativamente de modo similar, la exigencia de exactitud y de actualización se apreciará con especial rigor, sin perjuicio de los derechos de rectificación, actualización y supresión reconocidos por el régimen general de protección de datos y de su proyección sobre las inferencias, los resultados y los modelos previamente entrenados conforme a lo previsto en el Estándar en materia de derechos de las personas.

2.4.2. Cuando las inferencias o los resultados puedan producir efectos sobre una persona, deberán utilizarse teniendo en cuenta su carácter de hipótesis, deducción o propuesta y el grado de incertidumbre asociado, evitando tratarlos como si fueran hechos ciertos cuando no exista fundamento suficiente para ello y absteniéndose de presentarlos como tales ante la persona o ante terceros. Conforme a la regla de aplicación proporcional a los riesgos establecida en el Estándar en materia de ámbito y condiciones de aplicación, el rigor exigible en esa valoración se graduará atendiendo a la relevancia de los efectos que las inferencias o los resultados puedan producir sobre la persona. Ese carácter de la inferencia o del resultado no altera las garantías aplicables a los datos personales utilizados para generarlos.

2.4.3. La calidad y la exactitud exigibles podrán variar según la finalidad y la fase del tratamiento. En las actividades de desarrollo, entrenamiento, validación o prueba de sistemas de inteligencia artificial, siempre que resulte adecuado para la finalidad de esas actividades y no se produzcan efectos indebidos sobre las personas o sobre sus derechos, podrán utilizarse datos que no reúnan las mismas condiciones exigibles en la fase de utilización del sistema. Esta modulación se refiere a las condiciones de calidad y exactitud exigibles y no presupone ni habilita por sí sola la utilización de datos personales en dichas actividades, a las que resultan igualmente aplicables los principios de licitud, limitación de la finalidad y minimización y las técnicas de reducción del riesgo de identificación previstas en el Estándar.

2.5. Transparencia y explicabilidad

2.5.1. Cuando un tratamiento de datos personales incorpore un sistema de inteligencia artificial, el responsable deberá informar a las personas, en términos comprensibles conforme a este apartado, de la utilización del sistema, de la finalidad para la que se emplea, del papel que desempeña en el tratamiento y de las consecuencias que puede producir para ellas. La mera indicación de que se emplea inteligencia artificial o algoritmos no satisface el deber de informar.

2.5.2. La información permitirá comprender las categorías y los orígenes de los datos; si el sistema genera perfiles, inferencias, clasificaciones, recomendaciones o decisiones; los destinatarios y, cuando resulte relevante, la intervención de proveedores o terceros; los criterios de conservación y, cuando estén determinados, los plazos aplicables; el grado de intervención humana; las consecuencias razonablemente previsibles; y las vías de ejercicio de derechos y de impugnación. La descripción de la intervención humana expresará su función efectiva y no presentará como garantía una supervisión meramente formal.

2.5.3. Cuando sea necesario para comprender el tratamiento o para ejercer los derechos, se facilitará a la persona una explicación significativa, en lenguaje comprensible. La explicación comprenderá la lógica del sistema, los factores que influyen de forma relevante en el resultado, las fuentes de datos pertinentes, las limitaciones conocidas y las medidas de detección y corrección de errores y sesgos. No es exigible revelar el código fuente, el modelo completo ni información cuya divulgación comprometa la seguridad, los secretos empresariales o la propiedad intelectual, siempre que la persona conserve una comprensión suficiente del tratamiento. Esta limitación opera únicamente respecto del contenido de la explicación exigible y no dispensa de facilitar la información adicional que resulte necesaria para preservar dicha comprensión suficiente, sin perjuicio de lo dispuesto en este apartado sobre el contenido esencial del derecho de información. Tampoco afecta a las facultades de supervisión que el ordenamiento aplicable atribuya a las autoridades de protección de datos ni a los requisitos adicionales establecidos por la normativa específica aplicable en materia de transparencia y de acceso a la información pública. La explicación podrá apoyarse, además, en los resultados de auditorías, certificaciones o esquemas de sellos que ofrezcan garantías sobre el funcionamiento del sistema, sin que ello sustituya la explicación exigible ni reduzca su contenido.

2.5.4. La extensión y el detalle de la información y de la explicación se graduarán por el impacto del tratamiento de datos personales, el grado de automatización y la complejidad del sistema. Cuando se trate de una decisión individual automatizada, la información y la explicación facilitarán a la persona el conocimiento de las garantías y de las vías disponibles para solicitar la intervención humana, impugnar la decisión o ejercer los derechos que le correspondan.

2.5.5. La información podrá ofrecerse por capas o niveles de detalle, siempre que permita a las personas comprender el tratamiento y ejercer efectivamente sus derechos. La protección de la seguridad, los secretos empresariales o la propiedad intelectual no podrá invocarse para vaciar el contenido esencial del derecho de información ni para impedir el ejercicio efectivo de los derechos que el régimen general de protección de datos reconoce a las personas.

Capítulo III. Categorías especiales, inferencias sensibles y grupos de protección reforzada

3.1. Sensibilidad inferida

3.1.1. La salud, la ideología, la orientación sexual o las creencias pueden inferirse incluso a partir de variables aparentemente neutras. Pueden generarse asimismo otras inferencias especialmente sensibles, entre ellas las relativas a la vulnerabilidad económica. Estas inferencias, cuando no constituyan por sí mismas categorías especiales de datos, deberán valorarse en atención a sus efectos sobre las personas, conforme a la regla de aplicación proporcional a los riesgos y a la evaluación de impacto previstas en el Estándar, en los términos y con las condiciones allí establecidos. El régimen reforzado que el régimen general de protección de datos prevé para las categorías especiales de datos se determinará por la naturaleza efectiva de la información inferida y no sólo por la de los datos de entrada.

3.1.2. El régimen reforzado que el régimen general de protección de datos prevé para las categorías especiales de datos será aplicable cuando el sistema genere o utilice inferencias que permitan atribuir razonablemente a una persona una categoría especial de datos. La finalidad, el contexto y el impacto del tratamiento se tendrán en cuenta para valorar el carácter razonable de dicha atribución. El grado de fiabilidad o de exactitud de la inferencia no excluirá por sí solo la aplicación del régimen reforzado cuando dicha inferencia se utilice para adoptar o apoyar decisiones sobre la persona.

3.1.3. La presencia incidental de categorías especiales de datos en el desarrollo o el entrenamiento de un sistema deberá valorarse de forma diferenciada para cada una de esas fases, conforme al enfoque de ciclo de vida previsto con carácter general en el Estándar, atendiendo a si dichos datos son utilizados, pueden influir en el modelo o pueden revelarse mediante sus inferencias o resultados. El responsable adoptará medidas razonables para limitar su utilización y evitar que sirvan de fundamento para inferencias o decisiones sobre las personas cuando ello no resulte necesario y lícito.

3.1.4. El tratamiento de categorías especiales de datos o de inferencias que permitan atribuir razonablemente tales categorías, incluidas su conservación y su utilización, sólo podrá realizarse cuando concurra alguna de las excepciones que el régimen general de protección de datos prevea para dichas categorías. El Estándar no constituye por sí mismo habilitación para dicho tratamiento. Cuando el régimen general de protección de datos lo permita con la finalidad de detectar, evaluar o corregir efectos discriminatorios o sesgos, el tratamiento quedará sujeto a los principios de minimización y proporcionalidad y a las garantías reforzadas exigibles conforme al Estándar y a dicho régimen.

3.2. Personas o grupos en situación de vulnerabilidad

3.2.1. En los tratamientos con inteligencia artificial que afecten a personas o grupos en situación de vulnerabilidad, su interés superior deberá proyectarse tanto sobre la determinación del

tratamiento como sobre la valoración y la utilización de las inferencias y los resultados generados por el sistema.

3.2.2. La utilización de datos personales de personas o grupos en situación de vulnerabilidad para el desarrollo, entrenamiento, validación o mejora de sistemas de inteligencia artificial requerirá, en cada una de esas fases, una justificación especialmente rigurosa y la adopción de garantías reforzadas, atendiendo a su interés superior. El contenido de estas garantías se determinará en atención a los riesgos del tratamiento para este grupo y a las exigencias que el régimen general de protección de datos establezca para sus datos personales. Cuando dicho tratamiento no resulte necesario para la finalidad perseguida, deberán adoptarse medidas para evitar la utilización de estos datos.

3.2.3. En los sistemas dirigidos a personas o grupos en situación de vulnerabilidad, o que puedan afectarles de forma significativa, incluidos los supuestos de efectos discriminatorios, deberán aplicarse configuraciones especialmente protectoras y predeterminadas, un diseño adecuado a su edad y garantías reforzadas frente al perfilado, la publicidad comportamental y la obtención indebida de sus datos personales, conforme a la protección de datos desde el diseño y por defecto prevista en el Estándar. Estas medidas no habilitan por sí mismas el tratamiento, cuya licitud se rige por el régimen general de protección de datos y por las prohibiciones y limitaciones que este establezca. El perfilado, la publicidad comportamental y las inferencias que exploten indebidamente la situación de vulnerabilidad de estas personas o grupos quedan comprendidos en la incompatibilidad con los principios de protección de datos declarada con carácter general en el Estándar. Cuando, pese a las medidas adoptadas, persista un riesgo elevado para este grupo, se estará a lo previsto con carácter general en el Estándar para los supuestos en que el riesgo elevado persista. La información y la explicación exigibles conforme al Estándar se facilitarán en términos comprensibles atendiendo a su edad.

3.3. Neurodatos

3.3.1. El responsable someterá los tratamientos mediante sistemas de inteligencia artificial que utilicen neurodatos o generen inferencias sobre estados cognitivos, emocionales o conductuales a una especial evaluación de necesidad y proporcionalidad, atendiendo a su capacidad para revelar aspectos especialmente sensibles de la persona, al contexto en que se realizan, en particular cuando medien relaciones caracterizadas por una asimetría de poder, y a los riesgos que pueden plantear para la autonomía, la intimidad y los derechos fundamentales. Esta evaluación se entiende sin perjuicio de la evaluación de impacto que resulte exigible conforme a la normativa aplicable y al Estándar.

3.3.2. Son incompatibles con los principios de protección de datos, en los términos declarados con carácter general por el Estándar, los tratamientos que utilicen neurodatos o inferencias sobre estados cognitivos, emocionales o conductuales para manipular indebidamente la voluntad o el comportamiento de las personas o explotar situaciones de vulnerabilidad. Los tratamientos dirigidos a inferir convicciones íntimas o a realizar formas especialmente intrusivas de vigilancia sólo podrán admitirse de forma excepcional cuando concurra una base jurídica admisible conforme al régimen general de protección de datos respondan a una finalidad legítima especialmente relevante, cuya especial relevancia deberá justificarse, resulten necesarios y

proporcionados y no existan alternativas menos intrusivas, sin perjuicio de las prohibiciones, limitaciones o requisitos adicionales que establezca la normativa específica aplicable. La concurrencia de todas estas condiciones deberá poder acreditarse conforme al deber de documentación del tratamiento previsto en el Estándar. El consentimiento no debería constituir una base legitimadora cuando el responsable del tratamiento no esté en condiciones de determinar de antemano, con razonable precisión, qué inferencias se derivarán del tratamiento de los neurodatos.

Capítulo IV. Decisiones automatizadas y perfilado

4.1. Identificación de decisiones individuales automatizadas

4.1.1. El responsable identificará los supuestos en que existe una decisión basada únicamente en el tratamiento automatizado, incluido el perfilado, que produzca efectos jurídicos sobre la persona o la afecte significativamente, cualquiera que sea la naturaleza —social, económica o de otro orden— de esa afectación, distinguiendo las decisiones exclusivamente automatizadas, los sistemas de apoyo a la decisión y los supuestos de intervención humana efectiva conforme al criterio establecido en este capítulo. La identificación de estos supuestos determinará la aplicación de los deberes de transparencia y de explicación y de las garantías que el Estándar prevé frente a las decisiones individuales automatizadas.

4.1.2. Esta identificación se revisará cuando se produzcan cambios relevantes en el tratamiento o en el sistema de inteligencia artificial, a fin de determinar y adaptar las garantías aplicables teniendo en cuenta, entre otros aspectos, el carácter significativo de los efectos del tratamiento, la fase del ciclo de vida del tratamiento y las características del sistema.

4.2. Intervención humana efectiva

4.2.1. Sólo se excluirá el carácter exclusivamente automatizado de una decisión cuando exista una intervención humana previa, real y efectiva, realizada por una persona con capacidad para valorar y comprender el resultado del sistema y con autoridad suficiente para confirmarlo, modificarlo o apartarse de él. No constituirá intervención humana efectiva la que se limite a formalizar o refrendar el resultado del sistema sin reunir esas condiciones; en tal caso, la decisión conservará su carácter exclusivamente automatizado y le resultarán aplicables las garantías frente a las decisiones individuales previstas en este capítulo, así como los deberes de información y de explicación previstos en el Estándar en materia de transparencia y explicabilidad. Las condiciones organizativas necesarias para que esa intervención resulte efectiva, entre ellas la asignación de responsabilidades, la formación y la supervisión de quienes la realizan, se adoptarán conforme a las medidas de gobernanza algorítmica previstas en el Estándar.

4.2.2. La supervisión humana podrá apoyarse en herramientas automatizadas, como sistemas de alerta o detección de discrepancias, aunque estas no sustituirán la intervención humana cuando esta resulte exigible ni podrán presentarse como equivalente de ella. Su utilización no alterará las condiciones exigidas en este capítulo para que la intervención humana excluya el carácter exclusivamente automatizado de la decisión.

4.3. Efectividad de las garantías y no discriminación algorítmica

4.3.1. Las garantías aplicables frente a las decisiones individuales comprenden, al menos y en los términos previstos en el Estándar: la información sobre la utilización del sistema de inteligencia artificial, la finalidad para la que se emplea, el papel que desempeña en el tratamiento y las consecuencias que puede producir, así como la explicación significativa del tratamiento, conforme a lo dispuesto en el Estándar en materia de transparencia y explicabilidad; el conocimiento de las vías disponibles para solicitar la intervención humana, impugnar la decisión y ejercer los derechos que correspondan; y las medidas de prevención, detección y corrección de los efectos discriminatorios previstas en este capítulo. Dichas garantías serán accesibles, oportunas y reales en sus efectos, y no meramente formales, permitiendo a las personas comprender la decisión y sus consecuencias, solicitar la intervención humana cuando proceda y ejercer los derechos que les reconoce la normativa de protección de datos.

4.3.2. El responsable evaluará si el sistema puede producir efectos discriminatorios, incluidos los derivados del uso de variables aparentemente neutras, de diferencias relevantes en su funcionamiento respecto de distintos grupos de personas o de las inferencias generadas por el sistema, y adoptará las medidas necesarias para prevenir, detectar y corregir los sesgos que puedan afectar injustificadamente a las personas. La idoneidad y la suficiencia de estas medidas deberán poder acreditarse conforme al deber de documentación del tratamiento previsto en el Estándar, sin perjuicio de las competencias que correspondan a las autoridades de protección de datos conforme al ordenamiento aplicable.

Capítulo V. Derechos de las personas en sistemas de inteligencia artificial

5.1. Los derechos reconocidos por la normativa de protección de datos podrán ejercerse respecto de los datos personales tratados mediante sistemas de inteligencia artificial, incluidas las inferencias, perfiles y demás resultados que constituyan datos personales. La naturaleza de estos datos podrá determinar el alcance o la forma de ejercicio de cada derecho, sin perjuicio de lo establecido en dicha normativa.

5.2. En el ejercicio del derecho de acceso, la persona podrá conocer, cuando proceda conforme a la normativa de protección de datos, si sus datos han sido utilizados en un sistema de inteligencia artificial, si se han generado perfiles o inferencias sobre ella y si estos han influido o han sido utilizados para adoptar o apoyar decisiones que le afecten.

5.3. Cuando el derecho de rectificación se refiera a inferencias, perfiles o predicciones generadas por un sistema de inteligencia artificial, su ejercicio deberá adaptarse a su naturaleza. Cuando estas no puedan rectificarse directamente por su propia naturaleza, la persona podrá solicitar la corrección de los datos utilizados para generarlas, aportar información complementaria, impugnarlas o solicitar su revisión, especialmente cuando puedan producir efectos sobre ella.

5.4. El ejercicio de los derechos de supresión, limitación y oposición se regirá por el régimen general de protección de datos. El responsable procurará seleccionar, configurar o utilizar sistemas que incorporen, siempre que resulte técnicamente viable atendiendo al estado de la tecnología y a los medios razonablemente disponibles, mecanismos que faciliten el ejercicio efectivo de estos derechos durante todo el ciclo de vida del tratamiento, en particular mediante la eliminación, desvinculación o exclusión de los datos personales de futuros entrenamientos. La viabilidad técnica de dichos mecanismos no condiciona la exigibilidad de estos derechos y su apreciación deberá poder acreditarse conforme al deber de documentación del tratamiento previsto en el Estándar, sin perjuicio de las competencias que correspondan a las autoridades de protección de datos conforme al ordenamiento aplicable.

5.5. Cuando, debido al estado de la tecnología, los derechos de supresión, limitación u oposición no puedan ejercerse plenamente respecto de un modelo previamente entrenado, el responsable deberá justificar esta circunstancia y dejar constancia de ella conforme al deber de documentación del tratamiento previsto en el Estándar, informar de ello a la persona afectada y adoptar las medidas razonablemente disponibles para limitar el tratamiento y reducir su impacto, incluyendo, cuando sea posible, la supresión de los datos personales de los entornos en que se encuentren —entre otros, los conjuntos de datos de entrenamiento, los índices, las bases de representaciones vectoriales y las memorias persistentes— y su exclusión de futuros entrenamientos. Esta circunstancia no tendrá carácter definitivo y se apreciará conforme al estado de la tecnología existente en cada momento. Lo dispuesto en este apartado se entiende sin perjuicio del deber del responsable, al utilizar modelos o servicios de inteligencia artificial desarrollados por terceros, de verificar las condiciones de su utilización conforme a lo previsto en el Estándar en materia de actores y cadena de valor, y de la responsabilidad que a cada interviniente corresponda conforme a dicho régimen.

5.6. La portabilidad se proyecta sobre los datos aportados por la persona titular y, en su caso, sobre los observados que se hallen estructurados. Conforme al régimen general de protección de datos, la portabilidad no comprenderá, por regla general, los datos inferidos o derivados, sin que ello afecte a su condición de datos personales ni al ejercicio sobre ellos de los demás derechos reconocidos a la persona titular. Podrán facilitarse formatos estructurados que favorezcan la interoperabilidad y el cambio de proveedor.

Capítulo VI. Actores, cadena de valor y transferencias internacionales

6.1. La condición de responsable o de encargado y, cuando el ordenamiento aplicable reconozca esta figura, la de corresponsable, se determinarán respecto de cada tratamiento de datos personales; es responsable quien decide los fines y los medios esenciales del tratamiento concreto, aunque no entrene ni controle técnicamente el modelo. Cuando varios actores compartan dichas decisiones, las responsabilidades asumidas por cada uno deberán determinarse desde el diseño del tratamiento, definirse conforme a lo previsto en el Estándar sobre las relaciones entre los actores que intervienen en el tratamiento y documentarse conforme al deber de documentación del tratamiento previsto en el Estándar.

6.2. Los proveedores de modelos o de servicios de inteligencia artificial podrán actuar como responsables, encargados o corresponsables, según las funciones que desempeñen en cada tratamiento, apreciadas conforme al ciclo de vida del tratamiento y atendiendo a las decisiones adoptadas en fases anteriores que lo condicionen de forma relevante. En particular, cuando reutilicen datos personales para finalidades propias o determinen los fines o los medios esenciales del tratamiento, asumirán las responsabilidades que correspondan conforme a la normativa de protección de datos. Las condiciones de utilización de los datos para el desarrollo, entrenamiento o mejora de modelos deberán determinarse expresamente entre las partes, sin que dicha determinación supla la base jurídica que el tratamiento requiera ni lo previsto en el Estándar en materia de lealtad, finalidad y compatibilidad sobre la utilización de datos personales por proveedores o encargados para finalidades propias.

6.3. El responsable identificará los actores que intervienen en el desarrollo, suministro y utilización del sistema de inteligencia artificial, comprendidos los encargados, subencargados y proveedores sucesivos que intervengan en el tratamiento, así como los tratamientos o funciones que desempeña cada uno respecto de los datos personales y las transferencias internacionales que puedan producirse. Esta identificación deberá permitir determinar las responsabilidades de cada interviniente, formalizar las relaciones jurídicas correspondientes y disponer de la información necesaria para el cumplimiento de la normativa de protección de datos. Lo anterior se entiende sin perjuicio de la información que deba facilitarse a las personas sobre los destinatarios y sobre la intervención de proveedores o terceros, que se regirá por lo previsto en el Estándar en materia de transparencia y explicabilidad.

6.4. La utilización de modelos o servicios de inteligencia artificial desarrollados por terceros no exime al responsable de verificar que su utilización resulta compatible con la normativa de protección de datos ni de evaluar las garantías ofrecidas por dichos proveedores, especialmente cuando las características del modelo o las condiciones de prestación del servicio puedan afectar al tratamiento de datos personales. La selección de proveedores deberá atender, entre otros aspectos, a la naturaleza de los datos tratados y a las finalidades del tratamiento, a las garantías que ofrezcan para el cumplimiento de la normativa de protección de datos y a la información que pongan a disposición del responsable para acreditar dicho cumplimiento, comprendida la relativa a los actores que intervengan en la prestación del servicio, a las limitaciones que puedan afectar al tratamiento y a los mecanismos previstos para comunicar al responsable los cambios relevantes en la prestación. Lo dispuesto en este apartado se entiende sin perjuicio de la información y la documentación de que el responsable deba disponer sobre el sistema y sobre

los tratamientos previos realizados por el proveedor conforme al deber de documentación del tratamiento previsto en el Estándar.

6.5. Las relaciones entre los distintos actores que intervengan en el tratamiento, comprendidos los subencargados y los proveedores sucesivos, deberán definir claramente las responsabilidades asumidas por cada uno, en particular respecto del uso de los datos personales, incluidas la aplicación sobre ellos de técnicas de anonimización, su utilización para el desarrollo, entrenamiento o mejora de modelos conforme a lo previsto en el Estándar en materia de lealtad, finalidad y compatibilidad, el ejercicio de los derechos de las personas, las transferencias internacionales y aquellas obligaciones cuyo cumplimiento dependa de la actuación o de las capacidades o los medios de que disponga el proveedor.

6.6. Las transferencias internacionales de datos personales realizadas en el contexto del desarrollo, entrenamiento, despliegue o utilización de sistemas de inteligencia artificial quedarán sujetas al régimen general de protección de datos. Cuando el contexto del tratamiento o los riesgos asociados al uso de sistemas de inteligencia artificial lo requieran, deberán adoptarse medidas complementarias adecuadas respecto de los datos objeto de transferencia, tales como su minimización, el cifrado, la seudonimización o garantías contractuales adicionales sobre su uso, conservación o reutilización. La identificación de las transferencias que puedan producirse cuando intervengan proveedores o subencargados se regirá por lo previsto en este capítulo sobre identificación de los actores que intervienen en el tratamiento.

Capítulo VII. Gobernanza algorítmica, medidas proactivas y evaluación de impacto

7.1. La protección de datos desde el diseño y por defecto deberá aplicarse e integrarse en la gobernanza del tratamiento durante todo el ciclo de vida de los tratamientos que incorporen sistemas de inteligencia artificial, teniendo en cuenta los riesgos específicos derivados de su desarrollo, entrenamiento, despliegue, utilización y evolución, y proyectando sobre cada una de esas fases el principio de minimización en los términos previstos en el Estándar.

7.2. La documentación y el registro de las actividades de tratamiento cumplirán, respecto de los tratamientos que incorporen sistemas de inteligencia artificial, la función acreditativa que corresponde al deber de documentación del tratamiento previsto en el Estándar, e identificarán, cuando resulte relevante, las particularidades derivadas del uso de tales sistemas. Entre esas particularidades se comprenderán las categorías y el origen de los datos de entrada, los atributos, perfiles, inferencias o clasificaciones que el sistema genere o pueda razonablemente generar y las categorías de resultados que produzca, así como las finalidades y los usos previstos de tales resultados.

7.3. El responsable establecerá medidas de gobernanza algorítmica adecuadas, proporcionadas al riesgo y susceptibles de verificación, para garantizar el cumplimiento de la normativa de protección de datos durante todo el ciclo de vida del tratamiento. Estas medidas incluirán, cuando proceda, la asignación de responsabilidades, la gestión de riesgos, la supervisión humana, la formación adecuada de las personas que intervienen en el tratamiento y, en especial, en la configuración, la utilización o la supervisión del sistema de inteligencia artificial, la

coordinación entre las funciones relevantes de la organización y la revisión periódica y la mejora continua de las medidas adoptadas. La adecuación de estas medidas se apreciará en atención a la naturaleza, el contexto, el alcance y las finalidades del tratamiento y a los riesgos que pueda entrañar para los derechos y libertades de las personas, sin perjuicio de las competencias que correspondan a las autoridades de protección de datos conforme al ordenamiento aplicable.

7.4. Cuando proceda designar un delegado u oficial de protección de datos conforme al ordenamiento aplicable, este participará desde las fases iniciales de los tratamientos que incorporen sistemas de inteligencia artificial, con acceso a la información necesaria, la formación y los recursos adecuados para desempeñar sus funciones de forma efectiva y no meramente formal.

Cuando la organización disponga de funciones u órganos específicos de gobernanza de la inteligencia artificial, estos coordinarán su actuación con el delegado u oficial de protección de datos en las cuestiones que afecten al tratamiento de datos personales, respetando la autonomía funcional que corresponda a cada uno.

7.5. Cuando resulte exigible conforme a la normativa aplicable, el responsable deberá realizar la evaluación de impacto antes del inicio del tratamiento y de su despliegue, y revisarla cuando se produzcan cambios relevantes en los datos, las finalidades, los modelos y los sistemas de inteligencia artificial utilizados, los proveedores o los riesgos para las personas. Se recomienda realizarla, con independencia de dicha exigibilidad, cuando el tratamiento pueda entrañar un riesgo elevado para los derechos y libertades de las personas. La evaluación prestará especial atención a los riesgos derivados del uso de decisiones automatizadas, del perfilado y de los posibles efectos discriminatorios o adversos sobre las personas o grupos afectados cuando resulten relevantes para el tratamiento; cuando el tratamiento pueda producir tales efectos o afectar a personas o grupos que requieran una protección reforzada, se procurará recabar el punto de vista de las personas o colectivos potencialmente afectados o de sus representantes. El alcance y el nivel de documentación de la evaluación deberán ser proporcionales a la naturaleza, el contexto y los riesgos del tratamiento.

7.6. Cuando la evaluación identifique riesgos elevados, el responsable deberá adoptar medidas adicionales para mitigarlos. Si, pese a ello, persiste un riesgo elevado, deberá consultar a la autoridad de protección de datos competente con anterioridad al inicio del tratamiento o al despliegue del sistema o bien abstenerse de iniciar el tratamiento. Cuando la normativa aplicable prevea un procedimiento específico para estos supuestos, se actuará conforme a él. Requerirán una atención especialmente rigurosa en la identificación, la evaluación y la gestión del riesgo, entre otros, los tratamientos que impliquen decisiones automatizadas que produzcan efectos jurídicos sobre las personas o las afecten significativamente, la puntuación social o la evaluación generalizada de personas, la identificación biométrica o el reconocimiento o inferencia de emociones, la identificación biométrica remota en espacios públicos, el tratamiento de neurodatos o la generación de inferencias sobre estados cognitivos, emocionales o conductuales, el tratamiento de datos de personas o grupos en situación de vulnerabilidad, o la reutilización masiva de datos personales para el desarrollo o entrenamiento de sistemas de inteligencia artificial, todo ello conforme a lo previsto en el Estándar en materia de grupos de protección reforzada, neurodatos y patrones tecnológicos y de uso, y sin perjuicio de los tratamientos que el Estándar declara incompatibles con los principios de protección de datos.

7.7. La adhesión a códigos de conducta, normas técnicas o esquemas de certificación, así como la evaluación, la autorización, la certificación o la supervisión de un sistema de inteligencia artificial conforme a la normativa específica que le resulte aplicable, podrán facilitar la aplicación del Estándar y acreditar la adopción de buenas prácticas, pero no sustituyen la responsabilidad que corresponde al responsable de garantizar el cumplimiento de la normativa de protección de datos ni la adecuada gestión de los riesgos derivados del uso de sistemas de inteligencia artificial, y se entienden sin perjuicio de las competencias que correspondan a las autoridades de protección de datos conforme al ordenamiento aplicable

7.8. En los tratamientos que presenten riesgos elevados, la revisión periódica y la mejora continua de las medidas de gobernanza algorítmica previstas en este capítulo se intensificarán en atención a la entidad de esos riesgos, y se recomienda realizar además revisiones o auditorías independientes cuando contribuyan a verificar la eficacia de las medidas adoptadas, fortalecer la gobernanza algorítmica y mejorar la gestión de los riesgos. Cuando se realicen, se procurará que se apoyen en criterios de evaluación previamente definidos y proporcionados al riesgo y que puedan acreditarse la independencia, la competencia técnica y la ausencia de conflictos de interés de quienes las lleven a cabo.

7.9. La documentación que se genere en base a estos Estándares deberá estar a disposición de la autoridad de protección de datos competente en el ejercicio de sus funciones de supervisión, conforme al numeral 42.4 de los Estándares de Protección de Datos de los Estados Iberoamericanos.

Capítulo VIII. Medidas técnicas y organizativas de protección de datos y seguridad en el uso de sistemas de inteligencia artificial

8.1. Amenazas y riesgos específicos

El responsable identificará y mantendrá actualizada una relación de amenazas específicas de la inteligencia artificial aplicable a su sistema, que comprende, entre otras, la extracción de datos de entrenamiento, la inversión de modelo, los ataques de pertenencia, las fugas a través de las instrucciones al sistema, la exposición de las representaciones vectoriales, la reidentificación, la memorización, el envenenamiento de los datos de entrenamiento, la recuperación indebida de información en sistemas de generación aumentada por recuperación y la inyección de instrucciones.

La identificación de estas amenazas y la valoración del riesgo que comportan para los derechos y libertades de las personas servirán de base para definir las medidas de seguridad, las pruebas y la respuesta frente a incidentes adecuadas para la protección de los datos personales tratados. Cuando el tratamiento combine distintas fuentes de información, deberá valorarse el riesgo de agregación y de reidentificación.

8.2. Controles técnicos y pruebas específicas

Las medidas de seguridad deberán mantenerse y revisarse durante todo el ciclo de vida del tratamiento de datos personales. El responsable adoptará medidas legales, técnicas y organizativas adecuadas para hacer frente a las amenazas identificadas y a os riesgos concretos que comportan para el tratamiento de datos personales. Estas podrán incluir, entre otras, controles de acceso, segregación de entornos, cifrado, monitorización, aislamiento de entornos o componentes, mecanismos de filtrado y validación de entradas y salidas, medidas para limitar la memorización o divulgación de datos personales y otras medidas técnicas y organizativas adaptadas al estado de la técnica, a la evolución de las amenazas y al riesgo del tratamiento. Asimismo, realizará pruebas adecuadas para verificar la eficacia de estas medidas en el marco de su revisión y, en especial, cuando se produzcan modificaciones relevantes en el tratamiento o en el sistema de inteligencia artificial. Para ello podrá apoyarse en normas, metodologías o estándares técnicos generalmente reconocidos, siempre que resulten adecuados a los riesgos del tratamiento y contribuyan a la mejora continua de los controles adoptados, sin perjuicio del valor que el Estándar atribuye a los códigos de conducta, las normas técnicas y los esquemas de certificación.

8.3. Técnicas de reducción del riesgo de identificación

8.3.1. La seudonimización reduce la exposición, pero no excluye el tratamiento del régimen general de protección de datos. La anonimización sólo podrá considerarse efectiva cuando se acredite que no existe un riesgo razonable de identificación o reidentificación atendidos los medios que razonablemente puedan utilizarse, apreciación que exige especial prudencia cuando el tratamiento, el sistema o sus resultados puedan facilitar la reidentificación o revelar datos por memorización. Estas técnicas constituyen medidas de protección de datos desde el diseño y por defecto, en los términos previstos en el Estándar.

8.3.2. Según el caso, podrán emplearse técnicas como el filtrado previo de datos, la supresión de duplicados, la reducción del nivel de detalle de la información, la seudonimización, la privacidad diferencial, los datos sintéticos u otras medidas destinadas a reducir el riesgo de identificación. La utilización de estas técnicas no exime del cumplimiento de la normativa de protección de datos ni de la adopción de las medidas de seguridad adecuadas.

8.4. Respuesta a incidentes

8.4.1. El responsable contemplará en sus procedimientos de gestión de incidentes las vulnerabilidades específicas derivadas de la utilización de sistemas de inteligencia artificial, en particular aquellas que puedan dar lugar a la revelación, extracción o utilización indebida de datos personales o a la pérdida de su disponibilidad, integrando estas vulnerabilidades en los procedimientos generales de gestión de incidentes. Asimismo, preverá medidas adecuadas para su detección, contención y mitigación, así como para la evaluación del riesgo para los derechos y libertades de las personas. Cuando proceda, la respuesta al incidente se ajustará al régimen general de protección de datos aplicable a las violaciones de seguridad de los datos personales, incluida la notificación que conforme a dicho régimen corresponda a la autoridad de protección de datos competente y, en su caso, a las personas titulares.

Capítulo IX. Especialidades aplicables a determinados patrones tecnológicos y de uso

Además de las disposiciones generales contenidas en el Estándar, cuando un tratamiento incorpore alguno de los siguientes patrones tecnológicos o de uso deberán aplicarse las especialidades previstas en este capítulo, en la medida en que resulten pertinentes.

9.1. Inteligencia artificial generativa

En los tratamientos que incorporen sistemas de inteligencia artificial generativa deberán adoptarse medidas específicas para evitar la introducción innecesaria de datos personales, propios o de terceros, en las instrucciones, definir las condiciones de conservación de las instrucciones y de las salidas, así como de su eventual utilización para el entrenamiento o la mejora de los modelos, prevenir la divulgación, recuperación o memorización indebida de datos personales y verificar los resultados cuando puedan utilizarse para adoptar decisiones sobre las personas. Esta verificación se entiende sin perjuicio de la exactitud exigible a las inferencias y a los resultados conforme al Estándar. Cuando el tratamiento comprenda el desarrollo o entrenamiento de modelos, deberá disponerse de información suficiente sobre el origen de los datos utilizados, conforme al deber de documentación del tratamiento previsto en el Estándar.

9.2. Inteligencia artificial agéntica

Cuando un sistema de inteligencia artificial pueda acceder de forma autónoma a datos personales o ejecutar acciones sobre ellos, el responsable deberá determinar el nivel de autonomía admisible atendiendo a las acciones que pueda ejecutar y a sus posibles efectos sobre las personas, de modo que a mayor entidad de esos efectos correspondan un menor grado de autonomía y una supervisión más intensa; delimitar los datos, herramientas, servicios o recursos a los que el sistema puede acceder; controlar la utilización de memorias persistentes cuando impliquen la conservación o reutilización de datos personales, que quedará sujeta a los principios de finalidad y de limitación del plazo de conservación conforme al Estándar y a la normativa de protección de datos; garantizar la trazabilidad de las actuaciones, incluidos los accesos a los recursos autorizados y las interacciones con otros sistemas, servicios o agentes; y prever mecanismos de supervisión humana y de suspensión o interrupción de las actuaciones, que deberán poder activarse con independencia del propio sistema, cuando resulte necesario para proteger los derechos de las personas. Las actuaciones autónomas del sistema deberán mantenerse dentro del tratamiento y de las finalidades previamente determinadas por el responsable. Cuando la actuación autónoma del sistema comporte una finalidad distinta de la inicialmente prevista, deberá verificarse que cuenta con una base jurídica adecuada y, cuando proceda, que resulta compatible con la finalidad original, conforme a lo previsto en el Estándar en materia de finalidades y reutilización de datos. La actuación autónoma del sistema no exime ni desplaza las responsabilidades que corresponden a quienes intervienen en el tratamiento conforme al Estándar y a la normativa de protección de datos.

9.3. Biometría y reconocimiento de emociones

Los tratamientos que incorporen sistemas de inteligencia artificial y técnicas biométricas, con independencia de que estas sean utilizadas por el propio sistema, deberán distinguir, entre otros supuestos, la verificación, la identificación y otros usos de técnicas biométricas, conforme a la distinción entre identificación y verificación de la identidad establecida en el Estándar en materia de minimización y proporcionalidad, aplicando, cuando proceda, las garantías reforzadas exigibles conforme a la normativa de protección de datos, incluidas las derivadas del régimen reforzado de las categorías especiales de datos previsto en el Estándar en materia de sensibilidad inferida cuando resulten de aplicación, y al riesgo del tratamiento. Requerirán una atención especialmente rigurosa la identificación biométrica remota y, cuando impliquen el tratamiento de datos personales, el reconocimiento o la inferencia de emociones y otros usos que, por su naturaleza, su contexto o su alcance, puedan afectar significativamente a los derechos de las personas, sin perjuicio de las prohibiciones, limitaciones o requisitos adicionales establecidos por la normativa específica aplicable y de lo dispuesto en el Estándar para los tratamientos que generen inferencias sobre estados cognitivos, emocionales o conductuales. En estos supuestos deberá poder acreditarse, conforme a los deberes de calidad y exactitud y de documentación del tratamiento previstos en el Estándar, la idoneidad y la fiabilidad del sistema para la finalidad y el contexto en que se emplea. Estos tratamientos quedan comprendidos, además, entre los que requieren una atención especialmente rigurosa en la identificación, la evaluación y la gestión del riesgo y entre los que determinan la intensificación de la revisión periódica y la mejora continua de las medidas de gobernanza algorítmica, en los términos y con las condiciones establecidos en el Estándar.

9.4. Puntuación, selección y priorización

Los tratamientos que incorporen sistemas de inteligencia artificial destinados a la evaluación, selección, clasificación o priorización de personas deberán incorporar medidas específicas para prevenir resultados discriminatorios conforme a lo previsto en el Estándar en materia de efectividad de las garantías y no discriminación algorítmica, garantizar una transparencia adecuada a su impacto conforme a lo previsto en el Estándar en materia de transparencia y explicabilidad y asegurar el respeto de las garantías aplicables a las decisiones individuales automatizadas cuando proceda, sin perjuicio de las prohibiciones, limitaciones o requisitos adicionales establecidos por la normativa específica aplicable y de la incompatibilidad declarada con carácter general en el Estándar para los mecanismos de puntuación social que produzcan efectos discriminatorios o restricciones injustificadas de derechos.

9.5. Obtención de datos accesibles al público

La utilización de datos personales accesibles al público para el desarrollo, entrenamiento o mejora de modelos o sistemas de inteligencia artificial se rige por la regla establecida en el Estándar en materia de lealtad, finalidad y compatibilidad, conforme a la cual dicho carácter no legitima por sí solo esa utilización. Dicho carácter tampoco excluye la aplicación del principio de finalidad. La reutilización de esos datos para finalidades distintas de las que justificaron su publicación o recogida requerirá una base jurídica adecuada y, cuando proceda, garantías reforzadas, sin perjuicio del análisis de compatibilidad exigible conforme a lo previsto en el Estándar en materia de finalidades y reutilización de datos. Estas

garantías resultarán especialmente exigibles cuando la reutilización se realice mediante recopilaciones masivas o se fundamente en el interés legítimo, en los ordenamientos que reconozcan esta base jurídica, sin que la accesibilidad pública de los datos constituya por sí sola dicho interés legítimo ni dispense de la ponderación exigible conforme al Estándar. En la determinación de la base jurídica adecuada y de las garantías que procedan se tendrán en cuenta la finalidad y el contexto que justificaron la puesta a disposición pública de los datos, así como las expectativas razonables de las personas respecto de dicha reutilización. Entre dichas garantías podrán figurar las técnicas de reducción del riesgo de identificación previstas en el Estándar, aplicadas con especial atención a las categorías especiales de datos y a los datos de personas o grupos en situación de vulnerabilidad. Las condiciones establecidas en este numeral operan como garantías adicionales exigibles a la reutilización y no como habilitación de esta. Lo dispuesto en este numeral se entiende sin perjuicio de las garantías reforzadas y de las medidas de evitación exigibles conforme al Estándar respecto de los datos personales de personas o grupos en situación de vulnerabilidad.

RED IBEROAMERICANA DE PROTECCION DE DATOS

Red Iberoamericana
Rede Ibero-Americana

Financiado por
la Unión Europea

Documento financiado por la Unión Europea, en colaboración con la Fundación para la Internacionalización de las Administraciones Públicas (FIAP) en el marco de la Alianza Digital UE-ALC. Las opiniones y los argumentos expresados en este documento no reflejan necesariamente los puntos de vista oficiales de la Unión Europea, ni de sus Estados miembros.