

Alianza de Institutos de Investigación Sanitaria

Grupo de trabajo 5: Inteligencia Artificial



Miembros:

IR-HUVH (Barcelona)
IBSAL (Salamanca)
IIS-FJD (Madrid)
IBS Granada (Granada)
IISGM (Madrid)
IdISSC (Madrid)
INIBIC (A Coruña)
IISGS (Galicia Sur)
IIS Biobizkaia (Bizkaia)
IMIBIC (Córdoba)
IIS La Fe (Valencia)

Diciembre 2025

ÍNDICE

Subproyecto 1: Documento de posicionamiento: inteligencia artificial en la investigación sanitaria.....	7
Resumen ejecutivo.....	8
1. Introducción: un posicionamiento estratégico y necesario.....	8
1.1. Contexto y relevancia.....	8
1.2. Justificación del consenso	9
1.3. Objetivos, alcance y público destinatario	9
1.4. Metodología de elaboración y consenso	9
2. Conceptos clave para un entendimiento común	9
3. Diagnóstico estratégico: oportunidades, capacidades y desafíos	10
3.1. Estado del arte	10
3.2. Oportunidades y capacidades en el sistema español	10
3.3. Desafíos transversales.....	11
4. Marco ético y regulatorio para una inteligencia artificial responsable.....	11
4.1. Principios éticos fundamentales para los institutos	12
4.2. Marco regulatorio y legal aplicable.....	12
5. Guía de buenas prácticas en el ciclo de vida de la investigación con inteligencia artificial	13
6. Posicionamiento y recomendaciones estratégicas de los institutos.....	14
6.1. Declaración de principios rectores.....	14
6.2. Recomendaciones para el personal investigador y los grupos de investigación ..	16
6.3. Recomendaciones para la implementación institucional en los institutos	17
6.4. Recomendaciones para la gobernanza	18
7. Hoja de ruta para la implementación (2025-2030)	19
7.1. Acciones inmediatas (0-12 meses).....	19
7.2. Objetivos a corto plazo (1-3 años)	20
7.3. Metas a medio-largo plazo (3-5+ años)	21
8. Recursos, capacitación y colaboración.....	22
9. Limitaciones y líneas de futuro.....	22
10. Conclusiones y llamada a la acción	23

11. Anexos	25
Anexo I: Glosario de términos y acrónimos.....	25
Anexo II: Tabla resumen de responsabilidades por actor	27
Anexo III: Lista de verificación para proyectos de IA en investigación sanitaria.....	28
12. Bibliografía	29
13. Normativa consultada.....	30
14. Figuras y Tablas	31
Subproyecto 2: Encuesta Nacional: Mapa sobre la situación de la investigación con Big Data e IA en los IISA.....	35
1. Información General del Centro	36
2. Recursos Humanos y Capacitación	37
3. Infraestructura Tecnológica	38
4. Gobernanza y modelo de gestión	43
5. Ética y regulatoria.....	45
6. Proyectos con uso de Big Data e IA en investigación biomédica:	49
7. Obstáculos y Retos:.....	53
8. Perspectivas y necesidades	54
9. Observaciones Adicionales:.....	55
Subproyecto 3: Guía de requisitos técnicos y regulatorios para proyectos con Big Data e Inteligencia Artificial (IA) en salud	56
1. Objetivo y alcance	58
2. Pùblicos objetivo y usos previstos	59
2.1. Pùblicos objetivo	59
2.2. Usos previstos:	59
3. Terminología y definiciones clave.....	60
3.1. Acrónimos de uso frecuente:	60
3.2. Definiciones Clave:	62
4. Marco regulatorio aplicable	65
4.1. MDR/IVDR y cualificación de software de IA como producto sanitario	65
4.2. Reglamento europeo de IA (AI Act) e interacción con MDR/IVDR	66
4.3. Marco nacional español y protección de datos	66
4.4. Normas técnicas internacionales de referencia.....	67
4.5. Ciberseguridad y resiliencia en hospitales.....	67

5. Tipología de proyectos y niveles de madurez (TRL)	68
5.1. Tipología de proyectos	68
5.2. TRL adaptados a IA en salud: hitos y evidencias mínimas	69
5.3. Decisiones de paso y responsabilidades	70
5.4. Correspondencia con evaluación y financiación	70
6. Requisitos sobre datos de salud y gobernanza.....	71
6.1. Calidad, representatividad y mitigación de sesgos	71
6.2. Gestión del ciclo de vida y plan de gestión de datos (DMP)	71
6.3. Seguridad, privacidad y cumplimiento (RGPD/LOPDGDD)	72
6.4. Aprendizaje federado y colaboración multicéntrica	74
6.5. Documentación de datos y modelos (<i>DataSheet</i> y <i>ModelCard</i>)	76
7. Requisitos técnicos del sistema de IA	77
7.1. Diseño y validación del modelo.....	77
7.2. Explicabilidad y trazabilidad	78
7.3. Monitorización, <i>drift</i> y recalibración.....	79
7.4. Seguridad del software y ciclo de vida (MLOps)	80
8. Interoperabilidad e integración en el entorno clínico.....	82
8.1. Requisitos funcionales de integración	82
8.2. Estándares y modelado semántico	83
8.3. Seguridad, identidad y control de acceso	84
8.4. Flujo clínico, usabilidad y seguridad del paciente	85
8.5. Entorno de ejecución y despliegue	86
8.6. Pruebas de interoperabilidad y verificación en entorno clínico	87
8.7. Operación multicéntrica y portabilidad	88
8.8. Gobierno de la integración y acuerdos	89
9. Evaluación y madurez: criterios y baremación	90
9.1. Metodología de evaluación (común PI/ICPS).....	90
9.2. Baremación para PI (TRL 1–3) — 36 puntos.....	91
9.3. Baremación para ICPS (TRL ≥4) — 48 puntos	92
9.4. Procedimiento de evaluación y actas.....	93
9.5. Resultados y documentación de la evaluación	94
10. Implementación práctica y gestión del cambio	95
10.1. Gobernanza y liderazgo del cambio	95

10.2.	Análisis de preparación e impacto organizativo	96
10.3.	Formación, acreditación y soporte a usuarios	97
10.4.	Pilotos y despliegue por fases	97
10.5.	Gestión de cambios técnicos y versiones.....	98
10.6.	Seguridad del paciente y ética en operación	99
10.7.	Indicadores, resultados y cuadro de mando	100
10.8.	Sostenibilidad operativa y continuidad.....	101
10.9.	Participación de pacientes y comunicación	102
10.10.	Cierre, transferencia y retirada.....	103
11.	Sostenibilidad, escalabilidad y modelo económico	104
11.1.	Modelo económico y financiación	104
11.2.	Costes y recursos (CAPEX/OPEX).....	105
11.3.	Escalabilidad y despliegue multicéntrico	106
11.4.	Sostenibilidad operativa y continuidad.....	107
11.5.	Evaluación económica y resultados en salud.....	108
11.6.	Contratación, propiedad intelectual y licencias.....	108
11.7.	Riesgos estratégicos y mitigación.....	110
11.8.	Indicadores clave (KPIs) de sostenibilidad	110
12.	Referencias.....	113
12.1.	Normativa europea y española	113
12.2.	Guías reguladoras (MDCG/AEMPS).....	113
12.3.	Buenas prácticas clínicas e investigación	114
12.4.	Normas técnicas (calidad, software y riesgo)	114
12.5.	Ciberseguridad y resiliencia	114
12.6.	Datos, ética y gobernanza de la IA en salud.....	115
12.7.	Interoperabilidad y terminologías.....	115
13.	Anexos operativos	116
13.1.	Checklists transversales (verificación rápida)	116
13.2.	Formulario de evaluación PI (TRL 1–3).....	120
13.3.	Formulario de evaluación ICPS (TRL ≥4).....	120
13.4.	Plantilla DataSheet (dataset).....	121
13.5.	Plantilla ModelCard (modelo)	121
13.6.	Plan de monitorización posdespliegue (plantilla)	122
13.7.	Gestión de cambios y homologación por versión (plantilla).....	122

13.8.	Dossier de interoperabilidad (plantilla)	122
13.9.	Expediente regulatorio mínimo (si cualifica como PS).....	123
13.10.	DPIA/EIPD (esqueleto)	123
13.11.	Caso de negocio y evaluación económica (plantilla)	123
13.12.	Reportes tipo	123
13.13.	Política SBOM y boletín de vulnerabilidades (plantilla).....	124
13.14.	Procedimiento de divulgación de vulnerabilidades (VDP).....	124
13.15.	Plantilla de modelado de amenazas (STRIDE) y matriz 14971.....	124

Subproyecto 1: Documento de posicionamiento: inteligencia artificial en la investigación sanitaria

Versión: 1.2 · **Fecha:** 03/12/2025

Grupo de trabajo: GdT5

Coordinadores: IdISSC · IISGS

Control de versiones

VERSIÓN	FECHA	CAMBIOS PRINCIPALES
1.0	15/09/2025	Publicación versión inicial
1.1	30/09/2025	Revisión GdT5-IA
1.2	03/12/2025	Revisión GT-IA ISCIII

Objetivo: Establecer el posicionamiento común de los Institutos de Investigación Sanitaria sobre el uso de la inteligencia artificial (IA) en la generación de conocimiento científico y la gestión de datos sanitarios.

A quién va dirigida: Investigadores, instituciones, agencias financieras, responsables políticos, ciudadanía.

Qué aporta:

1. Una descripción del marco y compromiso que deben guiar el diseño de estudios, la gestión de datos, el desarrollo de modelos y la comunicación de resultados
2. Una hoja de ruta con acciones inmediatas, objetivos a corto plazo y metas a medio-largo plazo.

Estado del documento: en revisión final, versión semidefinitiva

Fecha de aprobación: [Pendiente]

Próxima revisión programada: [24 meses desde aprobación]

Subproyecto 1: Documento de posicionamiento: inteligencia artificial en la investigación sanitaria

Resumen ejecutivo

Los Institutos de Investigación Sanitaria del Instituto de Salud Carlos III establecen mediante este documento su posición común sobre el uso de inteligencia artificial en la generación de conocimiento científico y la gestión de datos sanitarios. Nos comprometemos a desarrollar una IA centrada en el paciente, transparente, equitativa y basada en evidencia científica sólida. Este compromiso se materializa en seis principios éticos fundamentales —beneficencia y no maleficencia, equidad y justicia, transparencia y explicabilidad, privacidad y protección de datos, responsabilidad, y supervisión humana— que guiarán el diseño de estudios, la gestión de datos, el desarrollo de modelos y la comunicación de resultados. La hoja de ruta contempla acciones inmediatas (creación de grupos de trabajo, formación urgente, protocolos de datos), objetivos a corto plazo (políticas comunes, proyectos piloto, infraestructuras federadas) y metas a medio-largo plazo (consolidación de buenas prácticas, escalado de soluciones validadas, posicionamiento internacional). Este posicionamiento se coordina con los subproyectos específicos del Grupo de Trabajo 5, dedicados al marco regulatorio y de datos, a la guía operativa para proyectos y a la implementación institucional, que desarrollan con mayor detalle los requisitos éticos, legales y organizativos para los IIS.

Invitamos a investigadores, instituciones, agencias financieras, responsables políticos y ciudadanos a implementar conjuntamente este marco común. La participación ciudadana será esencial para construir la confianza necesaria en estas tecnologías. Solo mediante acción coordinada la inteligencia artificial podrá cumplir su promesa de mejorar la salud de las personas mientras respeta sus derechos fundamentales.

1. Introducción: un posicionamiento estratégico y necesario

La inteligencia artificial permite analizar volúmenes masivos de datos sanitarios, generar predicciones precisas, acelerar el descubrimiento de tratamientos y apoyar decisiones clínicas complejas. España cuenta con capacidades digitales significativas y la Unión Europea impulsa iniciativas como la Ley de IA y el Espacio Europeo de Datos Sanitarios. Sin embargo, el potencial transformador se ve acompañado de riesgos relacionados con privacidad, sesgos algorítmicos, transparencia y fiabilidad. La dispersión de esfuerzos y la ausencia de un marco común dificultan el aprovechamiento ordenado de esta tecnología, razón por la cual el Instituto de Salud Carlos III ha encargado este consenso para guiar el empleo responsable de la IA y alinear la estrategia nacional con estándares internacionales.

1.1. Contexto y relevancia

La IA encuentra aplicaciones en todas las modalidades de investigación sanitaria: básica, clínica, translacional, epidemiológica y de servicios de salud. La Unión Europea promueve innovación responsable mediante la Ley de IA, que clasifica sistemas según riesgo y prohíbe prácticas como manipulación cognitiva o puntuación social. Los sistemas sanitarios se consideran de alto riesgo,

implicando requisitos estrictos. El Espacio Europeo de Datos Sanitarios facilitará acceso a historiales clínicos transfronterizos y uso secundario de datos para investigación con garantías de privacidad. España, con su digitalización sanitaria, está bien posicionada para liderar esta implementación, aunque debe abordar retos de gobernanza, financiación, competencias digitales, calidad de datos e interoperabilidad.

1.2. Justificación del consenso

La IA puede mejorar prevención, diagnóstico y tratamiento, pero también amplificar sesgos o generar resultados opacos. El RGPD y la LOPDGDD establecen principios operativos de licitud, transparencia, minimización de datos, exactitud, limitación temporal, integridad y confidencialidad, requiriendo seudonimización, privacidad desde el diseño, y consentimiento informado explícito para investigación biomédica. La Ley de IA considera de alto riesgo los sistemas en dispositivos médicos, obligando a cumplir requisitos estrictos de seguridad y rendimiento. Frente a este panorama regulatorio complejo y evolutivo, los institutos consideramos necesario un posicionamiento conjunto que oriente la investigación y promueva un uso ético, responsable y alineado con la normativa.

1.3. Objetivos, alcance y público destinatario

El objetivo principal es proporcionar un marco de referencia común para la utilización responsable de IA en investigación sanitaria bajo el ámbito del ISCIII. Buscamos armonizar definiciones, identificar oportunidades y desafíos, establecer principios éticos, explicar el marco regulador y ofrecer guías prácticas. El alcance incluye investigación básica, clínica, traslacional, epidemiológica y de servicios de salud. Los destinatarios son direcciones científicas, investigadores, técnicos, gestores, responsables de gobernanza, comités de ética, agencias financieras, autoridades sanitarias y ciudadanía interesada. En particular, las agencias financieras encontrarán en este documento criterios de referencia para incorporar requisitos explícitos de ética, seguridad, equidad y gobernanza de la IA en las convocatorias competitivas y en la evaluación de proyectos.

1.4. Metodología de elaboración y consenso

El consenso se elaboró mediante proceso deliberativo liderado por el Instituto de Investigación Sanitaria Galicia Sur y el Instituto de Investigación Sanitaria San Carlos. Se revisaron marcos éticos de OMS y UNESCO, normativa vigente en protección de datos, dispositivos médicos, investigación biomédica y ciencia, y se analizaron experiencias nacionales e internacionales. El borrador se compartió con los miembros del grupo de trabajo de IA de la Alianza de Institutos para incorporar aportaciones. Asimismo, los institutos se comprometen a revisar y actualizar este posicionamiento con una periodicidad aproximada de dos años, integrando cambios regulatorios, avances tecnológicos y experiencia acumulada en la red.

2. Conceptos clave para un entendimiento común

La comunicación efectiva entre disciplinas requiere lenguaje compartido. La **inteligencia artificial** designa técnicas computacionales que permiten a sistemas realizar tareas que típicamente

requieren razonamiento, aprendizaje, percepción o interacción adaptativa. El **aprendizaje automático** (machine learning) se refiere a algoritmos que encuentran patrones en datos y mejoran con la experiencia sin programación explícita para cada tarea. El **aprendizaje profundo** (deep learning) constituye un subconjunto basado en redes neuronales con múltiples capas que representan información jerárquicamente. La **IA generativa** describe modelos que producen contenido novedoso como textos o imágenes. El término **big data** se refiere a conjuntos de datos caracterizados por gran volumen, diversidad y velocidad de generación. Los **sistemas predictivos** estiman probabilidad de eventos futuros a partir de datos históricos.

Por coherencia terminológica, todas las siglas se utilizan tras definir previamente el término completo en castellano y se mantienen de manera sistemática los acrónimos oficiales de la normativa europea y española a lo largo del documento.

En este consenso, la **investigación sanitaria** abarca modalidades complementarias: básica (mecanismos biológicos fundamentales), clínica (estudios con personas o muestras biológicas humanas), translacional (traslado de hallazgos al paciente), epidemiológica (distribución y determinantes de enfermedades en poblaciones) y de servicios de salud (organización y eficiencia de sistemas sanitarios). La IA puede actuar en todas estas áreas con consideraciones metodológicas y éticas específicas en cada caso.

[Ver Tabla 1: Definiciones operativas de conceptos fundamentales]

3. Diagnóstico estratégico: oportunidades, capacidades y desafíos

3.1. Estado del arte

Las aplicaciones de IA en investigación sanitaria demuestran potencial transformador. En análisis de imágenes médicas, los algoritmos alcanzan precisión comparable o superior a especialistas humanos en detección de patologías. En genómica y medicina de precisión, la IA identifica variantes asociadas a enfermedades complejas mediante análisis de millones de variaciones. En farmacología, acelera significativamente el descubrimiento de compuestos al predecir propiedades moleculares y simular interacciones. Los sistemas predictivos se aplican en epidemiología para anticipar brotes y modelizar evolución de epidemias. En gestión de servicios, optimizan asignación de recursos y mejoran logística hospitalaria. Aunque estas aplicaciones han demostrado viabilidad técnica en condiciones controladas, muchas permanecen en fase experimental y requieren evaluación clínica rigurosa. La publicación de guías como SPIRIT-AI, CONSORT-AI, TRIPOD-AI, STARD-AI y CHART refleja el reconocimiento de la necesidad de estándares específicos que aseguren investigación robusta, transparente y reproducible.

3.2. Oportunidades y capacidades en el sistema español

El sistema español posee activos valiosos: biobancos integrados en la Red Nacional que ofrecen muestras bien caracterizadas vinculadas a datos clínicos longitudinales, redes de datos clínicos del SNS con potencial significativo cuando se superen desafíos de interoperabilidad, la Red Española de Supercomputación y centros de análisis de datos que aportan capacidad de procesamiento esencial. El EHDS conectará historiales clínicos a escala europea permitiendo uso secundario de datos seudonimizados para investigación con prohibición de usos dañinos. España está

relativamente bien posicionada por su digitalización sanitaria, aunque persisten desafíos: la gobernanza de datos requiere mayor flexibilidad, la financiación de infraestructuras debe asegurarse establemente, las competencias digitales necesitan fortalecerse, la calidad de datos clínicos debe mejorar y la interoperabilidad técnica y semántica requiere inversión sostenida. Existen ejemplos concretos de uso exitoso en proyectos de hospitales e institutos que han mejorado detección de enfermedades raras y análisis de imágenes radiológicas, demostrando no solo potencial técnico sino también disposición del sistema.

3.3. Desafíos transversales

El desarrollo responsable enfrenta desafíos que cruzan múltiples dominios. **Desde la perspectiva técnica y metodológica**, la calidad y representatividad de datos son determinantes. Los conjuntos pueden contener sesgos de múltiples fuentes (selección, medición, confusión) que se propagan y amplifican en modelos resultantes. La ausencia de estándares para preparación de datos limita reproducibilidad. El entrenamiento y validación requieren metodología rigurosa con separación estricta entre datos de desarrollo y evaluación, validación cruzada apropiada, y pruebas de generalización en poblaciones distintas.

En el ámbito ético y social, los sesgos y la equidad son preocupaciones centrales. La IA puede perpetuar desigualdades si no se evalúan datos y modelos con perspectiva de género, diversidad étnica, diferencias por edad y determinantes socioeconómicos. Un modelo entrenado predominantemente con datos de población masculina puede funcionar peor en mujeres; sistemas desarrollados con datos de hospitales urbanos de nivel terciario pueden no generalizar a entornos rurales. La protección de privacidad está regulada por RGPD y LOPDGDD, que exigen minimización de datos, transparencia, seudonimización y consentimiento informado. La seudonimización y anonimización presentan desafíos técnicos considerables: la combinación de múltiples variables aparentemente inocuas puede permitir reidentificación. La aceptación social depende críticamente de la confianza en que la IA respeta derechos fundamentales, se utiliza para beneficio colectivo, y permanece bajo supervisión de profesionales cualificados.

En la dimensión regulatoria y de gobernanza, se requiere coordinación entre instituciones y agilización de procesos normativos. La Ley de IA impone requisitos estrictos a sistemas de alto riesgo: gestión formal de riesgos, gobernanza rigurosa de datos, documentación técnica exhaustiva, transparencia, mecanismos de supervisión humana, métricas validadas de precisión y robustez, y salvaguardas de ciberseguridad. El Reglamento de Dispositivos Médicos establece que sistemas de IA utilizados como componentes de seguridad deben someterse a evaluación de conformidad y vigilancia postcomercialización continua. Los institutos deben establecer políticas claras, comités éticos con competencia en IA, y unidades especializadas de soporte. La financiación y formación son factores críticos: la escasez de recursos humanos especializados y mecanismos de carrera limitan la velocidad y calidad del desarrollo. La reforma de la Ley de la Ciencia aborda parcialmente estas cuestiones reforzando financiación estable, introduciendo itinerarios profesionales flexibles, y estableciendo requisitos de igualdad de género.

[Ver Figura 1: Marco Ético para IA en Investigación Sanitaria]

4. Marco ético y regulatorio para una inteligencia artificial responsable

4.1. Principios éticos fundamentales para los institutos

El desarrollo de IA en investigación sanitaria se guiará por seis principios éticos que constituyen compromisos operativos, no aspiraciones abstractas. La **beneficencia y no maleficencia** establece que los proyectos deben maximizar el beneficio para la salud mientras minimizan activamente el riesgo de daños, requiriendo evaluación cuidadosa de si el uso de IA es apropiado frente a alternativas más simples, consideración de consecuencias no intencionadas, y diseño de salvaguardas.

La **equidad y justicia** exige evitar perpetuar o amplificar sesgos y desigualdades. Las tecnologías desarrolladas deben beneficiar a todas las personas sin discriminación. Esto requiere representación adecuada de poblaciones diversas en datos de entrenamiento, evaluación desagregada del rendimiento en diferentes subgrupos, y ajustes cuando se detecten disparidades injustificadas. La equidad en IA no significa tratar a todos exactamente igual sino asegurar que nadie se vea sistemáticamente desfavorecido.

La **transparencia y explicabilidad** implica que los modelos deben ser comprensibles para profesionales y, en la medida posible, para pacientes afectados. Aunque ciertos modelos de aprendizaje profundo son inherentemente complejos, existen técnicas de IA explicable que pueden clarificar qué factores influyen en las predicciones. La transparencia se extiende a datos utilizados, métodos de desarrollo, limitaciones conocidas, y resultados de evaluaciones. Esta apertura facilita escrutinio crítico y genera confianza necesaria para adopción clínica.

La **privacidad y protección de datos** requiere salvaguardar identidad y derechos de personas cuyos datos se utilizan. El cumplimiento del RGPD implica implementar minimización de datos, adoptar técnicas de seudonimización o anonimización cuando sea apropiado, establecer períodos de retención limitados con destrucción segura posterior, y diseñar sistemas con privacidad incorporada desde el inicio (privacy by design).

La **responsabilidad y rendición de cuentas** implica que investigadores e instituciones asumen responsabilidad de resultados y pueden auditar modelos para comprender decisiones específicas. Deben existir cadenas claras de responsabilidad que identifiquen quién es responsable de cada fase del desarrollo y uso de sistemas. Cuando se detectan errores o consecuencias no intencionadas, deben existir mecanismos para corregirlos, aprender de ellos y comunicarlos apropiadamente.

La **supervisión humana cualificada** significa que decisiones críticas, especialmente aquellas que afectan directamente a salud o derechos de personas, deben mantener intervención y control de profesionales capacitados. Los sistemas de IA son herramientas de apoyo a la decisión humana, no sustitutos de ella. Los profesionales deben comprender capacidades y limitaciones de los sistemas que utilizan, poder cuestionar y anular recomendaciones algorítmicas cuando su juicio clínico así lo indique, y ser capaces de explicar y justificar sus decisiones a los pacientes.

Estos principios se alinean con los marcos éticos de OMS y UNESCO, que subrayan la protección de la autonomía humana, la promoción del bienestar y el interés público, el aseguramiento de transparencia e inteligibilidad, el fomento de responsabilidad clara, la garantía de inclusividad y equidad, y la promoción de IA responsiva a necesidades reales y sostenible a largo plazo.

4.2. Marco regulatorio y legal aplicable

La **Ley de Inteligencia Artificial de la UE** establece un enfoque basado en riesgos. Prohíbe aplicaciones de riesgo inaceptable (manipulación conductual subliminal, explotación de

vulnerabilidades, puntuación social, identificación biométrica en tiempo real en espacios públicos). Los sistemas de alto riesgo, categoría que incluye aplicaciones sanitarias, deben cumplir requisitos estrictos: gestión de riesgos, gobernanza de datos (conjuntos de entrenamiento de alta calidad, representativos, sin sesgos), documentación técnica exhaustiva, transparencia hacia usuarios, diseño que permita supervisión humana efectiva, y precisión, robustez y ciberseguridad demostradas. La ley reconoce que sistemas de IA incorporados en dispositivos médicos deben cumplir requisitos de ambos marcos regulatorios (Ley de IA y Reglamento de Dispositivos Médicos). La **Ley 14/2007 de Investigación Biomédica** (España) afirma integridad y dignidad humana, garantiza autonomía y consentimiento informado, prohíbe discriminación, asegura confidencialidad y exige donación libre de materiales biológicos. La **reforma de la Ley de la Ciencia** refuerza financiación estable, introduce contratos indefinidos, crea itinerarios postdoctorales, y obliga a instituciones a implementar planes de igualdad y protocolos contra acoso.

El **RGPD y la LOPDGDD** establecen principios vinculantes de licitud, transparencia, minimización de datos, exactitud, limitación temporal, integridad y confidencialidad. Para investigación biomédica se requiere consentimiento informado específico. Las evaluaciones de impacto en protección de datos son obligatorias cuando el tratamiento pueda implicar alto riesgo para derechos de las personas.

La **OMS** ha publicado directrices para modelos multimodales que enfatizan riesgos de privacidad, necesidad de evaluaciones de impacto, técnicas de preservación de privacidad (privacidad diferencial, aprendizaje federado), y corresponsabilidad entre desarrolladores y proveedores. La **UNESCO** articula principios de proporcionalidad, equidad, autonomía, privacidad, seguridad, transparencia, responsabilidad, sostenibilidad, supervisión humana y alfabetización.

El **Espacio Europeo de Datos Sanitarios** busca permitir que ciudadanos accedan y controlen sus historias clínicas en toda la UE (uso primario) y habilitar que datos anonimizados o seudonimizados se utilicen para investigación, innovación y formulación de políticas (uso secundario), con prohibición expresa de usos para decisiones perjudiciales o discriminación. La implementación requerirá estructura de gobernanza compleja e infraestructura técnica paneuropea. España tiene oportunidad de ejercer liderazgo, aunque debe abordar desafíos de gobernanza, financiación, capacidades técnicas, calidad de datos, interoperabilidad y educación.

En este marco, el Espacio Europeo de Datos Sanitarios (EHDS) distingue entre uso primario de datos, orientado a garantizar una atención sanitaria transfronteriza adecuada, y uso secundario, que obliga a los tenedores de datos de salud a ponerlos a disposición para fines legítimos definidos más allá de la mera investigación, como salud pública, planificación y evaluación de políticas.

Los Institutos de Investigación Sanitaria asumen que la participación en el EHDS requiere mecanismos robustos de rendición de cuentas, canales de comunicación accesibles para personas y organizaciones que deseen formular consultas o reclamaciones, y una documentación clara de responsabilidades en cada flujo de datos y proyecto basado en IA.

5. Guía de buenas prácticas en el ciclo de vida de la investigación con inteligencia artificial

El desarrollo responsable de IA requiere atención cuidadosa a lo largo de todo el ciclo de vida del proyecto, desde la concepción inicial hasta la difusión de resultados. La ética debe integrarse desde

el diseño, no añadirse posteriormente como requisito administrativo. Los estudios deben definir objetivos claros que justifiquen el uso de IA frente a métodos más simples, identificar tempranamente riesgos potenciales, e involucrar equipos multidisciplinarios que incluyan científicos de datos, clínicos, epidemiólogos, especialistas en ética y representantes de pacientes. La calidad de los datos determina fundamentalmente la calidad de los modelos. Se deben establecer criterios para asegurar que los datos son representativos de la población diana, minimizar errores y valores perdidos, y cumplir rigurosamente con RGPD y LOPDGDD mediante minimización de datos, seudonimización, y consentimiento informado explícito. La gestión debe incluir documentación exhaustiva del origen, contexto clínico y transformaciones de los datos, con mecanismos de monitorización de equidad y calidad a lo largo del tiempo.

El desarrollo de modelos debe mitigar proactivamente sesgos mediante técnicas como repesado de muestras, aprendizaje justo, o calibración por subgrupos. La validación rigurosa requiere conjuntos de datos independientes, siendo la validación externa en poblaciones o contextos diferentes el estándar oro para evaluar generalización. Las métricas deben evaluarse de manera desagregada por subgrupos relevantes para detectar disparidades.

La implementación experimental requiere supervisión continua mediante comités que monitoricen el impacto del sistema y garanticen que no vulnera derechos. La supervisión humana efectiva es esencial: los sistemas deben proporcionar información comprensible, alertar cuando la incertidumbre es alta, y permitir intervención humana fácil y rápida.

La difusión de resultados debe seguir estándares de transparencia y reproducibilidad (CONSORT-AI, SPIRIT-AI, TRIPOD-AI, STARD-AI, CHART). Es importante registrar protocolos en repositorios públicos, describir detalladamente algoritmos, datos y métricas, reportar honestamente limitaciones y resultados negativos, y facilitar acceso a modelos y códigos cuando sea posible. La autoría debe reflejar todas las contribuciones sustanciales, incluyendo las de pacientes y ciudadanos.

[Ver Tabla 2: Guía de buenas prácticas en el ciclo de vida]

6. Posicionamiento y recomendaciones estratégicas de los institutos

6.1. Declaración de principios rectores

Los Institutos de Investigación Sanitaria del Instituto de Salud Carlos III nos comprometemos a promover una inteligencia artificial centrada en las personas, basada en evidencia científica, transparente, equitativa y responsable. Este compromiso implica respetar derechos fundamentales y dignidad de las personas en todas nuestras actividades, asegurar calidad y robustez de los estudios mediante metodología rigurosa, valorar la diversidad de datos y contextos clínicos, y fomentar rendición de cuentas mediante mecanismos claros de responsabilidad.

Una IA centrada en las personas significa que todas las decisiones tecnológicas, metodológicas y de implementación priorizarán el bienestar humano por encima de consideraciones técnicas o económicas. Cada sistema se diseñará considerando necesidades reales de pacientes, profesionales sanitarios y comunidades. La IA no es un fin en sí misma sino una herramienta al servicio de la salud pública y el progreso científico.

Basar nuestro trabajo en evidencia científica significa que todos los sistemas se fundamentarán en datos validados, metodologías rigurosas y resultados reproducibles. Rechazamos aproximaciones especulativas o impulsadas exclusivamente por intereses comerciales que carezcan de sustento empírico sólido. Promoveremos cultura de investigación donde validación externa, revisión por pares y publicación abierta sean prácticas estándar.

La transparencia se erige como principio cardinal. Fomentaremos desarrollo de sistemas explicables que permitan comprender cómo se llega a conclusiones o recomendaciones. Esta transparencia se extenderá a aspectos técnicos de algoritmos y a procesos de toma de decisiones institucionales, criterios de selección de proyectos y fuentes de financiación, garantizando integridad científica y confianza pública. Equilibraremos la transparencia con consideraciones legítimas como protección de datos personales, confidencialidad cuando existan colaboraciones que lo requieran, y seguridad cuando la divulgación pudiera crear riesgos.

La equidad requiere atención especial a posibles fuentes de sesgo y discriminación. Implementaremos mecanismos de auditoría que identifiquen y corrijan sesgos algorítmicos, asegurando que los beneficios lleguen a todos los segmentos de la población sin excepción. Reconocemos que los sistemas pueden perpetuar o amplificar desigualdades existentes si no se diseñan con conciencia crítica sobre disparidades en salud relacionadas con género, etnia, nivel socioeconómico, edad o ubicación geográfica. Por ello, se promoverá la recogida y el análisis de datos desagregados que incluyan variables de diversidad étnica y social, nivel socioeconómico y contexto asistencial, siempre con garantías adecuadas de privacidad y proporcionalidad.

La responsabilidad implica establecer cadenas claras de rendición de cuentas donde se identifiquen responsables de cada fase del ciclo de vida de un sistema, desde diseño y desarrollo hasta implementación y monitorización. Esto incluye capacidad de rastrear decisiones, documentar procesos y responder ante posibles errores o consecuencias no previstas. Desarrollaremos marcos de gobernanza que definan roles, responsabilidades y mecanismos de supervisión efectivos.

El respeto a derechos fundamentales y dignidad de las personas constituye un límite infranqueable que ninguna innovación tecnológica puede traspasar. Garantizaremos que todos los proyectos se sometan a evaluaciones éticas rigurosas que consideren no solo beneficios potenciales sino también riesgos para derechos individuales y colectivos, incluyendo privacidad, autonomía, consentimiento informado, no discriminación y protección de datos personales sensibles.

Reconocemos la necesidad imperiosa de colaboración interinstitucional y diálogo con la sociedad para construir confianza y legitimidad. Los desafíos que plantea la IA trascienden capacidades de cualquier institución individual. La investigación requiere convergencia de múltiples disciplinas: medicina, informática, estadística, ética, derecho y ciencias sociales. Esta interdisciplinariedad solo puede lograrse mediante alianzas estratégicas que permitan compartir conocimientos, recursos e infraestructuras entre institutos, universidades, hospitales, organismos reguladores y organizaciones de pacientes.

El diálogo con la sociedad representa un compromiso con democratización del conocimiento y participación ciudadana en definición de prioridades de investigación. Desarrollaremos estrategias de comunicación pública que expliquen de manera accesible qué es la IA, cómo se aplica en salud, qué beneficios puede aportar y qué riesgos conlleva. Este diálogo será bidireccional, creando espacios genuinos de escucha donde preocupaciones, expectativas y valores de la ciudadanía informen decisiones de investigación. La participación de asociaciones de pacientes será especialmente crucial.

Reconocemos también la importancia de sostenibilidad e impacto ambiental de la investigación. El entrenamiento de modelos de aprendizaje profundo consume cantidades significativas de energía y genera emisiones de carbono considerables. Adoptaremos prácticas de computación sostenible: optimización de algoritmos, uso de infraestructuras eficientes energéticamente, preferencia por centros de datos alimentados con energías renovables, y consideración del impacto ambiental como criterio en evaluación de proyectos. En coherencia con este principio, los proyectos estratégicos de IA deberán estimar y monitorizar su consumo energético y huella de carbono, integrando estos indicadores en la priorización, financiación y seguimiento de las actuaciones.

6.2. Recomendaciones para el personal investigador y los grupos de investigación

Los investigadores deben desarrollar competencias en IA, ética y protección de datos mediante formación estructurada y continua. Necesitamos programas que aborden diferentes niveles, desde cursos introductorios hasta avanzados, con énfasis en aplicaciones prácticas en salud. La formación en ética debe integrarse transversalmente, desarrollando sensibilidad para identificar dilemas potenciales. La protección de datos requiere formación especializada en principios del RGPD, técnicas de anonimización y pseudonimización, y responsabilidades legales.

La colaboración con expertos de distintas disciplinas y con pacientes es esencial para definir preguntas relevantes. Las preguntas más importantes surgen frecuentemente en intersecciones entre disciplinas. Debemos crear estructuras que faciliten estas colaboraciones mediante grupos multidisciplinarios, espacios compartidos, seminarios interdisciplinarios y sistemas de evaluación que reconozcan contribuciones interdisciplinarias. La participación de pacientes representa un cambio paradigmático: pasan de objetos pasivos a participantes activos y co-creadores del conocimiento, aportando perspectivas únicas sobre qué problemas son prioritarios y qué resultados son verdaderamente importantes.

Es esencial diseñar estudios con protocolos claros utilizando guías como SPIRIT-AI y TRIPOD-AI, adoptar buenas prácticas de recolección y gestión de datos, y documentar meticulosamente cada fase. Un protocolo de investigación en IA debe especificar objetivos, características de la población, fuentes de datos, métodos de preprocesamiento, técnicas de IA empleadas, procedimientos de validación, métricas de rendimiento, consideraciones éticas y protección de datos, y planes de difusión.

Los equipos deben evaluar modelos con criterios de equidad y reproducibilidad, someter proyectos a revisión ética rigurosa, y comunicar hallazgos de forma transparente. La evaluación de equidad requiere ir más allá de métricas agregadas para examinar rendimiento en diferentes subgrupos mediante análisis de paridad demográfica, paridad de oportunidades y paridad predictiva. La reproducibilidad requiere establecer semillas aleatorias, documentar exhaustivamente hiperparámetros y decisiones de diseño, utilizar entornos computacionales controlados, compartir código completo, y publicar resultados negativos. En los usos asistenciales o de apoyo a la decisión clínica se recomienda fijar umbrales cuantitativos mínimos de rendimiento (por ejemplo, sensibilidad, especificidad, precisión, F1 o coeficiente kappa) y documentar su cumplimiento antes de cualquier despliegue o escalado.

La comunicación transparente implica reportar no solo éxitos sino también limitaciones, fracasos y incertidumbres, evitando "hype" y afirmaciones exageradas. El principio de ciencia abierta promueve publicación en acceso abierto, depósito de preprints, compartición de datos de

investigación con salvaguardas de privacidad, y publicación de código fuente. La igualdad de género y diversidad deben guiar la composición de equipos y la selección de datos, reconociendo que equipos diversos producen mejor ciencia y que la inclusión es cuestión tanto de justicia como de excelencia.

6.3. Recomendaciones para la implementación institucional en los institutos

Los institutos deben establecer políticas institucionales que incluyan principios éticos, procedimientos de evaluación y requisitos de transparencia. Estas políticas deben ser documentos formales aprobados por órganos de gobierno que establezcan el marco normativo interno, no como declaraciones genéricas sino como instrumentos operativos que guíen decisiones concretas y establezcan responsabilidades claras. Deben especificar cómo se revisarán proyectos antes de su inicio, durante desarrollo y tras finalización, con criterios claros para determinar qué proyectos requieren evaluación ética específica para IA.

Es necesario crear unidades de apoyo en ciencia de datos y protección de datos que asesoren a grupos de investigación. Una unidad de ciencia de datos debe reunir expertos en estadística avanzada, aprendizaje automático, ingeniería de datos y computación de alto rendimiento, con funciones que incluyan asesoramiento en diseño de estudios, apoyo en selección de técnicas analíticas, desarrollo e implementación de modelos, validación técnica de resultados, formación de investigadores, gestión de infraestructuras computacionales y desarrollo de herramientas reutilizables. Esta unidad debe trabajar en colaboración estrecha con grupos como socios que participan en conceptualización de proyectos.

La unidad de protección de datos debe contar con expertos en derecho de protección de datos, seguridad de la información y técnicas de anonimización, con funciones que incluyan asesoramiento sobre cumplimiento del RGPD, realización de evaluaciones de impacto en protección de datos, diseño de procedimientos de consentimiento informado, implementación de medidas de seguridad, gestión de incidentes, formación, y relación con autoridad de protección de datos. Debe actuar como facilitadora de la investigación mediante soluciones prácticas como plantillas de consentimiento, procedimientos estandarizados de anonimización y acuerdos de tratamiento de datos.

Las instituciones deben garantizar acceso a infraestructuras de datos y supercomputación seguras y fomentar interoperabilidad con el EHDS. Las infraestructuras de datos seguras deben implementar múltiples capas de protección en niveles físicos, técnicos y organizativos, incluyendo controles de acceso estrictos, cifrado de datos en reposo y en tránsito, autenticación multifactor, control de acceso basado en roles, auditoría exhaustiva y sistemas de detección de intrusiones. El acceso a supercomputación es cada vez más necesario; los institutos deben evaluar diferentes modelos (infraestructura propia, acceso a infraestructuras compartidas, servicios en la nube) considerando volumen y sensibilidad de datos, necesidades computacionales, competencias técnicas, restricciones presupuestarias y requisitos normativos.

La interoperabilidad requiere promoción activa de estándares reconocidos internacionalmente como HL7 FHIR, DICOM, SNOMED CT, LOINC e ICD. Los institutos deben prepararse para participar en el EHDS adaptando infraestructuras, implementando estándares de interoperabilidad, participando en procesos de gobernanza, formando al personal y contribuyendo con sus datos al espacio común europeo cuando sea apropiado. En particular, se priorizará el uso de modelos de

datos e interoperabilidad basados en estándares como OMOP, HL7 FHIR o Beacon, identificando un estándar principal y otros complementarios en dominios emergentes o innovadores. Cuando se utilicen datos genéticos u ómicos, los institutos mantendrán la custodia en la institución de origen, favorecerán el procesamiento local mediante algoritmos portables y compartirán preferentemente resultados agregados y validados.

La inversión en formación continua y atracción de talento es crucial. Los programas de formación deben estar integrados en planes de carrera con oportunidades reales de desarrollo, reconocimiento de competencias adquiridas y progresión laboral. La atracción de talento requiere desarrollar estrategias integrales que enfaticen oportunidad de trabajar en problemas con impacto social directo, libertad intelectual, ambiente colaborativo, oportunidades de formación y desarrollo, estabilidad laboral, y equilibrio entre vida profesional y personal. La igualdad de género y diversidad deben guiar selección y promoción mediante medidas proactivas que establezcan objetivos cuantitativos de representación femenina, implementen procesos que minimicen sesgos de género, aseguren igualdad salarial, proporcionen medidas de conciliación, y promuevan modelos de referencia femeninos.

6.4. Recomendaciones para la gobernanza

La gobernanza debe basarse en transparencia, responsabilidad y participación. La transparencia tiene múltiples dimensiones: a nivel técnico implica que sistemas sean explicables; a nivel organizativo que procesos de toma de decisiones sean abiertos; a nivel institucional que institutos comuniquen públicamente sus políticas, proyectos desarrollados, resultados obtenidos y gestión de cuestiones éticas. La responsabilidad implica que existan cadenas claras de rendición de cuentas donde se identifiquen responsables de diferentes decisiones y acciones. La participación se refiere a inclusión de diferentes actores relevantes en procesos de gobernanza, incluyendo pacientes y ciudadanos, profesionales sanitarios, expertos en ética, derecho y ciencias sociales, y representantes de grupos potencialmente afectados por sesgos.

Los institutos deberían reforzar sus comités de ética de la investigación y crear comités específicos sobre IA que incluyan expertos en datos, clínicos, juristas, pacientes y ciudadanos. El reforzamiento debe incluir formación de miembros en particularidades éticas de IA, incorporación de expertos con competencias específicas, desarrollo de procedimientos y criterios específicos para evaluación ética de proyectos de IA, y establecimiento de procedimientos de seguimiento. La composición multidisciplinaria de comités específicos es esencial, con pacientes y ciudadanos como miembros de pleno derecho que aporten perspectivas únicas sobre qué riesgos son aceptables y qué beneficios son verdaderamente valiosos. Estos comités deberían integrar, como miembros estables o asesores ad hoc, profesionales con experiencia acreditada en ciencia de datos e inteligencia artificial, y contar con acceso ágil a asesoría externa en ética y gobernanza de datos.

Las funciones de estos comités deben incluir evaluación inicial de proyectos, supervisión de implementación mediante seguimiento durante desarrollo, y propuesta de mejoras trabajando constructivamente con investigadores. Los comités deben verse como aliados en búsqueda de excelencia ética, no como obstáculos burocráticos. Además, los institutos promoverán un foro permanente interdisciplinario sobre IA en salud que acompañe a los proyectos, armonice criterios entre comités y unidades técnicas, y facilite respuestas rápidas a consultas complejas.

La gobernanza debe alinearse con la Ley de IA y con guías de OMS y UNESCO, que exigen evitar prácticas prohibidas, proteger autonomía y garantizar supervisión humana. La Ley de IA establece

enfoque basado en riesgos con requisitos estrictos para sistemas de alto riesgo que institutos deben cumplir mediante implementación de sistemas de gestión de calidad, realización de evaluaciones de conformidad, mantenimiento de documentación técnica completa, y establecimiento de sistemas de vigilancia.

Las políticas institucionales deben contemplar mecanismos de evaluación periódica y rendición de cuentas, así como procedimientos para gestionar incidentes. La evaluación periódica debe abordar proyectos individuales durante desarrollo, al completarse y en seguimiento post-implementación; programas de investigación valorando si generan resultados valiosos, alineados con prioridades estratégicas, utilizando recursos eficientemente y cumpliendo con principios éticos; y políticas y gobernanza examinando si están siendo efectivas y qué mejoras serían necesarias. En las decisiones de adquisición y contratación de soluciones de IA se establecerán procedimientos específicos de gestión de conflictos de interés con proveedores, incluyendo la declaración pública de vínculos previos, la trazabilidad de interacciones y la evaluación independiente de las propuestas.

La rendición de cuentas implica que institutos informen regularmente sobre actividades de investigación en IA a diferentes audiencias mediante informes anuales que incluyan secciones específicas sobre actividades de IA. También debe incluir mecanismos para que actores externos puedan solicitar información, plantear preocupaciones o presentar quejas. Los procedimientos para gestionar incidentes deben especificar cómo se reportan mediante canales claros con garantías de no represalia, cómo se evalúa gravedad, qué acciones inmediatas deben tomarse, cómo se investigan causas, qué medidas correctivas se implementan, y cómo se comunican incidentes. La gestión debe realizarse con cultura de aprendizaje más que de culpabilización.

La coordinación con agencias financieras y autoridades sanitarias es esencial para asegurar coherencia reguladora y promover adopción de estándares. Los institutos deben mantener diálogo activo con agencias financieras para informar sobre necesidades, promover armonización de requisitos, y asegurar que requisitos sean apropiados y proporcionales. La coordinación con autoridades sanitarias es igualmente esencial para comprender requisitos regulatorios aplicables, anticipar cambios futuros, y contribuir con expertise técnico a elaboración de regulaciones apropiadas. La adopción de estándares técnicos y metodológicos es fundamental; institutos deben participar activamente en procesos de desarrollo de estándares y promover su adopción una vez establecidos. Esta coordinación incluirá la incorporación progresiva de requisitos explícitos de ética, seguridad, equidad y sostenibilidad de la IA en las convocatorias, bases reguladoras y guías de evaluación de proyectos.

[Ver Tabla 3 en Anexo: Resumen de Responsabilidades por Actor]

7. Hoja de ruta para la implementación (2025-2030)

La implementación efectiva requiere planificación temporal con prioridades claras y objetivos alcanzables en tres fases progresivas. La calendarización propuesta se alinea con los hitos de entrada en vigor de la Ley de Inteligencia Artificial de la UE, de forma que los institutos ajusten progresivamente sus estructuras, procesos y recursos a las nuevas obligaciones regulatorias.

7.1. Acciones inmediatas (0-12 meses)

Los primeros doce meses deben centrarse en establecer fundamentos organizativos y de conocimiento. Constituir grupos de trabajo en cada instituto con composición multidisciplinaria

para realizar diagnóstico exhaustivo de la situación actual mediante inventario detallado de capacidades existentes, identificación de necesidades y brechas, y detección de oportunidades aprovechables. Este diagnóstico debe coordinarse a nivel nacional para obtener visión panorámica del conjunto de la red.

Desarrollar programas de formación urgente en IA, ética y protección de datos para personal investigador y técnico. La formación en IA debe abordar aspectos conceptuales y prácticos con énfasis en aplicaciones específicas en salud. La formación en ética debe ir más allá de introducción superficial para abordar dilemas concretos. La formación en protección de datos debe proporcionar conocimiento operativo del RGPD y normativa nacional.

Establecer protocolos de consentimiento y gestión de datos acordes con RGPD y LOPDGDD. Los protocolos de consentimiento específicos para proyectos de IA deben explicar de manera comprensible qué es la IA, qué datos se recolectarán, cómo se procesarán, qué inferencias se generarán, cómo se protegerá privacidad, y qué derechos tienen participantes. Los protocolos de gestión deben abordar todo el ciclo de vida especificando procedimientos de recolección, estándares de documentación, medidas de seguridad, procedimientos de anonimización o pseudonimización, reglas de acceso, y procedimientos de respaldo.

Definir procedimientos de revisión ética específicos para proyectos con IA que complementen procedimientos estándar con consideraciones adicionales. Desarrollar formularios de solicitud ampliados que incluyan secciones específicas sobre descripción técnica del sistema, datos de entrenamiento, procedimientos de desarrollo y validación, métricas de rendimiento, análisis de equidad, explicabilidad, supervisión humana, y plan de monitorización.

Iniciar diálogos con organismos responsables del EHDS para preparar integración de infraestructuras, comprender requisitos técnicos y organizativos, influir en diseño desde perspectiva de necesidades de investigación, y establecer colaboraciones con otros institutos europeos.

7.2. Objetivos a corto plazo (1-3 años)

El período de uno a tres años representa fase de consolidación y expansión. Implementar políticas y guías comunes en todos los institutos mediante proceso participativo que involucre representantes de todos los institutos, con consulta amplia, pilotaje en algunos institutos, y refinamiento basado en experiencia. Desarrollar proyectos piloto colaborativos que permitan probar modelos en diferentes contextos clínicos y evaluar su impacto. Los proyectos deben incluir evaluación rigurosa no solo del rendimiento técnico sino también de impacto clínico real, usabilidad, aceptabilidad, eficiencia y equidad. En esta fase se priorizará una evaluación rigurosa de la integración de la IA en entornos clínicos reales, con indicadores de impacto, seguridad, equidad, aceptabilidad y eficiencia que informen decisiones de continuidad o escalado.

Crear infraestructuras de datos federadas y seguras que permitan análisis de datos distribuidos sin centralización física. Las infraestructuras federadas ofrecen ventajas en protección de privacidad y facilitan colaboración entre instituciones con restricciones sobre compartir datos directamente. Deben complementarse con medidas de seguridad robustas incluyendo cifrado, autenticación fuerte, control de acceso granular, auditoría de operaciones, y técnicas avanzadas como computación segura multipartita o privacidad diferencial.

Consolidar formación avanzada e incorporación de perfiles multidisciplinarios mediante programas estructurados de desarrollo profesional que proporcionen trayectorias claras de progresión.

Impulsar participación en iniciativas europeas y consorcios internacionales mediante estrategias proactivas para identificar oportunidades relevantes, preparar propuestas competitivas, y gestionar efectivamente proyectos financiados.

Evaluar proyectos con métricas de equidad, impacto clínico y retorno social. Las métricas de equidad evalúan si sistema funciona de manera justa para todos los grupos. Las métricas de impacto clínico evalúan si sistema mejora realmente resultados de salud o calidad de atención. Las métricas de retorno social evalúan valor más amplio que proyecto genera para sociedad. Desarrollar marcos de evaluación que integren estas múltiples dimensiones de manera coherente.

7.3. Metas a medio-largo plazo (3-5+ años)

El horizonte de tres a cinco años y más allá representa fase de maduración. Consolidar buenas prácticas que se institucionalizan y convierten en forma normal de trabajar mediante documentación exhaustiva, formación sistemática de todo personal nuevo, mecanismos de aseguramiento de calidad, y cultura organizacional que valore adherencia a buenas prácticas.

Escalar soluciones validadas expandiendo sistemas que han demostrado ser efectivos, seguros y aceptables en proyectos piloto a poblaciones más amplias y contextos más diversos. El escalado requiere consideración cuidadosa a nivel técnico, organizativo, regulatorio y económico. Reconocer que no todas las soluciones validadas serán apropiadas para escalado; desarrollar criterios claros basados en evidencia de efectividad, seguridad, equidad, aceptabilidad, viabilidad y sostenibilidad. Evaluar sistemáticamente impacto clínico y en salud pública de sistemas de IA mediante evaluación sistemática, planificada desde el inicio, realizada de manera continua, que incluya comparaciones con práctica habitual y considere no solo efectos intencionados sino también posibles efectos no intencionados. Integrar modelos con evidencia robusta en atención sanitaria requiriendo que sistemas se diseñen considerando contexto de uso desde inicio, evaluación rigurosa antes de implementación amplia, y aseguramiento de que sistemas operen siempre bajo supervisión humana efectiva.

Consolidar infraestructuras de datos y participación en EHDS permitirán posicionar institutos como referentes internacionales. Las infraestructuras consolidadas deben ser técnicamente robustas, interoperables, bien gobernadas y sostenibles. La participación activa en EHDS significa que institutos no solo están técnicamente preparados sino que están utilizándolo activamente en proyectos y contribuyendo con sus datos al espacio común europeo.

Establecer marcos para colaboración con industria y garantizar transferencia responsable de tecnología. Los marcos deben establecer principios y procedimientos claros que protejan interés público y aseguren que colaboraciones sean mutuamente beneficiosas y éticamente apropiadas, abordando qué tipos de colaboración son aceptables, cómo se gestionarán conflictos de interés, cómo se protegerán datos de pacientes, cómo se compartirán resultados, y cómo se gestionará propiedad intelectual.

Seguimiento permanente de evolución tecnológica es esencial dado ritmo extremadamente rápido de cambio. Los institutos deben establecer mecanismos para mantenerse actualizados sobre desarrollos relevantes mediante participación en conferencias, seguimiento de literatura, membresía en redes profesionales, y establecimiento de comités que revisen periódicamente estado del arte. El seguimiento debe informar actualización periódica de políticas y estrategias, identificación y gestión de riesgos emergentes, e identificación de oportunidades emergentes.

[Ver Figura 2: Hoja de Ruta para Implementación 2025-2030]

8. Recursos, capacitación y colaboración

El avance en IA demanda convergencia de competencias humanas, recursos materiales y colaboraciones estratégicas. El desarrollo requiere talento humano con formación especializada en ciencia de datos, estadística, informática, bioética, gestión, ética y Derecho. Los programas de formación deben estar integrados en planes de carrera y promover interdisciplinariedad, traduciéndose en oportunidades reales de desarrollo profesional y progresión laboral. Los IIS desplegarán programas estructurados de capacitación continua en IA, ética y gobernanza de datos, organizados por niveles de competencia y accesibles de forma equitativa para personal investigador, clínico, técnico y de gestión.

La igualdad de género y diversidad deben guiar selección y promoción del personal mediante medidas concretas como criterios objetivos y transparentes, composición equilibrada de comités de evaluación, promoción de modelos de referencia diversos, y políticas de conciliación. Las instituciones deben desarrollar infraestructuras de datos seguras y centros de supercomputación accesibles, cumpliendo con más altos estándares de seguridad y garantizando accesibilidad mediante políticas equitativas y transparentes.

La participación en EHDS facilitará acceso a datos estandarizados y seudoanonymizados, ampliando volumen y diversidad de datos disponibles mientras garantiza cumplimiento de estándares europeos de calidad, interoperabilidad y protección de privacidad. La colaboración con universidades, hospitalares, centros tecnológicos y empresas es esencial para compartir conocimiento y recursos. Las alianzas público-privadas deben regirse por principios de transparencia y beneficio mutuo, con términos claramente definidos y comunicados, asegurando que interés público prevalece y resultados revertan en beneficio de sociedad.

La participación ciudadana fortalece legitimidad y facilita recopilación de datos de calidad. Los pacientes deben ser informados y tener voz en diseño de proyectos. Cuando ciudadanos son informados claramente sobre objetivos, métodos e implicaciones, y tienen oportunidad de expresar opiniones y preocupaciones, se genera confianza esencial para aceptación y éxito de soluciones desarrolladas. La participación activa permite incorporar perspectivas y conocimientos experienciales, asegurando que soluciones respondan efectivamente a necesidades reales de usuarios finales mediante encuestas, grupos focales, incorporación de representantes en órganos de gobernanza, garantizando que voz de ciudadanos sea escuchada en todas las fases. Se impulsarán mecanismos concretos de participación como comités asesores de pacientes, procesos de co-diseño de proyectos y consultas públicas que orienten prioridades y criterios de evaluación.

9. Limitaciones y líneas de futuro

La investigación en IA aplicada a salud se desarrolla en entorno en constante transformación que demanda humildad epistémica. Reconocemos que cualquier posicionamiento debe concebirse como documento vivo sujeto a revisión continua. Sabemos parte del camino, ignoramos mucho sobre desafíos que enfrentaremos, y necesitamos aprender sistemáticamente a través de experiencia, evidencia emergente y debate interdisciplinar.

Persisten lagunas de conocimiento críticas. Aunque modelos de IA muestran resultados prometedores en estudios controlados, existe escasa evidencia sobre impacto real en práctica clínica cotidiana. Los sistemas deben evaluarse en entornos complejos, diversos y cambiantes que reflejen atención sanitaria real. Desconocemos en gran medida cómo estos sistemas influyen en

trabajo de investigadores y profesionales clínicos. La equidad constituye otro desafío clave: el sesgo algorítmico puede reproducir y amplificar desigualdades estructurales de maneras sutiles no siempre evidentes. Necesitamos metodologías más sofisticadas que permitan evaluar cómo diferentes definiciones de equidad se traducen en consecuencias prácticas y cómo equilibrar tensión entre maximizar precisión global y asegurar rendimiento equitativo.

La generación de datos sintéticos emerge como recurso prometedor pero plantea dudas técnicas y éticas que requieren investigación adicional sobre cuándo preservan propiedades estadísticas relevantes de datos reales, cómo garantizar que no permiten reidentificación, y cómo normativas de protección de datos se aplican a datos sintéticos. La investigación debe afrontar debate sobre autoría y autenticidad en resultados generados con IA. Surgen interrogantes sobre si IA es mero instrumento o si su contribución requiere reconocimiento explícito. La proliferación de deepfakes y datos manipulados amenaza integridad de evidencia científica. Los institutos deben definir políticas claras sobre uso de IA generativa, establecer estándares de transparencia, y reforzar mecanismos de validación de datos e imágenes.

El desarrollo de modelos multimodales abre nuevas oportunidades pero incrementa riesgos de privacidad porque pueden inferir información sensible de combinaciones de datos. La OMS advierte sobre necesidad de evaluaciones rigurosas de impacto y adopción de técnicas avanzadas como privacidad diferencial o aprendizaje federado. La complejidad inherente plantea limitaciones en explicabilidad, comprometiendo aceptación en investigación y práctica clínica.

Las líneas de futuro incluyen desarrollo de metodologías de auditoría y certificación que aporten confianza, exploración de sinergias con tecnologías emergentes como computación cuántica o neuromórfica, y evaluación del impacto ambiental de investigación en IA. El entrenamiento de modelos grandes puede consumir tanta energía como consumo anual de varios hogares. Adoptar prácticas de computación sostenible es responsabilidad ambiental y eficiencia económica. Los institutos deben desarrollar métricas de eficiencia energética, priorizar cuando sea posible algoritmos más eficientes, utilizar infraestructuras alimentadas por energías renovables, y considerar impacto ambiental como criterio en evaluación de proyectos. Entre estas líneas se priorizarán metodologías de auditoría y certificación de sistemas de IA en salud, así como la creación de una guía en línea inspirada en PROSPERO y PRISMA para el registro, seguimiento y revisión de proyectos de IA, y el desarrollo de un catálogo de activos de datos basado en el perfil de aplicación HealthDCAT-AP.

Este posicionamiento se compromete a revisiones periódicas, estructuradas y participativas que incorporen avances científicos, desarrollos regulatorios y cambios tecnológicos. Proponemos establecer ciclo de revisión bienal donde grupo de trabajo ad hoc analice vigencia de recomendaciones, identifique áreas que requieren actualización, y proponga modificaciones sometidas a consulta amplia. Solo desde actualización continua, con visión crítica pero constructiva y orientada al bien común, la investigación en IA en salud podrá consolidarse como motor de innovación responsable genuinamente centrada en las personas. Estas revisiones se plantean con una periodicidad de aproximadamente dos años, salvo circunstancias regulatorias o tecnológicas que aconsejen adelantar la actualización.

10. Conclusiones y llamada a la acción

La inteligencia artificial ofrece una oportunidad genuina y excepcional para mejorar la investigación sanitaria y, en última instancia, la salud de las personas. Esta oportunidad solo se materializará si

abordamos esta transformación con la seriedad, rigor y responsabilidad que demanda. Este documento establece un marco común construido mediante proceso deliberativo amplio que armoniza definiciones, identifica oportunidades y riesgos de manera equilibrada, recoge principios éticos compartidos, explica el marco regulador de manera accesible, y propone buenas prácticas concretas junto con recomendaciones accionables.

Los Institutos de Investigación Sanitaria del ISCIII reafirmamos nuestro compromiso institucional con una IA centrada en el paciente, basada en evidencia científica sólida, transparente, equitativa y responsable. Estos no son lemas abstractos sino compromisos operativos que se traducirán en políticas institucionales específicas, procedimientos de evaluación ética rigurosos, estructuras de apoyo técnico y normativo, e inversiones sostenidas en formación e infraestructuras. Invitamos a investigadores a adoptar las guías metodológicas específicas, a colaborar de forma genuinamente interdisciplinar, a incorporar consideraciones éticas desde el diseño inicial, y a comunicar hallazgos con transparencia. Invitamos a instituciones a crear estructuras de apoyo efectivas, a establecer políticas claras, a invertir sostenidamente, y a promover activamente igualdad de género y diversidad. Invitamos a agencias financieras a promover políticas y financiación estables, a incluir criterios de equidad e impacto social en evaluación, a requerir planes robustos de gestión de datos, y a apoyar desarrollo de infraestructuras compartidas. Invitamos a responsables políticos a desarrollar marcos regulatorios equilibrados, a invertir en infraestructuras digitales interoperables y seguras, a facilitar participación en iniciativas europeas, y a promover diálogo entre todos los actores. Invitamos a la sociedad a participar activamente en debate sobre cómo queremos que IA se desarrolle, a exigir transparencia, a aportar perspectivas sobre qué problemas son prioritarios, y a confiar en que ciencia desarrollada en instituciones públicas se guía por compromiso con bien común, aunque esta confianza debe ganarse continuamente mediante demostración de responsabilidad.

Solo mediante acción coordinada y sostenida de todos los actores relevantes, la inteligencia artificial podrá cumplir su promesa de contribuir significativamente al bienestar de las personas y a la innovación en salud. El camino no será sencillo ni lineal. Enfrentaremos desafíos técnicos que requerirán innovación metodológica, dilemas éticos que demandarán deliberación cuidadosa, y tensiones entre diferentes valores legítimos que necesitarán equilibrios prudentes. Pero si permanecemos fieles a los principios articulados en este documento, si mantenemos como guía constante el beneficio de los pacientes y de la sociedad, y si aprendemos humildemente tanto de nuestros éxitos como de nuestros errores, podremos avanzar con paso firme hacia un futuro donde la inteligencia artificial sirva verdaderamente a la salud y al bienestar de todas las personas.

11. Anexos

Anexo I: Glosario de términos y acrónimos

Anonimización: Proceso de transformación irreversible de datos personales de modo que impida la identificación del individuo, convirtiendo los datos en no personales bajo el RGPD.

Aprendizaje automático (Machine Learning): Conjunto de técnicas que permiten a los sistemas aprender de los datos y mejorar su rendimiento sin ser explícitamente programados para cada tarea específica.

Aprendizaje federado (Federated Learning): Técnica de aprendizaje automático donde un modelo se entrena utilizando datos distribuidos en múltiples sitios sin centralizar los datos, preservando la privacidad.

Aprendizaje profundo (Deep Learning): Subconjunto del aprendizaje automático basado en redes neuronales artificiales con múltiples capas que pueden representar información de manera jerárquica.

Big Data: Conjuntos de datos caracterizados por gran volumen, alta diversidad y velocidad de generación, cuyo análisis requiere infraestructuras y técnicas específicas.

CHART: Chatbot Assessment Reporting Tool, guía para reportar estudios de sistemas conversacionales basados en IA.

CONSORT-AI: Extensión de las guías CONSORT para reportar ensayos clínicos que involucran intervenciones basadas en IA.

DECIDE-AI: Guía para evaluar sistemas de IA en etapas tempranas de uso clínico.

EHDS: European Health Data Space (Espacio Europeo de Datos Sanitarios), iniciativa de la UE para facilitar el acceso y uso secundario de datos de salud.

IA explicable (XAI - Explainable AI): Técnicas que permiten comprender cómo los sistemas de IA llegan a sus conclusiones o recomendaciones.

IA generativa: Modelos de IA capaces de producir contenido novedoso como textos, imágenes o datos sintéticos a partir del aprendizaje de patrones en datos de entrenamiento.

Inteligencia Artificial (IA): Conjunto de técnicas computacionales que permiten a los sistemas realizar tareas que típicamente requieren razonamiento, aprendizaje, percepción o interacción adaptativa.

Interoperabilidad: Capacidad de diferentes sistemas de intercambiar datos y utilizarlos de manera efectiva mediante estándares técnicos y semánticos comunes.

LOPDGDD: Ley Orgánica de Protección de Datos y Garantía de Derechos Digitales (España).

Modelo multimodal: Sistema de IA que integra y procesa diferentes tipos de datos (texto, imágenes, audio, datos estructurados) simultáneamente.

Privacidad diferencial: Técnica matemática que proporciona garantías cuantificables de privacidad al añadir ruido controlado a los datos o resultados.

RGPD: Reglamento General de Protección de Datos de la Unión Europea.

Sesgo algorítmico: Tendencia sistemática de un algoritmo a producir resultados injustos o discriminatorios para ciertos grupos.

Seudonimización: Proceso de sustitución de identificadores directos por códigos, reduciendo el riesgo de identificación pero manteniendo la utilidad analítica de los datos.

SPIRIT-AI: Guía para elaborar protocolos de ensayos clínicos que involucran intervenciones basadas en IA.

STARD-AI: Guía para reportar estudios de precisión diagnóstica que utilizan IA.

Supervisión humana: Principio que requiere que sistemas de IA, especialmente los de alto riesgo, mantengan la intervención y control de profesionales capacitados.

TRIPOD-AI: Guía para reportar estudios de desarrollo y validación de modelos predictivos basados en IA.

Validación externa: Evaluación del rendimiento de un modelo en datos completamente independientes de una población o contexto diferente al utilizado para su desarrollo.

Anexo II: Tabla resumen de responsabilidades por actor

Actor	Responsabilidades principales	Acciones clave
Investigadores individuales	Formación continua en IA, ética y protección de datos; Diseño riguroso de estudios; Gestión responsable de datos; Comunicación transparente de resultados	Completar formación específica en IA; Aplicar guías SPIRIT-AI, TRIPOD-AI, CONSORT-AI; Evaluar modelos con criterios de equidad; Publicar en acceso abierto
Grupos de investigación	Colaboración interdisciplinaria; Participación de pacientes; Evaluación de impacto ético y social; Reproducibilidad	Formar equipos multidisciplinarios; Incorporar representantes de pacientes; Realizar análisis de impacto; Compartir datos y código cuando sea posible
Institutos de investigación	Políticas institucionales; Unidades de apoyo especializadas; Infraestructuras seguras; Formación de personal; Igualdad y diversidad	Aprobar políticas sobre IA; Crear unidades de ciencia de datos y protección de datos; Invertir en infraestructuras; Implementar planes de igualdad
Comités de ética	Evaluación ética específica para IA; Supervisión continua; Formación especializada	Desarrollar criterios de evaluación para IA; Realizar seguimiento de proyectos; Incorporar expertos en IA y representantes de pacientes
Direcciones científicas	Estrategia institucional; Asignación de recursos; Promoción de buenas prácticas; Rendición de cuentas	Priorizar inversiones en IA; Reconocer contribuciones interdisciplinares; Reportar actividades de IA; Participar en redes nacionales e internacionales
Agencias financiadoras	Políticas de financiación alineadas con principios éticos; Promoción de estándares; Requisitos de transparencia	Incluir criterios de equidad en evaluación; Requerir planes de gestión de datos; Financiar infraestructuras compartidas; Promover ciencia abierta
Autoridades sanitarias	Marco regulatorio apropiado; Coordinación interinstitucional; Protección de derechos fundamentales	Desarrollar normativas equilibradas; Facilitar participación en EHDS; Armonizar requisitos; Supervisar cumplimiento
Pacientes y ciudadanos	Participación informada; Aportación de perspectivas; Ejercicio de derechos	Participar en diseño de proyectos; Expresar prioridades y preocupaciones; Solicitar información sobre uso de datos

Anexo III: Lista de verificación para proyectos de IA en investigación sanitaria

Fase de diseño

- [] ¿Se ha justificado el uso de IA frente a métodos más simples?
- [] ¿El equipo incluye expertise en IA, clínica, ética y protección de datos?
- [] ¿Se han incluido representantes de pacientes en la planificación?
- [] ¿Se ha realizado un análisis de riesgos éticos y de privacidad?
- [] ¿El protocolo sigue las guías SPIRIT-AI cuando corresponda?
- [] ¿Se ha planificado la gestión de datos desde el inicio?

Fase de datos

- [] ¿Los datos son representativos de la población diana?
- [] ¿Se ha verificado la calidad de los datos?
- [] ¿Se han implementado medidas de seudonimización o anonimización?
- [] ¿Se ha obtenido consentimiento informado apropiado?
- [] ¿Existe documentación completa del origen y transformaciones de los datos?
- [] ¿Se cumplen los requisitos del RGPD y LOPDGDD?

Fase de desarrollo

- [] ¿Se han considerado estrategias de mitigación de sesgos?
- [] ¿Se utilizan conjuntos separados para entrenamiento, validación y prueba?
- [] ¿Se han definido métricas apropiadas incluyendo medidas de equidad?
- [] ¿Se documenta exhaustivamente el desarrollo del modelo?
- [] ¿Se ha realizado validación externa cuando sea posible?
- [] ¿Se evaluó el rendimiento desagregado por subgrupos relevantes?

Fase de implementación

- [] ¿Existe un plan de supervisión humana efectiva?
- [] ¿Se ha diseñado monitorización continua del rendimiento?
- [] ¿Hay procedimientos claros para gestionar incidentes?
- [] ¿Se ha formado adecuadamente a los usuarios del sistema?
- [] ¿Existe un comité de supervisión para la implementación?

Fase de comunicación

- [] ¿Se ha registrado el protocolo en repositorio público?
- [] ¿El manuscrito sigue las guías apropiadas (CONSORT-AI, TRIPOD-AI, etc.)?
- [] ¿Se reportan limitaciones y resultados negativos honestamente?
- [] ¿Se facilita acceso a código, modelos o datos cuando sea posible?
- [] ¿La autoría refleja todas las contribuciones sustanciales?
- [] ¿La comunicación pública es clara sobre beneficios y limitaciones?

12. Bibliografía

- Al Manir S, Levinson MA, Niestroy J, Churas C, Parker JA, Clark T. FAIRSCAPE: An Evolving AI-readiness Framework for Biomedical Research. *bioRxiv*. 2025;2024.12.23.629818. doi:10.1101/2024.12.23.629818
- Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. 2024;385:e078378. doi:10.1136/bmj-2023-078378
- EIT Health. Implementation of European Health Data Space in Spain: is it really feasible? 2024. Disponible en: https://www.eit.europa.eu/sites/default/files/2024-12/Report%20EHDS_Spain_ENG.pdf
- Gao S, Fang A, Huang Y, Giunchiglia V, Noori A, Schwarz JR, et al. Empowering biomedical discovery with AI agents. *Cell*. 2024;187(22):6125-6151. doi:10.1016/j.cell.2024.09.022
- Ibrahim H, Liu X, Rivera SC, et al. Reporting guidelines for clinical trials of artificial intelligence interventions: the SPIRIT-AI and CONSORT-AI guidelines. *Trials*. 2021;22:11. doi:10.1186/s13063-020-04951-6
- Ioannidis JPA, Collins TA, Baas J. Evolving patterns of extreme publishing behavior across science. *Scientometrics*. 2024;129:5783-5796. doi:10.1007/s11192-024-05117-w
- The CHART Collaborative. Reporting guideline for chatbot health advice studies: the Chatbot Assessment Reporting Tool (CHART) statement. *BJS*. 2025;112(8):znaf142. doi:10.1093/bjs/znaf142
- UNESCO. Guía para la utilización de la IA Generativa en educación e investigación. 2023. Disponible en: <https://unesdoc.unesco.org/ark:/48223/pf0000389227>
- Universidad Complutense de Madrid. Guía básica para la protección de datos en la investigación. Disponible en: <https://www.ucm.es/dpd/file/guía-básica-protección-datos-en-investigación>
- White House. America's AI Action Plan. 2025. Disponible en: <https://www.whitehouse.gov/articles/2025/07/white-house-unveils-americas-ai-action-plan/>
- World Health Organization. Big data and artificial intelligence. Disponible en: <https://www.who.int/teams/health-ethics-governance/emerging-technologies/big-data-and-artificial-intelligence>

13. Normativa consultada

- Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos (Reglamento General de Protección de Datos).
- Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial).
- Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.
- Ley 14/2007, de 3 de julio, de Investigación biomédica.
- Ley 17/2022, de 5 de septiembre, por la que se modifica la Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación.

14. Figuras y Tablas

Figura 1

Principios Éticos Fundamentales para la IA en Investigación Sanitaria

Beneficencia y No Maleficencia

Maximizar el beneficio para la salud y minimizar el daño potencial

Equidad y Justicia

Evitar sesgos y garantizar beneficios para todas las personas sin discriminación

Transparencia y Explicabilidad

Modelos comprensibles para profesionales y pacientes, con motivos claros de decisiones

Privacidad y Protección de Datos

Salvaguardar la identidad, cumplir con el RGPD, adoptar seudonimización y anonimización

Responsabilidad y Rendición de Cuentas

Asumir responsabilidad de resultados, con capacidad de auditar modelos

Supervisión Humana Cualificada

Mantener intervención y control de profesionales capacitados en decisiones críticas

Estos principios se alinean con los marcos éticos de la OMS y UNESCO, enfatizando la protección de la autonomía, la seguridad, la transparencia, la equidad, la sostenibilidad y la supervisión humana.

Figura 2

Hoja de Ruta para la Implementación de Inteligencia Artificial en Investigación Sanitaria

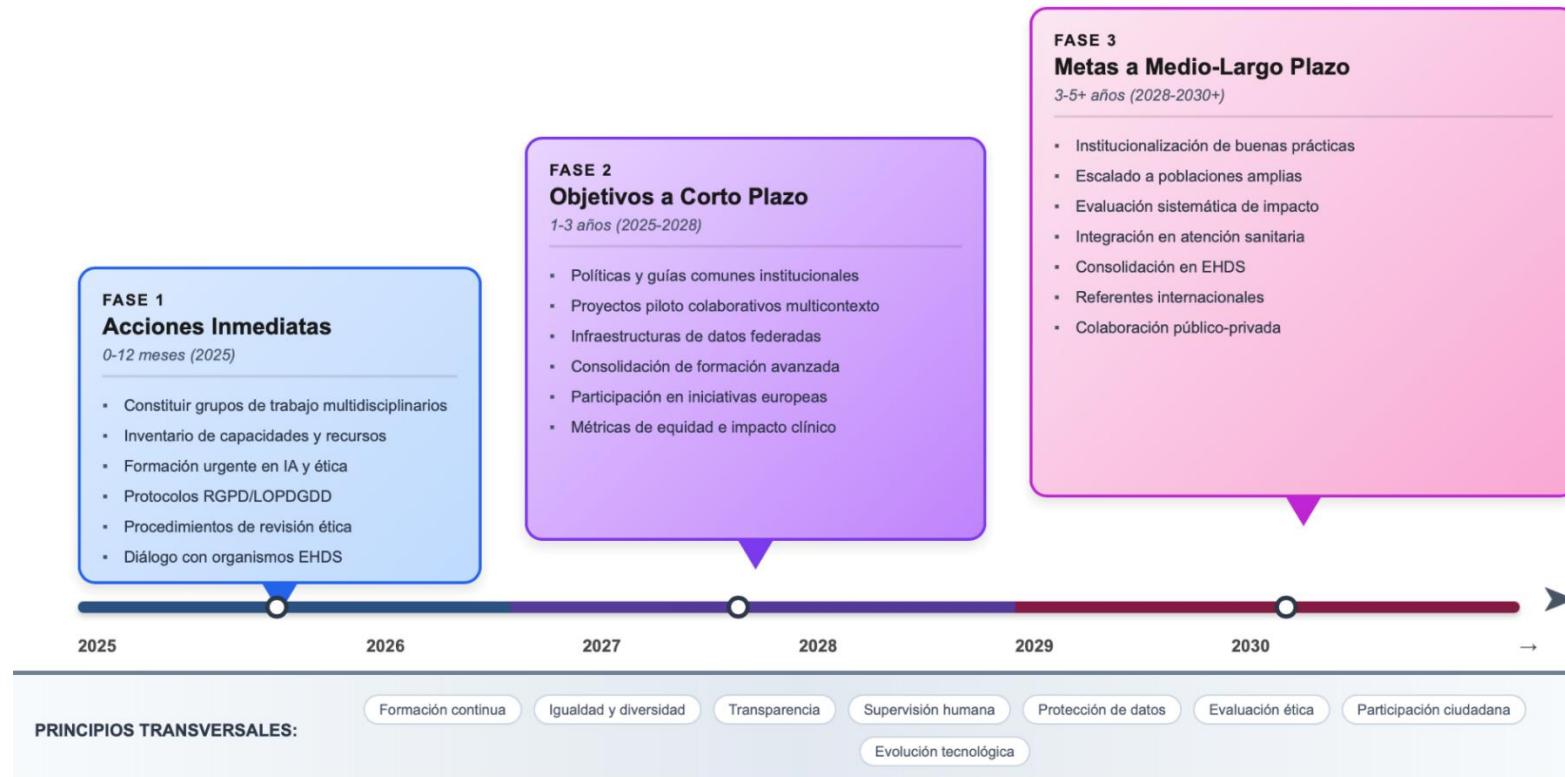


Figura 2. Representación visual de la implementación progresiva de inteligencia artificial en investigación sanitaria. Las fases se muestran como bloques ascendentes que reflejan el aumento en complejidad y alcance, sustentadas por principios transversales que permanecen constantes durante todo el proceso.

Tabla 1. Definiciones operativas de conceptos fundamentales

Concepto	Definición operativa	Aplicación en investigación sanitaria
TECNOLOGÍAS DE IA		
Inteligencia Artificial	Conjunto de técnicas que permiten a sistemas realizar tareas que requieren razonamiento, aprendizaje, percepción o interacción adaptativa	Ánálisis de imágenes médicas, predicción de riesgos, apoyo a decisiones clínicas
Machine Learning (Aprendizaje automático)	Algoritmos que encuentran patrones en datos y mejoran con la experiencia sin programación explícita	Estratificación de pacientes, predicción de respuesta a tratamientos
Deep Learning (Aprendizaje profundo)	Redes neuronales con múltiples capas que representan información jerárquicamente	Diagnóstico por imagen, análisis genómico, procesamiento de historias clínicas
IA Generativa	Modelos que producen contenido novedoso (textos, imágenes) a partir de datos de entrenamiento	Generación de datos sintéticos, asistencia en redacción científica, simulaciones
Big Data	Conjuntos de datos caracterizados por gran volumen, diversidad y velocidad de generación	Integración de datos multi-ómicos, registros clínicos poblacionales, vigilancia epidemiológica
Sistemas Predictivos	Modelos que estiman probabilidad de eventos futuros a partir de datos históricos	Predicción de desenlaces clínicos, modelización de brotes, pronóstico de enfermedades
MODALIDADES DE INVESTIGACIÓN SANITARIA		
Investigación Básica	Profundiza en mecanismos biológicos fundamentales a nivel molecular, celular o sistémico	Identificación de dianas terapéuticas, comprensión de patogénesis
Investigación Clínica	Estudios con personas o muestras biológicas humanas para evaluar intervenciones	Ensayos clínicos, estudios observacionales, validación de biomarcadores
Investigación Traslacional	Traslada hallazgos de investigación básica a aplicaciones clínicas	Desarrollo de nuevas terapias, medicina de precisión
Investigación Epidemiológica	Estudia distribución y determinantes de enfermedades en poblaciones	Vigilancia de salud pública, identificación de factores de riesgo
Investigación de Servicios de Salud	Analiza organización y eficiencia de sistemas sanitarios	Optimización de recursos, mejora de procesos asistenciales

Nota: La IA puede actuar en todas las modalidades de investigación sanitaria, aunque con consideraciones metodológicas y éticas específicas en cada caso. La comprensión compartida de estos conceptos facilita la comunicación efectiva entre clínicos, científicos de datos, bioestadísticos, especialistas en ética, juristas y otros profesionales que deben colaborar en proyectos multidisciplinarios.

Tabla 2. Guía de buenas prácticas en el ciclo de vida de la investigación con IA

Fase	Principios clave y acciones	Guías específicas y consideraciones
1. Planificación y Diseño	<ul style="list-style-type: none"> Ética desde el diseño <ul style="list-style-type: none"> Justificar el uso de IA frente a alternativas Identificar riesgos potenciales tempranamente Incluir análisis de impacto ético y social Multidisciplinariedad <ul style="list-style-type: none"> Equipos con científicos de datos, clínicos, epidemiólogos, especialistas en ética y pacientes 	<p>SPIRIT-AI Protocolos de ensayos</p> <ul style="list-style-type: none"> Descripción detallada del algoritmo Información sobre datos de entrenamiento Estrategias de validación Plan de actualización del modelo Procedimientos de supervisión humana
2. Recolección y Gestión de Datos	<ul style="list-style-type: none"> Calidad y representatividad <ul style="list-style-type: none"> Diversidad en edad, género, etnia, nivel socioeconómico Minimizar errores, valores perdidos, inconsistencias Estandarizar procedimientos de recolección Cumplimiento normativo <ul style="list-style-type: none"> Minimización de datos (solo lo necesario) Seudonimización temprana Consentimiento informado específico para IA 	<p>Requisitos RGPD/LOPDGDD:</p> <ul style="list-style-type: none"> Documentar origen y contexto de datos Registrar transformaciones aplicadas Implementar medidas de seguridad multicapa Planificar preservación o destrucción segura <p>Interoperabilidad:</p> <ul style="list-style-type: none"> Estándares: HL7 FHIR, DICOM, SNOMED CT, LOINC
3. Desarrollo y Validación	<ul style="list-style-type: none"> Mitigación de sesgos <ul style="list-style-type: none"> Repesado de muestras para compensar desequilibrios Técnicas de aprendizaje justo (fairness-aware ML) Calibración por subgrupos Validación rigurosa <ul style="list-style-type: none"> Separación: entrenamiento/validación/prueba Validación externa en datos independientes Evaluación desagregada por subgrupos 	<p>TRIPOD+AI Modelos predictivos (27 elementos)</p> <ul style="list-style-type: none"> Datos fuentes y variables predictoras Arquitectura del algoritmo e hiperparámetros Procedimientos de preprocessamiento Métricas de rendimiento y equidad Validación interna y externa <p>Métricas de equidad: Paridad demográfica, igualdad de oportunidades, paridad predictiva</p>
4. Implementación Experimental	<ul style="list-style-type: none"> Supervisión continua <ul style="list-style-type: none"> Comenzar en entornos controlados Comités de supervisión con informes periódicos Monitorización del impacto y derechos Supervisión humana efectiva <ul style="list-style-type: none"> Información comprensible sobre predicciones Alertas ante alta incertidumbre Intervención humana fácil y rápida 	<p>DECIDE-AI Evaluación temprana</p> <ul style="list-style-type: none"> Evaluación en etapas tempranas Reporte transparente para replicabilidad Capacidad de retirar o ajustar el modelo <p>Gestión de incidentes:</p> <ul style="list-style-type: none"> Procedimientos claros de reporte Investigación de causas Acciones correctivas y preventivas
5. Publicación y Difusión	<ul style="list-style-type: none"> Transparencia y reproducibilidad <ul style="list-style-type: none"> Registro previo de protocolos Descripción detallada de algoritmos, datos, métricas Discusión honesta de limitaciones Facilitar acceso a código y modelos Autoría responsable <ul style="list-style-type: none"> Reconocer contribuciones técnicas sustanciales Incluir participación de pacientes 	<p>CONSORT-AI Resultados de ensayos</p> <p>STARD-AI Precisión diagnóstica</p> <p>CHART Sistemas conversacionales</p> <ul style="list-style-type: none"> Acceso abierto a publicaciones Compartir datos sintéticos, acceso controlado, enclaves seguros Repositorios: GitHub (código), Hugging Face (modelos), Zenodo (datos) <p>Comunicación: Clara y honesta sobre beneficios y limitaciones</p>

Principio transversal: La colaboración entre científicos de datos y clínicos debe mantenerse a lo largo de todo el ciclo de vida. Los clínicos aportan conocimiento del dominio y criterio sobre utilidad clínica; los científicos de datos contribuyen comprensión de capacidades y limitaciones algorítmicas, técnicas avanzadas de validación, e implementación eficiente.

Subproyecto 2: Encuesta Nacional: Mapa sobre la situación de la investigación con Big Data e IA en los IISA

Versión: 1.2 · **Fecha:** 03/12/2025

Grupo de trabajo: GdT5

Coordinadores: VHIR · FJD · IISGM

Control de versiones

VERSIÓN	FECHA	CAMBIOS PRINCIPALES
1.0	15/09/2025	Publicación versión inicial
1.1	30/09/2025	Revisión GdT5-IA
1.2	02/12/2025	Revisión GT-IA ISCIII

Objetivo: Encuesta dirigida a conocer el estado actual de la capacitación para realizar investigación con Big Data e IA en los IISA, identificar obstáculos y áreas de mejora, y analizar la evolución temporal de dichas capacitaciones en nuestro país.

A quién va dirigida: Institutos de Investigación Sanitaria Acreditados de España.

Qué aporta:

1. Un **mapa de capacidades actuales para la investigación con Big Data e IA** en los IISA de España, con oportunidad para identificar barreras y obstáculos comunes y modelos de éxito exportables a otros centros.
2. Un **seguimiento de la evolución** de capacitaciones a lo largo del tiempo.

Estado del documento: en revisión final, versión semidefinitiva

Fecha de aprobación: [Pendiente]

Próxima revisión programada: [12 meses desde aprobación]

Subproyecto 2: Encuesta Nacional: Mapa sobre la situación de la investigación con Big Data e IA en los IISA

Contacto

Correo electrónico para dudas y aclaraciones (opcional)
(texto libre, formato: xxxx@xxx.xxx)

_____ @ _____ . _____

1. Información General del Centro

1.1. Nombre del IISA:

(texto libre)

1.2. Comunidad Autónoma:

(seleccionar una)

- Andalucía
- Aragón
- Asturias
- Baleares
- Canarias
- Cantabria
- Castilla y León
- Castilla-La Mancha
- Cataluña
- Comunidad Valenciana
- Extremadura
- Galicia
- La Rioja
- Madrid
- Murcia
- Navarra
- País Vasco
- Ceuta
- Melilla

1.3. Centros vinculados al IISA:

(opción multiple, en caso de selección centro, indicaqr nombre en formato texto libre)

- Hospital/-es: _____
- Centro/-s de investigación: _____

- Universidad/-es: _____
- Otros: _____

1.4. Nº aproximado de investigadores vinculados:

(respuesta única)

- < 100
- 100-249
- 250-499
- 500-999
- ≥ 1000 (especificar) _____

1.5. Nº personal de gestión y administración (soporte a la investigación, incluyendo servicios científico-técnicos):

(respuesta única)

- < 20
- 20-49
- 50-99
- 100-199
- ≥ 200 (especificar) _____

2. Recursos Humanos y Capacitación

2.1. ¿Dispone su centro de las siguientes unidades o servicios?

(respuesta múltiple, si selecciona otros, especificar en formato texto libre)

- Bioestadística
- Bioinformática / Big data
- Oficina de gestión de datos clínicos (data office) sin servicio de análisis de datos
- Oficina de IA aplicada a la investigación
- Otros (especificar) _____

2.2. ¿Dispone su centro de profesionales con estos perfiles? ¿De cuántos profesionales dispone en cada caso?

(respuesta múltiple, indicar cuántos. Considerar solo el personal perteneciente a estructura, no los pertenecientes a grupos de investigación). En otros: especificar perfil y cantidad

- Bioestadística; _____
- Bioinformática; _____
- Ingeniería de ciencia de datos; _____
- Ingeniería Inteligencia Artificial; _____
- Expertos en protección de datos de carácter personal; _____
- Expertos en regulatoria de IA; _____
- Otros (especificar); _____

2.3. ¿Dispone su centro de programas de formación interna en IA/Big Data?

(respuesta única)

- Sí, estructurada y regular (especificar) _____
- Sí, ocasional (especificar) _____
- En desarrollo
- No

3. Infraestructura Tecnológica

3.1. Infraestructura de almacenamiento de datos

3.1.1. ¿Cómo almacena el IIS los datos?

(selección múltiple)

- Infraestructura física on premise-centralizada en el propio IIS
- Infraestructura física on premise-centralizada en el centro hospitalario
- Infraestructura física on premise-no centralizada, en los grupos de investigación.
- Infraestructura en la nube
- Otros (especificar) _____

En caso de respuesta múltiple, describa el reparto:

3.1.2. ¿Dispone el IIS de una estrategia para repercutir los costes del uso del almacenamiento al grupo de investigación usuario?

(respuesta única)

- Sí, definida
- Sí, en desarrollo
- No
- No aplica / no lo sabe

3.1.3. En el caso de la infraestructura de almacenamiento en la nube, seleccionar el modelo de nube utilizado:

(respuesta múltiple)

- Nube pública (Azure, AWS, Google, etc.)
- Nube gestionada por una institución externa con uso privado
- Nube privada gestionada por el propio IIS / centro hospitalario
- Otros (especificar) _____

Nube gestionada por institución externa: infraestructura de terceros (p.ej., BSC, CSUC, RedIRIS). Servicio compartido (multi-tenant), acceso por proyectos/cuotas/convocatorias, escalado según oferta. Coste por uso/acuerdos marco/subvención. Cumplimiento/seguridad: políticas del proveedor; el IIS encaja su DPIA, ENS/ISO en esos marcos.

Nube privada gestionada por IIS/hospital: infraestructura propia (on-premise), control total (HW, red, contenedores, colas, políticas). Modelo single-tenant o multi-tenant interno, prioridades/cuotas propias. Con Costes CAPEX+OPEX. Cumplimiento/seguridad: controles propios alineados a ENS/ISO, integración directa con redes clínicas, identidades corporativas y residencia del dato local.

3.1.4. Modelo de contratación de los datos en la nube:

(respuesta múltiple)

- Pago por uso
- Financiación institucional / subvencionado
- Otro (especificar) _____

3.1.5. Capacidades técnicas de almacenamiento (en caso de más de una, especificar en cada lugar):

3.1.5.1. Almacenamiento alta velocidad centralizada en el IIS:

(respuesta única)

- < 50 TB
- 50–200 TB
- 200–500 TB
- 500 TB – 1PB
- > 1PB (especificar) _____
- No aplica

3.1.5.2. Almacenamiento alta velocidad centralizada en el centro hospitalario:

(respuesta única)

- < 50 TB
- 50–200 TB
- 200–500 TB
- 500 TB – 1PB
- > 1PB (especificar) _____
- No aplica

3.1.5.3. Almacenamiento total alta velocidad no-centralizada en los grupos de investigación:

(respuesta única)

- < 50 TB
- 50–200 TB
- 200–500 TB
- 500 TB – 1PB
- 1PB (especificar) _____
- No aplica

3.1.5.4. Almacenamiento alta velocidad en la nube:

(respuesta única)

- < 50 TB
- 50–200 TB
- 200–500 TB

- 500 TB – 1PB
- > 1PB (especificar) _____
- No aplica

3.1.5.5. Sistema almacenamiento-centralizado, con redundancia y seguridad:

(respuesta única)

- No
- Si, RAID (especificar Nivel RAID 1 a 6): _____
- Si, sin RAID (especificar sistema de redundancia empleado) _____

3.1.5.6. Sistema centralizado de copias de seguridad en cinta:

(respuesta única)

- Si , Cinta (LTO-Linear Tape-Open u otras)
- No
- Otro sistema de backup (especificar: disco, nube, etc): _____

3.1.5.7. Periodo de retención (p.ej. 30 días, 3 meses, 5 años):

(texto libre)

3.1.6. Administración de almacenamiento (quién administra -instala, configura, gestiona, gobierna, realiza copias de seguridad- los sistemas de almacenamiento):

(respuesta única)

- Se administra centralizadamente por el departamento de IT o similares
- Cada grupo de investigación administra sus sistemas de almacenamiento
- Mixto (especificar) _____

3.2. Infraestructura de computación y supercomputación

3.2.1. ¿Dónde se encuentra la infraestructura de computación?

(respuesta múltiple)

- Infraestructura física on premise en el IIS
- Infraestructura física on premise-local en grupos de investigación
- Infraestructura externa (en la nube)
- No dispone de infraestructura propia

En caso de respuesta múltiple, detallar _____

3.2.2. En el caso de la infraestructura on premise, ¿ dónde se encuentran los servidores de computación y supercomputación?

(respuesta múltiple)

- En el Centro de Procesamiento Datos centralizado y ubicado en el IIS
- En el CPD del hospital
- Cada grupo de investigación tiene / gestiona sus servidores

En caso de selección múltiple, detallar _____

3.2.3. ¿Dispone el IIS de una estrategia para repercutir los costes del uso de los servicios de computación al grupo de investigación usuario?

(respuesta única)

- Sí, definida
- Sí, en desarrollo
- No
- No aplica

3.2.4. En el caso de la infraestructura de computación en la nube, seleccionar el modelo de nube utilizado:

(respuesta múltiple)

- Nube pública (Azure, AWS, Google, etc.)
- Nube gestionada por una institución externa (BSC, CSUC, REDIRIS...)
- Nube privada gestionada por el propio IIS / hospital
- Otros (especificar) _____

Nube gestionada por institución externa: infraestructura de terceros (p.ej., BSC, CSUC, RedIRIS). Servicio compartido (multi-tenant), acceso por proyectos/cuotas/convocatorias, escalado según oferta. Coste por uso/acuerdos marco/subvención. Cumplimiento/seguridad: políticas del proveedor; el IIS encaja su DPIA, ENS/ISO en esos marcos.

Nube privada gestionada por IIS/hospital: infraestructura propia (on-prem/housing), control total (HW, red, contenedores, colas, políticas). Modelo single-tenant o multi-tenant interno, prioridades/cuotas propias. Costes CAPEX+OPEX. Cumplimiento/seguridad: controles propios alineados a ENS/ISO, integración con redes clínicas e identidades corporativas.

3.2.5. Modelo de contratación de los servicios de computación en la nube:

(respuesta múltiple)

- Pago por uso
- Financiación institucional / subvencionado
- Contrato marco con proveedor público (ej. RedIRIS, CSUC)
- Otro (especificar) _____

3.2.6. Capacidades técnicas de computación centrada en CPD IIS:

(respuesta única)

- Básica (< 200 cores/núcleos)
- Media (200–1000 cores/núcleos)
- Alta (> 1000 cores) (especificar/núcleos) _____
- No aplica

3.2.7. Capacidad de supercomputación (GPU) centrada en CPD IIS:

(respuesta única)

- Nula
- Baja (1–10 GPUs)
- Media (11–50 GPUs)
- Alta (> 50 GPUs) (especificar número GPUs) _____

- No aplica

3.2.8. Capacidades técnicas de computación centrada en CPD Hospital:

(respuesta única)

- Básica (< 200 cores/nucleos)
- Media (200–1000 cores/nucleos)
- Alta (> 1000 cores) (especificar/nucleos) _____
- No aplica

3.2.9. Capacidad de supercomputación (GPU) centrada en CPD Hospital:

(respuesta única)

- Nula
- Baja (1–10 GPUs)
- Media (11–50 GPUs)
- Alta (> 50 GPUs) (especificar número GPUs) _____
- No aplica

3.2.10. Capacidades técnicas de computación no centralizada, en servidores locales gestionadas por cada grupo:

(respuesta única)

- Básica (< 200 cores/nucleos)
- Media (200–1000 cores/nucleos)
- Alta (> 1000 cores) (especificar/nucleos) _____
- No aplica

3.2.11. Capacidad de supercomputación (GPU) no centralizada, en servidores locales gestionados por cada grupo:

(respuesta única)

- Nula
- Baja (1–10 GPUs)
- Media (11–50 GPUs)
- Alta (> 50 GPUs) (especificar número GPUs) _____
- No aplica

3.2.12. Capacidades técnicas de computación externalizada (CSUC,BSC,...) :

(respuesta única)

- Básica (< 200 cores/nucleos)
- Media (200–1000 cores/nucleos)
- Alta (> 1000 cores) (especificar/nucleos) _____
- No aplica

3.2.13. Capacidad de supercomputación (GPU) externalizada (CSUC,BSC,RedIris,...):

(respuesta única)

- Nula
- Baja (1–10 GPUs)

- Media (11–50 GPUs)
- Alta (> 50 GPUs) (especificar número GPUs) _____
- No aplica

3.2.14. ¿Quién administra (instala, configura, gestiona, gobierna, realiza copias de seguridad,...) la infraestructura local de los grupos?

(respuesta única)

- Departamento de IT
- El propio grupo de investigación
- Otros (especificar) _____

3.2.15. ¿Quién administra los sistemas de computación y almacenaje centralizados?

(respuesta única)

- Departamento de IT
- Otros (especificar) _____

3.3. Estrategia y gestión de la ciberseguridad en el IIS

3.3.1. ¿Dispone el IIS de una política institucional de ciberseguridad que incluya procedimientos, recursos y formación específica para la protección de datos en investigación biomédica?

(respuesta única)

- Sí, existe una estrategia formal y consolidada
- Si, está en desarrollo
- No
- No existe estrategia formal, pero se aplican medidas puntuales (especificar)

3.3.2. En caso afirmativo, ¿qué ámbitos están cubiertos en la estrategia de ciberseguridad?

(respuesta múltiple)

- Planes de respuesta ante incidentes de seguridad
- Protocolos de continuidad de negocio y recuperación ante desastres
- Formación periódica al personal investigador y técnico
- Auditorías internas/externas de seguridad
- Certificaciones (ISO 27001, ENS u otras)
- Otros (especificar) _____

4. Gobernanza y modelo de gestión

4.1. ¿Dispone el IIS de una política de gobernanza de datos clínicos definida?

(respuesta única)

- Sí, aprobada y publicada
- Sí, existente pero no publicada
- Sí, en desarrollo
- No

4.2. En caso afirmativo, ¿está esta política consensuada con el hospital de referencia?

(respuesta única)

- Sí, aprobada conjuntamente
- En proceso de consenso
- No

4.3. Tecnologías empleadas para gestión de datos:

(respuesta múltiple)

- RedCAP
- Omibio
- OpenClinica
- Castor EDC
- Medidata Rave
- OpenEDC
- Plataforma propia interna
- Otros (especificar) _____

4.4. Tecnologías de orquestación / contenedores:

(respuesta múltiple)

- Kubernetes
- Docker
- Singularity / Apptainer
- Otros (especificar)

4.5. ¿Cómo accede el IIS a los datos clínicos del hospital?

(respuesta única)

- No accede
- Acceso in situ únicamente dentro del recinto del hospital
- Acceso remoto mediante VDI (Virtual Desktop Infrastructure), VPN (Virtual Private Network), VLAN (Virtual Local Area Network)/ escritorio seguro
- Otros (especificar) _____

4.6. Modelo de compartición de datos entre el hospital y el IIS:

(respuesta única)

- Pseudoanonymizados
- Anonimizados

- En crudo (sin anonimizar)
- No se comparten

4.7. Formas de compartición de datos clínicos con otros centros de investigación

(respuesta múltiple)

- Anonimización
- Pseudoanonymización
- Plataformas federadas (los datos no salen del centro)
- Generación y compartición de datos sintéticos
- Acuerdos de transferencia de datos (DTA)
- Plataformas europeas (EGA, EHDS, etc.)
- Otros (especificar) _____

5. Ética y regulatoria

5.1. ¿Se ha previsto alguna política/planificación relativa a la futura aplicación del Reglamento europeo sobre IA?

(respuesta múltiple)

- Sí
- No
- En proceso

5.2. ¿Dispone su IIS de certificación ENS (Esquema Nacional de Seguridad)?

(respuesta única)

- Sí
- En proceso
- No

5.2.1. En caso de disponer de certificación ENS, nivel de seguridad de la certificación ENS:

(respuesta única)

- Bajo
- Medio
- Alto

5.2.2. En caso de disponer de certificación ENS, tipología de datos bajo la certificación ENS:

(respuesta múltiple)

- Datos de carácter general / administrativos (Ejemplo: información de contacto, datos de proyectos, facturación, inventario de equipos...)

- Datos personales no sensibles (Ejemplo: nombre y apellidos de investigadores, curricular, datos profesionales.)
 - Datos personales sensibles (categorías especiales del RGPD) (Ejemplo: datos de salud de pacientes, historia clínica, datos genómicos, biomarcadores.)
 - Datos de investigación biomédica agregados / anonimizados (no reversible) (Ejemplo: cohortes estadísticas anonimizadas, big data pseudonimizado para IA.)
 - Datos de investigación biomédica agregados pseudoanonimizados (reversible) (Ejemplo: cohortes estadísticas anonimizadas, big data pseudonimizado para IA.)
 - Datos críticos para continuidad del servicio (Ejemplo: sistemas de supercomputación, CPDs, infraestructuras que soportan ensayos clínicos o proyectos europeos.)
 - Otros
-

5.3. ¿Dispone su IIS de certificación ISO 27001?

(respuesta única)

- Sí
- En proceso
- No

5.4. ¿Dispone su centro de un Comité de Ética de Datos?

(respuesta única)

- Sí
- En proceso
- No

5.5. En caso afirmativo, el Comité de Ética de Datos es:

(respuesta única)

- Propio (dentro del CEIm)
- Independiente
- Compartido

5.6. ¿Se ha desarrollado una estrategia interna para integrar los principios éticos en los Proyectos de IA?:

(respuesta única)

- Sí
- En proceso
- No

5.7. Adherencia a la normativa ética y regulación aplicable:

(respuesta múltiple)

- Declaración de Helsinki
- RGPD
- EU Artificial Intelligence Act
- Reglamento europeo sobre Productos Sanitarios

- Beneficencia / no maleficencia
- Equidad / ausencia de sesgos
- Transparencia / explicabilidad
- Privacidad / protección de datos
- Supervisión humana
- Ninguno formalizado

5.8. ¿Cuenta su CEIm con expertise específico para evaluación de proyectos sobre IA?

(respuesta única)

- Si
- No

5.9. Las bases de datos utilizadas para el desarrollo de IA (entrenamiento, validación, prueba), ¿cumplen con la normativa de protección de datos y su utilización?

(respuesta única)

- Si
- No

5.10. En el desarrollo de sistemas de IA que tienen un fin médico, ¿se prevén desde fases tempranas (idea/diseño) los recursos necesarios para la certificación del sistema de IA como producto sanitario?

(respuesta única)

- Si, en todos los casos
- Sí, en algunos casos
- No

5.11. En el caso de los sistemas de IA ya desarrollados, ¿se ha procedido a la certificación como Productos Sanitarios con IA de acuerdo con su finalidad de uso antes de su implementación clínica?

(respuesta única)

- Si, en todos los casos
- Sí, en algunos casos
- No

5.12. ¿Se aplican en su centro guías de referencia para garantizar las Buenas Prácticas en Investigación con IA?

(respuesta única)

- Sí, en todos los casos
- Sí, en algunos casos
- No

5.13. ¿Qué guías de referencia se utilizan?

(respuesta múltiple)

- SPIRIT-AI
- CONSORT-AI
- TRIPOD-AI
- STARD-AI
- Otros _____
- No aplica

5.14. ¿Participa su instituto en proyectos relacionados con el Espacio Europeo de Datos de Salud (EHDS)?

(respuesta única)

- Sí, con proyectos en marcha
- No, en proceso de incorporación
- No

5.15. ¿Dispone el IIS de una estrategia institucional para acompañar a los investigadores a desarrollar evaluaciones de impacto de protección de datos de carácter personal?

(respuesta única)

- Sí, con proyectos en marcha
- No, en proceso de incorporación
- No

5.16. ¿Dispone el IIS de una estrategia institucional para acompañar a los investigadores a desarrollar evaluaciones de impacto derivadas del Reglamento Europeo de Inteligencia Artificial?

(respuesta única)

- Sí, con proyectos en marcha
- No, en proceso de incorporación
- No

5.17. ¿Dispone el IIS de una estrategia institucional de gobernanza que permita detectar si se cumplen las obligaciones que se derivan del Reglamento Europeo de Inteligencia Artificial?

(respuesta única)

- Sí, con proyectos en marcha

- No, en proceso de incorporación
- No

6. Proyectos con uso de Big Data e IA en investigación biomédica:

6.1. ¿Existe en su centro un circuito definido para la preparación, planificación y aprobación de proyectos con uso de Big Data e IA?

(respuesta única)

- Sí
- En proceso de definición
- No
- Otros: (especificar) _____

6.2. ¿Con qué tipología de datos se trabaja en los proyectos de investigación del centro?

(respuesta múltiple)

- Historia Clínica Electrónica:
 - Datos Clínicos
 - Datos de Laboratorio
 - Imagen médica
 - Otros
- Wearables / dispositivos médicos
- Genómicos / ómicos
- Biobanco
- Administrativos / gestión
- Socioeconómicos
- Ambientales
- Registros nacionales/autonómicos
- Otros (especificar) _____

6.3. Respecto a los datos clínicos, ¿existe en su centro u hospital asociado la posibilidad de obtener volcado masivo para uso secundario en investigación?

(respuesta única)

- Sí
- En proceso de construcción
- No
- Otros: (especificar) _____

6.4. Respeto a los datos clínicos, ¿con qué nivel de estandarización están disponibles?

(respuesta única)

- Datos clínicos no estandarizados.
- Datos parcialmente estandarizados (OMOP, i2b2, CDISC, FHIR, SNOMED, LOINC, OMOP, etc.)
- Datos completamente estandarizados (OMOP, i2b2, CDISC, FHIR, SNOMED, LOINC, OMOP, etc)

6.5. ¿Qué nivel de integración tienen los datos clínicos entre centros?

(respuesta múltiple)

- Local (entre centros pertenecientes al IIS)
- Provincial (entre centros de la región o provincia)
- Autonómico (entre centros de la Comunidad Autónoma)
- Ninguna de las anteriores

6.6. Número de proyectos activos con Big Data / IA en el año previo:

(respuesta única)

- 0
- 1-9
- 10-24
- 25-49
- ≥ 50 (especificar): _____
- NS/NC

6.7. Número de proyectos activos con Big Data / IA en el año actual

(respuesta única)

- 0
- 1-9
- 10-24
- 25-49
- ≥ 50 (especificar): _____
- NS/NC

6.8. Número de proyectos con Big Data / IA finalizados en el año previo

(respuesta única)

- 0
- 1-9
- 10-24
- 25-49

- ≥ 50 (especificar): _____
- NS/NC

6.9. Nº de investigadores implicados en proyectos con Big Data/IA:
(respuesta única)

- 0
- 1-9
- 10-24
- 25-49
- ≥ 50 (especificar): _____
- NS/NC

6.10. Estos proyectos son estudios:
(indicar % en cada categoría, suma debe ser 100%)

- (%) prospectivos
- (%) retrospectivos
- (%) Mixtos

6.11. ¿En qué estado de madurez se encuentran estos proyectos?
(indicar % en cada categoría, suma debe ser 100%)

- (%) Fase de idea
- (%) Fase de diseño
- (%) Fase de desarrollo
- (%) Fase de validación
- (%) Fase de marcado CE
- (%) Fase de implementación en la práctica clínica

6.12. ¿Están sus proyectos protegidos intelectualmente?
(respuesta única)

- Si, en su totalidad
- Sí, parcialmente
- No

6.13. ¿En qué áreas se aplican los proyectos de investigación con IA?
Indique el % aproximado

(respuesta múltiple, suma debe ser 100%)

- (%) Investigación clínica
- (%) Epidemiología / salud poblacional
- (%) Imagen médica (radiología, anatomía patológica, oftalmología, etc.)
- (%) Genómica / multi-ómica
- (%) Gestión hospitalaria / eficiencia
- (%) Otras (especificar) _____

6.14. Fuentes de financiación de los proyectos de investigación con Big Data e IA:

(respuesta múltiple, suma de % debe ser 100%)

- Fondos europeos (Horizon Europe, Digital Europe, etc.) ____%
- Fondos nacionales (ISCIII / Agencia Estatal de Investigación) ____%
- Fondos autonómicos ____%
- Fondos privados (industria farmacéutica, tecnológica) ____%
- Fondos propios del IIS ____%
- Otros (donaciones,... especificar) _____ %

6.15. ¿Participa el IIS en colaboraciones público-privadas para el desarrollo de proyectos de investigación con Big Data e IA?

(respuesta única)

- Sí
- No

6.16. ¿Cuenta su IIS con participación ciudadana y de pacientes en proyectos de IA?

(respuesta única)

- Sí, estructurada (comités, grupos de trabajo)
- Sí, puntual
- No

6.17. ¿Participa su centro en consorcios multicéntricos en proyectos de IA?

(respuesta múltiple)

- Sí, a nivel local/regional
- Sí, a nivel nacional
- Sí, a nivel internacional
- Sin participación en estudios multicéntricos

6.18. ¿Dispone de proyectos en marcha con aprendizaje federado?

(respuesta única)

- Sí, en operación
- Planificados
- No

6.19. En caso afirmativo, indique qué plataformas utiliza: (respuesta múltiple)

- OTwin
- Flower

- FATE
- Google FL
- Otras (especificar) _____

6.20. ¿Se utiliza IA Generativa en los proyectos de investigación del IIS?

(respuesta única)

- Sí, en varios proyectos
- Sí, en piloto
- No
- NS/NC

6.21. ¿Se utilizan datos sintéticos en los proyectos de investigación del IIS?

(respuesta única)

- Sí
- En exploración
- No

7. Obstáculos y Retos:

7.1. ¿Cuáles son las principales barreras para llevar a cabo los proyectos con Big Data e IA en su instituto?

(respuesta múltiple)

- Demasiadas obligaciones regulatorias
- Falta de conocimiento de las obligaciones regulatorias
- Técnicas (interoperabilidad, estandarización)
- Financiación insuficiente
- Infraestructura insuficiente
- Escasez de personal cualificado
- Dificultades de colaboración entre centros
- Dificultad de obtención del dato para investigación
- Otros (especificar) _____

7.2. Tiempo medio de aprobación de proyectos Big Data/IA en CEIm:

(respuesta única)

- < 3 meses
- 3–6 meses
- > 6 meses

7.3. Una vez aprobado el proyecto, tiempo medio de espera para recepción de los datos clínicos para su desarrollo:

(respuesta única)

- < 3 meses
- 3–6 meses
- > 6 meses

7.4. ¿Cuáles son las dificultades de la transferencia de activos de IA generados en investigación?

(respuesta múltiple)

- Financieros
 - Aprobaciones por organismos reguladores.
 - Falta de conocimiento del proceso de transferencia.
 - Alta competencia internacional
 - Política de implementación en el sistema nacional de salud.
 - Otros (especificar)
-

8. Perspectivas y necesidades

8.1. Principales necesidades inmediatas:

(respuesta múltiple)

- Infraestructura tecnológica
 - Personal especializado
 - Formación
 - Marco normativo
 - Gobernanza
 - Financiación
 - Colaboración con otros centros
 - Otras (especificar)
-

8.2. Que ayudaría más al IIS para poder desarrollar Proyectos de IA (ordenar de 1 a 5 del más importante al menos importante):

(respuesta múltiple)

- (____) Financiación específica
 - (____) Formación
 - (____) Infraestructuras
 - (____) Adaptación de la normativa europea a nivel español
 - (____) Otros (especificar) _____
-

8.3. Expectativas de evolución en 3–5 años:



(especificar)

9. Observaciones Adicionales:

9.1. ¿Considera útil repetir la encuesta periódicamente?

- Sí, cada año
- Sí, cada 2 años
- No

9.2. Otras observaciones finales:

Subproyecto 3: Guía de requisitos técnicos y regulatorios para proyectos con Big Data e Inteligencia Artificial (IA) en salud

Versión: 1.2 · Fecha: 03/12/2025

Grupo de trabajo: GdT5

Coordinadores: IBSAL · IIS La Fe · INIBIC

Control de versiones

VERSIÓN	FECHA	CAMBIOS PRINCIPALES
1.0	15/09/2025	Publicación versión inicial
1.1	30/09/2025	Revisión GdT5-IA
1.2	02/12/2025	Revisión GT-IA ISCIII

Nota legal y de alcance

Este documento ofrece **criterios operativos** para proyectos de **IA** y **big data** en salud. **No sustituye** a la normativa vigente; en caso de discrepancia, **prevalece el MDR (UE) 2017/745, IVDR (UE) 2017/746, AI Act, RGPD/LOPDGDD, RD 192/2023** y las guías de **AEMPS/MDCG**. El término **ICPS** se usa aquí como **investigación clínica con producto sanitario** (estudios regulados). El **despliegue asistencial** se refiere al uso rutinario fuera del marco de investigación.

Agradecimientos

A los miembros del **GdT5 AI — Subgrupo 3**, y a las instituciones colaboradoras (**ISCIII, IBSAL, IIS La Fe e INIBIC**), por sus aportaciones técnicas, regulatorias y clínicas.

A los equipos de **datos, TI y ciberseguridad** de los centros participantes por su apoyo en interoperabilidad e integración.

Resumen ejecutivo

Esta guía ofrece un marco práctico y verificable para diseñar, evaluar e implantar proyectos de inteligencia artificial (IA) y big data en salud en el ámbito del SNS español. Su objetivo es facilitar decisiones informadas por parte de equipos investigadores, gestores clínicos y unidades de apoyo, alineando los proyectos con los requisitos científicos, ético-legales y organizativos vigentes, y acelerando su traslación a la práctica clínica.

A quién va dirigida

- Investigadores y grupos de investigación que desarrollan o implementan modelos de IA/big data.
- Unidades de Investigación Clínica (UIC) y plataformas de soporte que acompañan los estudios.

- Comités y unidades transversales (ética, protección de datos, seguridad de la información, calidad).
- Aliados externos (empresas medtech/pharma, pymes tecnológicas y administraciones sanitarias).

Qué aporta

- Un **mapa regulatorio aplicado al caso**: identificación de normativa, guías y roles implicados; cronograma de obligaciones del AI Act y de producto sanitario cuando aplique.
- Un **itinerario por niveles de madurez (TRL)** con evidencias mínimas por hito y criterios de puerta de paso.
- Un **modelo de evaluación y baremación** transparente, con acta, hoja de puntuación por bloques y condiciones de financiación/seguimiento.
- Plantillas, listas de verificación y salidas normalizadas para anexar a memorias, PNT, HIP/CI y expedientes regulatorios.
- Pautas de **interoperabilidad y gobierno del dato** (calidad, trazabilidad, catálogo, intercambio seguro con redes/consorcios).
- Requisitos esenciales de **ciberseguridad y protección de datos** desde el diseño, con controles técnicos y organizativos mínimos.
- Recomendaciones de **formación** priorizada para equipos (protección de datos, IA responsable, producto sanitario y CTIS).

Estado del documento: finalizado.

Fecha de aprobación: [Pendiente]

Próxima revisión programada: [24 meses desde aprobación]

Subproyecto 3: Guía de requisitos técnicos y regulatorios para proyectos con Big Data e Inteligencia Artificial (IA) en salud

1. Objetivo y alcance

Propósito de la guía: integrar en un único marco los requisitos técnicos, éticos y regulatorios aplicables a proyectos de Inteligencia Artificial (IA) y Big Data en salud, de forma práctica y verificable, para facilitar su diseño, evaluación, autorización y despliegue seguro en el Sistema Nacional de Salud (SNS). La guía busca homogeneizar criterios entre investigadores, evaluadores y órganos de supervisión, reducir la incertidumbre regulatoria y promover el valor clínico y la protección de las personas.

Alcance: aplica a proyectos de investigación, validación y despliegue asistencial en el SNS que utilicen datos de salud y/o algoritmos de IA, incluyendo tanto software cualificado como producto sanitario y proyectos cuyo software no tiene como finalidad ni consideración de producto sanitario, pero emplea analítica avanzada.

- **Ciclo de vida cubierto:** Concepción y diseño → preparación y gobernanza de datos → desarrollo y validación (interna/externa) → evaluación clínica → implantación en entorno asistencial → monitorización posdespliegue, mantenimiento y retirada.
- **Tipología de proyectos:**
 - i. Proyecto de Investigación (PI) – investigación y desarrollo algorítmico con datos retrospectivos/prospectivos (TRL 1–3);
 - ii. Investigación Clínica con Producto Sanitario (ICPS; estudios regulados, TRL ≥4). *Cuando existe uso rutinario fuera del estudio, se denomina despliegue asistencial (no ICPS).*
- **Ámbitos y entornos:** unidades clínicas y de apoyo, investigación translacional, redes multicéntricas y escenarios de aprendizaje federado.
- **Tipos de datos:** historia clínica electrónica (HCE), imagen médica (incluido DICOM), señales, datos ómicos, informes clínicos, Patient-Reported Outcome Measures / Patient-Reported Experience Measures (PROMs/PREMs) y datos administrativos u operacionales.

Límites y exclusiones: quedan fuera del alcance las aplicaciones no sanitarias (p. ej., bienestar/fitness sin finalidad médica), comunicaciones comerciales o publicitarias y ámbitos de investigación básica en modelos animales. Esta guía no sustituye a la normativa vigente; en caso de discrepancia prevalece la regulación aplicable en la UE y en España.

Qué exige la guía: la guía exige: (i) un **marco regulatorio aplicado al caso**, con la cualificación (incluida la de producto sanitario si procede), la clase y las obligaciones clave (MDR/IVDR, AI Act), así como las referencias MDCG/AEMPS y la definición de roles; (ii) **evidencias por TRL** y la correspondiente **decisión de paso** documentada; (iii) **interoperabilidad y gobierno del dato**, incluyendo catálogos, estándares, medidas de calidad y mecanismos de intercambio seguro; (iv) **ciberseguridad by design**, con análisis y tratamiento de riesgos, controles técnicos y organizativos mínimos, continuidad y respuesta a

incidentes; y (v) **evaluación y baremación**, con acta, hoja de puntuación por bloques, condiciones de financiación y registro de seguimiento.

Nota de uso: las secciones posteriores detallan requisitos verificables y anexos con listas de comprobación y plantillas. La numeración y las referencias internas facilitan su adopción en convocatorias y evaluaciones del ISCIII y otros agentes.

2. Públicos objetivo y usos previstos

2.1. Públicos objetivo

- Investigadores clínicos y expertos en análisis de datos e IA, promotores y responsables científicos de proyectos de IA y big data en salud.
- Equipos clínico-asistenciales (médicos, enfermería y otros profesionales sanitarios) y equipos de apoyo (farmacia, radiología, laboratorio, anatomía patológica, documentación clínica, etc.).
- Unidades de innovación y transformación digital, oficinas de datos y áreas de tecnologías de la información hospitalarias (TI); responsables de interoperabilidad y estándares (p. ej., **HL7 FHIR**, **DICOM**, **openEHR**, **OMOP**, **IHE**), sin limitarse exclusivamente a ellos).
- Delegados de Protección de Datos (**DPD**), responsables y encargados del tratamiento, y equipos de seguridad/ciberseguridad.
- Comités de Ética de la Investigación (**CEIm**) y órganos de gobernanza ética.
- Autoridades y agencias reguladoras (p. ej., **AEMPS**) y servicios de salud autonómicos.
- Fabricantes y proveedores tecnológicos (incluido software como producto sanitario).
- Evaluadores y financiadores públicos y privados responsables de convocatorias del **ISCIII**, procesos de compra pública de innovación, etc.).
- Pacientes y asociaciones, en lo relativo a participación, información y coevaluación del valor clínico; Comités de Seguridad del Paciente y Calidad Asistencial, responsables de validar planes de mitigación y supervisar incidentes en la práctica clínica (en línea con MDR GSPR y normativa nacional de seguridad del paciente); y Unidades de Evaluación de Tecnologías Sanitarias (HTA) y Agencias de Calidad, cuyos informes son determinantes para la adopción, financiación y escalado de proyectos de IA en el SNS (p. ej., RedETS, EUnetHTA).

2.2. Usos previstos:

- **Diseño del proyecto:** apoyar la definición de objetivos clínicos y de negocio, población diana, variables y criterios de valoración (*endpoints*), así como la planificación del ciclo de vida (desde la concepción y diseño hasta la monitorización posdespliegue).

- **Gobernanza de datos y cumplimiento normativo:** estructurar políticas de protección de datos y seguridad conforme al RGPD y la LOPDGDD, incluyendo bases jurídicas, Evaluación de Impacto en Protección de Datos (DPIA), minimización, control de acceso, trazabilidad y políticas de retención/eliminación.
- **Desarrollo y validación técnica:** fijar criterios mínimos de calidad de datos, métricas de rendimiento (incluida calibración), validación interna y externa y estrategias para la gestión de sesgos.
- **Regulación y ética:** determinar si el software se califica como producto sanitario, su clase de riesgo y la documentación exigible; preparar la presentación del proyecto para su evaluación por el Comité de Ética de la Investigación (**CEIm**) y, cuando proceda, ante la Agencia Española de Medicamentos y Productos Sanitarios (**AEMPS**).
- **Interoperabilidad e integración clínica:** orientar la integración con la Historia Clínica Electrónica (**HCE**), estándares como **HL7 FHIR**, **DICOM** y otros sistemas, asegurando además requisitos de experiencia de usuario clínica (**UX**) y seguridad.
- **Evaluación y financiación:** aplicar criterios y baremos para proyectos de investigación (PI) e investigación clínica con producto sanitario (ICPS), considerando la viabilidad, la madurez tecnológica (**TRL**), y el apoyo a decisiones de inversión y financiación.
- **Implantación y operación:** planificar pilotos, formación de usuarios, gestión del cambio, monitorización posdespliegue (detección de *drift*, seguridad, rendimiento) y planes de actuación controlada.
- **Escalabilidad y sostenibilidad:** facilitar el escalado multicéntrico, la cooperación (incluido el aprendizaje federado), el mantenimiento y la evaluación económica (p. ej., coste total de propiedad **TCO**, retorno de la inversión **ROI** y razón incremental coste-efectividad **ICER**).

3. Terminología y definiciones clave

Este apartado define los términos esenciales para interpretar los requisitos de la guía. Las siglas se explican la primera vez que aparecen. Cada definición incluye, cuando es relevante, el contexto normativo o técnico que condiciona su aplicación en proyectos de IA y Big Data en salud.

3.1. Acrónimos de uso frecuente:

A. Regulatorio, ético y protección de datos (UE/España)

- **AEPD:** Agencia Española de Protección de Datos.
- **EDPB/CEPD:** European Data Protection Board / [Comité Europeo de Protección de Datos](#).
- **MDCG:** Medical Device Coordination Group (Grupo de Coordinación de Productos Sanitarios).
- **EUDAMED:** Base europea de datos de productos sanitarios.

- **UDI:** Identificador único de producto sanitario.
- **ON (NB):** Organismo Notificado (Notified Body).
- **PMS/VPC:** Post-Market Surveillance / Vigilancia poscomercialización.
- **PMCF:** Post-Market Clinical Follow-up (seguimiento clínico poscomercialización).
- **GSPR:** General Safety and Performance Requirements (Requisitos generales de seguridad y funcionamiento del MDR).
- **ISO 13485:** Sistemas de gestión de la calidad para productos sanitarios.
- **ISO 14971:** Gestión de riesgos para productos sanitarios.
- **ISO 14155:** Investigación clínica de productos sanitarios en sujetos humanos (BPC).
- **IEC 62304:** Ciclo de vida del software de productos sanitarios.
- **NIS2:** Directiva europea de ciberseguridad para sectores esenciales.
- **ENS:** Esquema Nacional de Seguridad.

B. Interoperabilidad y estándares clínicos

- **HL7 v2 / CDA:** Mensajería HL7 versión 2 / Clinical Document Architecture.
- **IHE:** Integrating the Healthcare Enterprise (perfiles de integración).
- **openEHR:** Modelo de información clínica abierto.
- **OMOP CDM (OHDSI):** Common Data Model del consorcio OHDSI.
- **PACS / RIS / LIS:** Picture/ Radiology/ Laboratory Information System.
- **SNOMED CT / LOINC / ICD-10-ES / ATC / RxNorm:** Terminologías clínicas y de fármacos.
- **CDS Hooks / SMART on FHIR:** Integración de soporte a la decisión y apps sobre FHIR.
- **ONNX:** Open Neural Network Exchange (portabilidad de modelos).

C. Datos, seguridad y operación

- **ETL / ELT:** Extracción-Transformación-Carga / Extracción-Carga-Transformación.
- **RBAC / ABAC:** Control de acceso basado en roles / atributos.
- **IAM / SSO / OIDC / OAuth2 / TLS / PKI:** Gestión de identidades, inicio único y protocolos de autenticación/cifrado.
- **SIEM / SOC:** Gestión de eventos de seguridad / Centro de operaciones de seguridad.
- **SBOM:** Software Bill of Materials.
- **CVE / CWE:** Vulnerabilidades y debilidades comunes.
- **SAST / DAST:** Análisis estático / dinámico de seguridad.
- **RPO / RTO:** Objetivo de punto de recuperación / de tiempo de recuperación.
- **CI/CD:** Integración y despliegue continuos.
- **SLA/SLO:** Acuerdo / objetivo de nivel de servicio.
- **RACI:** Matriz de responsabilidades (Responsible, Accountable, Consulted, Informed).

- **DTA:** Data Transfer Agreement (acuerdo de transferencia de datos).
- **DSA:** Data Sharing Agreement (acuerdo de compartición de datos).
- **DPA:** Data Processing Agreement (contrato de encargo de tratamiento de datos).
- **JCA:** Joint Controller Agreement (acuerdo de corresponsabilidad de tratamiento).
- **UX:** Experiencia de usuario.

D. IA, métricas y evaluación

- **AUC-ROC / AUPRC:** Área bajo la curva ROC / PR.
- **F1 / PPV (VPP) / NPV (VPN):** Medidas de rendimiento de clasificación.
- **ECE / Brier:** Error de calibración esperado / puntuación de Brier.
- **KPI/KPIs:** Indicadores clave de desempeño.
- **HTA:** Health Technology Assessment (evaluación de tecnologías sanitarias).
- **BIA:** Budget Impact Analysis (análisis de impacto presupuestario).
- **QALY / DALY:** Años de vida ajustados por calidad / perdidos por discapacidad.

E. Compra y financiación

- **CPI:** Compra Pública de Innovación.
- **ROI / TCO / ICER:** Retorno de inversión / coste total de propiedad / razón incremental coste-efectividad.

3.2. Definiciones Clave:

A. Inteligencia Artificial y sistemas relacionados

Sistema de IA (AI Act): sistema basado en técnicas estadísticas, lógicas o computacionales que, a partir de objetivos definidos por personas, es capaz de inferir, predecir, clasificar, recomendar o tomar decisiones con impacto en entornos físicos o virtuales. El AI Act establece que, en el ámbito de la salud, estos sistemas suelen considerarse de **alto riesgo**, lo que conlleva obligaciones reforzadas en materia de gestión de riesgos, supervisión humana, transparencia y seguridad.

B. Datos y privacidad

Big Data en salud: conjuntos de datos sanitarios masivos, heterogéneos, multimodales y/o longitudinales (p. ej., HCE, imagen médica, ómicas, señales, texto libre) que requieren arquitecturas y procesos específicos de gobernanza, calidad, seguridad y análisis.

Datos personales de salud: información relativa a la salud física o mental de una persona identificada o identificable (incluye datos inferidos). Su tratamiento exige bases jurídicas reforzadas y medidas de seguridad avanzadas conforme al Reglamento General de Protección de Datos (RGPD) y la Ley Orgánica de Protección de Datos y Garantía de Derechos Digitales (LOPDGDD).

Pseudonimización: tratamiento de datos personales de manera que ya no puedan atribuirse a un interesado sin utilizar información adicional, siempre que dicha información adicional se mantenga por separado y esté protegida mediante medidas técnicas y organizativas adecuadas (art. 4.5 RGPD). La AEPD aclara que la información pseudonimizada sigue siendo **dato personal**, dado que puede reidentificarse a través de la información adicional, que debe conservarse separada y debidamente protegida

Anonimización: conversión de datos personales en datos que no puedan utilizarse para identificar a ningún individuo. La anonimización debe ser irreversible en términos razonables, mediante la aplicación de medidas técnicas y organizativas que eviten la reidentificación. Los datos anonimizados quedan fuera del ámbito de aplicación del RGPD (Considerando 26).

C. Tipología de proyectos y madurez

Proyecto PI (Investigación/Desarrollo): iniciativas orientadas al desarrollo y prueba de concepto de algoritmos con datos retrospectivos o prospectivos, sin impacto directo en la práctica clínica. Corresponde a niveles de madurez tecnológica (TRL, *Technology Readiness Level*) 1–3.

Proyecto ICPS (Investigación clínica con producto sanitario): estudio regulado conforme al Reglamento de Productos Sanitarios (MDR/IVDR) y a la normativa española, autorizado por un Comité de Ética de la Investigación (CEIm) y, cuando proceda, por la Agencia Española de Medicamentos y Productos Sanitarios (AEMPS). Su objetivo es generar evidencia clínica sobre un producto sanitario, incluido el software como dispositivo médico (SaMD). Puede requerir integración técnica en el entorno asistencial, pero su uso está acotado al marco de investigación y a los procedimientos del estudio y no constituye operación rutinaria del servicio.

Despliegue asistencial (operación): uso rutinario del sistema en la práctica clínica fuera del marco de investigación. Requiere monitorización posdespliegue (*Post-Market Surveillance*, PMS y *Post-Market Clinical Follow-up*, PMCF cuando aplique) y modelo de gobernanza operativo. No se considera investigación clínica con producto sanitario (**ICPS**).

Software de uso médico (SaMD): software que, por sí solo, cumple una finalidad médica. Cuando se destina a diagnóstico, prevención, monitorización o apoyo a decisiones clínicas, puede cualificar como producto sanitario (PS) y quedar sujeto al Reglamento de PS (MDR/IVDR) y a la normativa nacional aplicable.

Fabricante legal: persona física o jurídica responsable del diseño, fabricación, marcado CE, documentación técnica y vigilancia poscomercialización de un PS (incluido software).

Evaluación de impacto en protección de datos (DPIA/EIPD): análisis sistemático para identificar y mitigar riesgos que un tratamiento de datos personales puede generar para los derechos y libertades de las personas, conforme al artículo 35 del RGPD. Es obligatoria en tratamientos de alto riesgo, como los que implican IA en salud.

Base jurídica del tratamiento: fundamento legal que habilita el tratamiento de datos personales (p. ej., misión en interés público, investigación científica, obligación legal o consentimiento cuando proceda), en coherencia con las categorías especiales de datos del artículo 9 del RGPD.

D. Conceptos técnicos clave

Calidad de datos: grado en que los datos son completos, exactos, coherentes, oportunos y representativos de la población diana; incluye la gestión de valores perdidos, anómalos (*outliers*) y sesgos de selección.

Validación interna / externa: la validación interna estima el rendimiento del modelo en los datos de desarrollo (mediante partición o remuestreo); la validación externa lo mide en centros o poblaciones distintas para evaluar la capacidad de generalización.

Generalización: capacidad del modelo para mantener rendimiento y calibración fuera del conjunto de entrenamiento, en contextos y poblaciones distintas.

Calibración: concordancia entre probabilidades estimadas y frecuencias observadas; esencial para decisiones clínicas basadas en riesgo.

Sesgo algorítmico y equidad: desviación sistemática del rendimiento o de la utilidad entre subgrupos (p. ej., por sexo, edad, etnia, centro), que requiere medición, mitigación y seguimiento.

Explicabilidad: capacidad de justificar por qué un modelo toma una decisión concreta.

Interpretabilidad: capacidad para comprender la estructura interna de un modelo. Por ejemplo, las regresiones lineales, que se interpretan en términos de sus coeficientes.

Trazabilidad y registro (*logging*): capacidad de reconstruir el historial de datos, versiones de modelo, parámetros, umbrales, decisiones y cambios a lo largo del ciclo de vida.

Supervisión humana efectiva: mecanismos que permiten a profesionales cualificados entender, intervenir, anular y responsabilizarse de decisiones apoyadas por IA.

Cambio significativo (modelo/software): modificación que puede afectar a la seguridad o al rendimiento clínico (p. ej., incorporación de nuevos datos, cambio de población diana o de arquitectura), y que exige control de cambios y, en su caso, nueva evaluación o regulación.

Monitorización posdespliegue: seguimiento continuo del rendimiento, seguridad, incidencias, drifts y uso real, con planes de mantenimiento, recalibración y retirada.

Drift:

- **Data drift:** cambio en la distribución de los datos de entrada respecto al conjunto de entrenamiento.
- **Concept drift:** cambio en la relación entre variables y el objetivo clínico.

Ambos fenómenos pueden degradar el rendimiento del modelo si no se detectan y corrigen

MLOps: conjunto de prácticas para desarrollar, validar, versionar, desplegar y mantener modelos de IA de forma reproducible y segura, incluyendo automatización, control de cambios, auditoría y gobernanza.

Interoperabilidad (HL7 FHIR, DICOM): capacidad de intercambiar e interpretar datos clínicos de forma segura y estructurada. HL7 FHIR se utiliza para recursos clínicos y DICOM para imágenes y sus metadatos.

E. Interoperabilidad y operación

Historia clínica electrónica (HCE): sistema corporativo que almacena y gestiona la información clínica longitudinal del paciente.

Patient-Reported Outcome Measures (PROMs)/ Patient-Reported Experience Measure (PREMs): resultados comunicados por pacientes (PROMs) y experiencia percibida por pacientes (PREMs), útiles como variables de resultado (*endpoints*) y para evaluar impacto asistencial.

ModelCard / DataSheet: plantillas de documentación estandarizada para documentar modelos y conjunto de datos, incluyendo origen, población, limitaciones, métricas, riesgos, usos previstos y prohibidos, y requisitos de mantenimiento.

Aprendizaje federado: paradigma de entrenamiento y validación en el que los datos permanecen en origen compartiendo únicamente parámetros o actualizaciones en el modelo bajo un marco de gobernanza y seguridad definido.

Seguridad del paciente: principio y conjunto de medidas para prevenir daños derivados del uso del sistema, incluyendo incluye gestión de riesgos, alertas, salvaguardas y validación de usabilidad clínica.

Evaluación económica: análisis del coste total de propiedad (**TCO**), retorno de la inversión (**ROI**) y coste-efectividad incremental (**ICER**) para apoyar decisiones de implantación y escalado.

UX clínica: diseño de interfaces y flujos seguros intuitivos, accesibles y eficientes, seguros), validados con usuarios y alineados con la HCE; incluye manejo de alertas, privacidad y seguridad.

IA de propósito general (General Purpose AI (GPAI)): modelos o sistemas de **IA** de alcance general, capaces de realizar una **amplia gama de tareas** y de **integrarse** en otras aplicaciones. El **AI Act** establece requisitos específicos de **documentación, transparencia y uso responsable** cuando se emplean como base de otros sistemas.

Indicador de nivel de servicio (Service Level Indicator (SLI)): métrica cuantitativa que mide un aspecto del servicio (p. ej., latencia p95, disponibilidad, tasa de errores) y sirve de base para definir acuerdos de nivel de servicio (**SLA**) y objetivos (**SLO**).

4. Marco regulatorio aplicable

Esta guía se alinea con el marco europeo (MDR/IVDR y Reglamento de IA) y con el marco nacional español (AEMPS, RD 192/2023 y protección de datos — RGPD/LOPDGDD). Se describen obligaciones clave y su cronograma de aplicación.

4.1. MDR/IVDR y cualificación de software de IA como producto sanitario

- **Cualificación y clasificación:** El software destinado a fines médicos (diagnóstico, prevención, monitorización, apoyo a decisiones, etc.) puede **cualificar como producto sanitario (PS)** bajo el **MDR (UE) 2017/745** o el **IVDR (UE) 2017/746**, con obligaciones de evaluación clínica/desempeño y gestión de riesgos. La **MDCG** mantiene guías específicas para software (p. ej., **MDCG 2020-1** sobre evaluación clínica de software). [European Commission / Guidance - MDCG endorsed documents and other guidance](#)
- **España – autoridad competente y licencias:** La **AEMPS** es autoridad competente. El **Real Decreto 192/2023** regula, entre otros, licencias previas de funcionamiento de instalaciones, entidades notificadas, reprocesamiento y vigilancia de PS en España. [AEMPS - Legislación sobre productos sanitarios](#)

- **Guías nacionales de apoyo:** La AEMPS publica guías operativas (p. ej., **Guía para la comercialización de productos sanitarios en España, 2025**), útiles para fabricantes, importadores y distribuidores. [Guía para la comercialización de PS en España](#)

4.2. Reglamento europeo de IA (AI Act) e interacción con MDR/IVDR

- **Ámbito y cronograma:** El AI Act entró en vigor el 1 de agosto de 2024. Prohibiciones y obligaciones de alfabetización en IA aplican desde el 2 de febrero de 2025; reglas de gobernanza y obligaciones para GPAI aplican desde el 2 de agosto de 2025; para **sistemas de IA de alto riesgo integrados en productos regulados** (p. ej., PS bajo MDR/IVDR) la aplicación completa se extiende hasta el 2 de agosto de 2027. [AI Act European Commission - Artificial Intelligence in healthcare](#)
- **Alto riesgo en salud:** Los **sistemas de IA que sean componentes de seguridad** de productos regulados (incluidos PS) se consideran **de alto riesgo** y deben cumplir requisitos de **gestión de riesgos, gobernanza y calidad de datos, documentación técnica, trazabilidad y registro de eventos (logging), transparencia y supervisión humana**, normalmente integrados en el procedimiento de evaluación de conformidad del producto sanitario. [European Commission - Artificial Intelligence in healthcare](#)

4.3. Marco nacional español y protección de datos

- **AEMPS y RD 192/2023:** En España, la AEMPS aplica el MDR/IVDR y el RD 192/2023 para licencias, vigilancia, fabricación (incluida *in-house* cuando proceda) y distribución, con trámites y registros específicos a nivel estatal/autonómico. [AEMPS - Legislación sobre productos sanitarios](#)
- **RGPD y LOPDGDD:** El RGPD (UE) 2016/679 regula el tratamiento de datos de salud, que forman parte de las categorías especiales de datos personales. La base jurídica de legitimación debe encontrarse siempre en el **artículo 6 del RGPD** (p. ej., art. 6(1)(e) misión en interés público o art. 6(1)(a) consentimiento en determinados supuestos), en conexión con lo previsto en las leyes de los Estados Miembros. El **artículo 9 RGPD** no constituye en sí mismo una base jurídica, sino que establece excepciones a la prohibición general de tratar categorías especiales de datos; en investigación sanitaria suele aplicarse la excepción del art. 9(2)(j) (fines de investigación científica), junto con las garantías del art. 89(1) (p. ej., seudonimización). La LOPDGDD complementa y desarrolla esta regulación en España. [Reglamento \(UE\) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.](#)
- **Derechos y salvaguardas:** Deben evaluarse **DPIA/EIPD**, minimización, control de accesos, retención y ejercicio de derechos. La **AEPD** ofrece orientaciones específicas sobre datos de salud. [Comité Europeo de Protección de Datos, Agencia Española de Protección de Datos](#)

4.4. Normas técnicas internacionales de referencia

Además de los marcos regulatorios europeos y nacionales, existen normas técnicas internacionales que orientan el desarrollo y evaluación de productos sanitarios basados en IA:

- **ISO 13485:** gestión de calidad en productos sanitarios.
- **ISO 14971:** gestión de riesgos en dispositivos médicos.
- **IEC 62304:** ciclo de vida del software sanitario.
- **ISO 14155:2020:** Buenas Prácticas Clínicas en investigaciones con productos sanitarios.
- **ICH E6(R3) – Good Clinical Practice:** actualmente en proceso de adopción internacional, actualiza las Buenas Prácticas Clínicas y será de aplicación transversal a los ensayos clínicos, incluidos aquellos con productos sanitarios basados en IA. Su incorporación refuerza la armonización global de requisitos éticos, regulatorios y metodológicos.

Nota práctica: En proyectos PI (TRL 1–3) y ICPS (TRL ≥4), documenta la **calificación del software**, la **clase de riesgo**, la **base jurídica** de datos, el **plan de evaluación clínica/desempeño**, y el **encaje AI Act** (alto riesgo), alineando el expediente técnico con MDR/IVDR y, cuando proceda, con los hitos de aplicación del AI Act.

4.5. Ciberseguridad y resiliencia en hospitales.

La ciberseguridad constituye un elemento crítico en la implantación de sistemas de IA y Big Data en salud, dado que los hospitales forman parte de las infraestructuras críticas nacionales y son objetivos frecuentes de ciberataques. La protección de los sistemas debe garantizar la confidencialidad, integridad y disponibilidad de los datos clínicos, así como la continuidad asistencial en caso de incidentes.

Los proyectos deberán cumplir con los requisitos del [Esquema Nacional de Seguridad \(ENS\)](#) y de la [Directiva NIS2](#), incluyendo planes de prevención, detección, respuesta y recuperación ante incidentes. Además, se espera la definición de protocolos específicos de contingencia que contemplen la coordinación con los equipos de seguridad del centro sanitario y con los equipos de expertos en seguridad informática ([CSIRT](#), Computer Security Incident Response Team) competentes.

En este contexto, es recomendable vincular los planes de ciberresiliencia hospitalaria con las capacidades desarrolladas a nivel europeo. El [Reglamento de Cibersolidaridad \(UE\) 2025/38](#) establece la **Reserva de Ciberseguridad de la Unión Europea**, un mecanismo que pone a disposición de los Estados miembros recursos especializados de detección, análisis y respuesta frente a incidentes graves que afecten a infraestructuras críticas, incluyendo las sanitarias.

Asimismo, el **Comité Europeo de las Regiones** ha señalado la importancia de los **centros regionales de apoyo en ciberseguridad**, concebidos como hubs descentralizados que ofrecen soporte directo a hospitales y centros de investigación biomédica con menor capacidad propia. Estos centros se coordinan con las autoridades nacionales, con ENISA y con los Security Operation Centres (SOC) europeos, garantizando una vigilancia en tiempo real y una capacidad de respuesta reforzada.

La integración de estas capacidades europeas en los planes institucionales de ciberseguridad hospitalaria aporta **redundancia, soporte especializado y capacidad de escalado** en caso de incidentes

de gran magnitud. Al mismo tiempo, sitúa a los proyectos de IA y Big Data en salud en un marco de **ciberresiliencia alineado con la estrategia digital europea**, fortaleciendo la confianza de profesionales, gestores y pacientes en el uso seguro de estas tecnologías.

Salida esperada del apartado 4: Mapa regulatorio aplicado al caso (RGPD/LOPDGDD; MDR/IVDR si procede; AI Act; NIS2 y ENS; EHDS; CRA; normas IEC/ISO pertinentes). Cualificación y clase del sistema (si es producto sanitario), rol del operador (fabricante legal / responsable del despliegue) y estrategia de conformidad (GSPR, evaluación clínica/desempeño, vigilancia poscomercialización); cronograma del AI Act para el proyecto con hitos y obligaciones aplicables (transparencia, alto riesgo, documentación técnica, monitorización, registro); Listado de guías MDCG/AEMPS referenciadas y trazabilidad de su aplicación al proyecto (enlaces y versión); Alineación con ciberresiliencia: plan ENS/NIS2, coordinación con CSIRT competente y, cuando aplique, integración con la Reserva de Ciberseguridad de la UE y centros regionales de apoyo.

5. Tipología de proyectos y niveles de madurez (TRL)

Diferenciamos dos familias de proyectos por su finalidad y nivel de madurez: **PI** (investigación y desarrollo, **TRL 1–3**) e **ICPS** (Investigación clínica con producto sanitario, **TRL ≥4**). La escala **TRL** se adapta aquí a **IA en salud** para orientar hitos, evidencias y decisiones de paso. Esta adaptación permite alinear los niveles de madurez con los criterios de evaluación y financiación, así como con los marcos normativos aplicables (MDR/IVDR, AI Act, RGPD/LOPDGDD), facilitando la trazabilidad regulatoria y la planificación del despliegue asistencial.

5.1. Tipología de proyectos

- **PI (Investigación/Desarrollo):** generación y prueba de conceptos algorítmicos con datos retrospectivos o prospectivos controlados. Enfoque principal en metodología, calidad de datos, rendimiento y generalización inicial. No hay uso asistencial real ni impacto directo sobre pacientes. En la fase PI no aplica la regulación como producto sanitario ni el AI Act, salvo que el software ya realice funciones críticas en salud. Sí se recomienda documentar la **finalidad prevista** y la **base jurídica del tratamiento de datos**, especialmente cuando se utilicen datos personales.
- **ICPS (Investigación clínica con PS):** estudio regulado con CEIm/AEMPS cuando proceda, orientado a generar evidencia clínica sobre el PS/SaMD. Puede requerir integración técnica en el hospital para ejecutar el protocolo, pero el uso queda circunscrito al ámbito del estudio. Esta fase debe incluir un **plan de monitorización posdespliegue (PMS/PMCF)**, control de cambios y trazabilidad operativa, así como el cumplimiento de los requisitos de **seguridad del paciente** y la garantía de una **supervisión humana efectiva**.
- **Despliegue asistencial (operación):** uso rutinario fuera del ensayo/estudio. El despliegue asistencial queda implícito en los TRL 8–9 como fase posterior a ICPS. En esta guía, los TRL se han **adaptado específicamente al contexto de IA en salud**, incorporando requisitos técnicos,

regulatorios y clínicos. Cada nivel debe documentarse con **artefactos verificables** que respalden la progresión hacia fases posteriores.

5.2. TRL adaptados a IA en salud: hitos y evidencias mínimas

Los niveles de madurez tecnológica (TRL) se han **adaptado específicamente al ámbito de la IA** en salud, incorporando evidencias técnicas, regulatorias y clínicas. Cada nivel debe documentarse mediante **artefactos verificables** (protocolos, informes, validaciones, aprobaciones regulatorias) que permitan trazar de forma clara la progresión hacia etapas posteriores.

A continuación, se describen los **nueve niveles de madurez tecnológica (TRL)** adaptados a proyectos de IA en salud, junto con los hitos y evidencias mínimas que deben acompañar cada fase.

TRL 1 – Principios básicos observados. Conocimiento clínico y metodológico preliminar; definición del problema y de los posibles **endpoints**.

Evidencias: revisión del estado del arte, hipótesis clínica, boceto de **finalidad prevista (intended use)**.

TRL 2 – Concepto formulado. Diseño conceptual del enfoque algorítmico y plan de datos.

Evidencias: protocolo PI de alto nivel, esquema de variables, criterios de inclusión/exclusión, plan de gobernanza y **DPIA** preliminar.

TRL 3 – Prueba de concepto. Desarrollo inicial con datos de investigación; validación interna básica.

Evidencias: dataset curado (diccionario y linaje), métricas primarias, control de sobreajuste, análisis de sesgos inicial, documentación reproducible.

TRL 4 – Mínimo producto viable (entorno controlado). Versión ejecutable con **SOP** técnicas y de datos; validación interna robusta y **validación externa** inicial.

Evidencias: paquete software trazable (versionado), **ModelCard** y **DataSheet** iniciales, plan de riesgos, seguridad y ciberseguridad.

TRL 5 – Validación técnica en entorno relevante. Ensayos con datos de centros distintos y condiciones próximas a la clínica (sin influir decisiones).

Evidencias: validación externa multicéntrica, análisis de **calibración**, rendimiento por subgrupos y mitigación de sesgos, plan de interoperabilidad.

TRL 6 – Piloto *shadow mode* (pruebas en la sombra) en clínica. Integración con sistemas (**HCE/FHIR/DICOM**) y ejecución en paralelo al flujo clínico, sin sustituir juicio profesional.

Evidencias: evaluación de usabilidad, tiempos y carga asistencial; registros (**logging**) y alertas; **plan de cambio** predefinido.

TRL 7 – Uso supervisado con impacto asistencial acotado. Despliegue controlado en unidades piloto con supervisión humana efectiva.

Evidencias: aprobación ética/organizativa, indicadores de seguridad del paciente, procedimiento de **override** clínico, acuerdo de soporte y mantenimiento.

TRL 8 – Preparación para adopción regulada/operativa. Documentación técnica completa y procesos de calidad para operación sostenida; si aplica, expediente para **marcado CE** y cumplimiento de requisitos del **AI Act** para alto riesgo.

Evidencias: gestión de riesgos consolidada, **post-market plan** (o plan de monitorización posdespliegue), contratos y **SLA/SLO**.

TRL 9 – Operación a escala y mejora continua. Sistema en servicio estable, monitorizado y con mejora controlada.

Evidencias: métricas en vida real, **drift** y recalibración, auditorías periódicas, ciclo **MLOps** y revisión anual.

5.3. Decisiones de paso y responsabilidades

- **PI → ICPS (\geq TRL 4):** requiere validación externa, análisis de sesgos por subgrupos, **DPIA** actualizada, plan de interoperabilidad y supervisión humana definida.
- **Pre-uso asistencial (TRL 6–7):** en **ICPS**, corresponden a pilotos y ejecución “**shadow/uso supervisado**” **dentro del estudio**; cualquier decisión clínica se rige por el protocolo y la supervisión humana definida.
- **Operación sostenida (TRL 8–9):** requiere plan de monitorización posdespliegue, control de cambios y evidencias de formación de usuarios. Suelen corresponder a **despliegue asistencial** (poscomercialización) con **PMS/PMCF** cuando aplique.

Cada transición entre TRL debe estar respaldada por evidencias documentadas, y los **gateways** pueden utilizarse como criterios objetivos de evaluación en convocatorias públicas.

5.4. Correspondencia con evaluación y financiación

- **Evaluación diferenciada:** los criterios y baremos se aplican de forma distinta en **PI** (calidad científica y de datos, metodología, potencial de generalización) e **ICPS** (seguridad, interoperabilidad, impacto asistencial, sostenibilidad).
- **Madurez mínima orientativa:** **PI** debe alcanzar **TRL 3** para optar a escalado; **ICPS** debería situarse en **TRL 6–7** para pilotos clínicos y en **TRL 8–9** para despliegue estable.

Esta adaptación TRL es operativa para proyectos de IA en salud y sirve para alinear hitos técnicos, éticos y regulatorios con decisiones de inversión y adopción. Además, los TRL permiten establecer **umbrales de madurez para la financiación pública, la evaluación ética y regulatoria, y la planificación del despliegue asistencial**. Se recomienda vincular cada TRL con los **formularios de evaluación PI/ICPS incluidos en los anexos**, de forma que exista coherencia entre la progresión tecnológica y los criterios de valoración aplicados.

Salida esperada del apartado 5: Nivel TRL declarado y justificado (PI: TRL 1–3; ICPS: TRL \geq 4). Evidencias mínimas por TRL alcanzadas y trazables (para pasar de PI a ICPS: validación externa inicial, análisis de sesgos por subgrupos, DPIA actualizada y evidencias básicas de ciberresiliencia acordes al riesgo). Hoja de ruta de madurez con hitos TRL, criterios de “cambio significativo” y plan inicial de ciberresiliencia para la transición a ICPS. Decisión de paso documentada (acta del comité/órgano competente) y referencia a los formularios de evaluación PI/ICPS anexos.

6. Requisitos sobre datos de salud y gobernanza

Este apartado establece los **criterios verificables** que deben cumplirse en relación con la obtención, preparación, protección y uso de datos de salud en proyectos de IA. Su objetivo es garantizar la **calidad**, la **representatividad**, el **cumplimiento normativo** y la **sostenibilidad del dato** a lo largo de todo el ciclo de vida del sistema, desde su concepción hasta su retirada.

6.1. Calidad, representatividad y mitigación de sesgos

La calidad de los datos es un requisito esencial para asegurar que los sistemas de IA operen de forma fiable y segura en el contexto clínico previsto. Es necesario que los datos reflejen adecuadamente la **población diana** y el **entorno asistencial** en el que se evaluará o desplegará el sistema.

- Para ello deben definirse claramente los **criterios de inclusión y exclusión**, realizar un análisis de **valores ausentes**, controlar **duplicados**, garantizar la **coherencia temporal** y aplicar una estandarización rigurosa de **unidades y codificaciones clínicas**. Además, se recomienda el uso de **muestrado estratificado** que permita representar adecuadamente subgrupos relevantes como sexo, edad, centro o comorbilidad.
- La **equidad algorítmica** debe abordarse mediante la medición del rendimiento por subgrupos, la detección de sesgos de selección e información y la aplicación de estrategias de mitigación (reponderación, recolección adicional de datos o rediseño del modelo). Estas estrategias deben revisarse con profesionales clínicos para garantizar su plausibilidad.
- Antes de avanzar hacia fases reguladas, se exige al menos una **validación externa** en un centro o población distinta, justificando la **transferibilidad** entre entornos mediante un análisis comparativo de condiciones clínicas, demográficas y operativas.

Evidencias esperadas: diccionario de variables con linaje, informe de calidad y representatividad, análisis de sesgos con métricas por subgrupos y documentación de las decisiones tomadas para su mitigación.

6.2. Gestión del ciclo de vida y plan de gestión de datos (DMP)

El **plan de gestión de datos (Data Management Plan, DMP)** es el documento que define cómo se organizarán, protegerán, versionarán y conservarán los datos utilizados en un proyecto de IA en salud. Su elaboración y aplicación permiten garantizar la **trazabilidad del dato** desde su captura inicial hasta su retirada definitiva, asegurando en todo momento la **calidad técnica**, el **cumplimiento normativo** y la **sostenibilidad operativa**. Cuando procede, este plan debe alinearse con los **principios FAIR**, facilitando la reutilización responsable y la interoperabilidad entre sistemas.

Elementos clave del DMP:

- **Inventario y procedencia:** debe incluir un catálogo completo de fuentes de datos, identificando responsables, bases jurídicas y finalidades específicas; **acreditar el origen legítimo de los datos** (evitando prácticas informales como la extracción directa de hojas Excel de la historia clínica

sin garantías); y acompañarse de trazabilidad de todas las transformaciones realizadas (ETL/ELT) y del versionado de los conjuntos de datos.

- **Metadatos y estandarización:** apoyarse en modelos de información clínica reconocidos y terminologías normalizadas; registrar diccionarios de variables, reglas de limpieza y mapeos semánticos que garanticen coherencia e interoperabilidad.
- **Retención y eliminación:** definir políticas de conservación alineadas con la finalidad y la normativa; establecer procedimientos de eliminación segura y mantener registros que acrediten la destrucción de los datos.
- **Roles y responsabilidades:** formalizar una matriz (p. ej., RACI) que identifique al propietario del dato, al custodio, al delegado de protección de datos, al equipo de seguridad y al de ciencia de datos.
- **PI vs ICPS:** en proyectos PI se prioriza la **reproducibilidad** de los resultados y la **compartición controlada**; en proyectos ICPS se exige **continuidad asistencial, trazabilidad operativa robusta y sostenibilidad del dato** para su mantenimiento en entornos clínicos reales.

Evidencias esperadas: DMP aprobado y versionado; control de versiones de los datasets; bitácora de cambios con fecha, responsable, motivo del cambio y versión afectada; acto formal de entrega de los datos que acredite su origen legítimo; plan de copias de seguridad y recuperación ante desastres con objetivos de punto y tiempo de recuperación (RPO/RTO) definidos

6.3. Seguridad, privacidad y cumplimiento (RGPD/LOPDGDD)

El tratamiento de datos en proyectos de IA en salud debe regirse por principios de **seguridad, privacidad y cumplimiento normativo**, garantizando la protección de los derechos de los pacientes y la confianza de los profesionales. Este apartado establece los **criterios verificables** que deben cumplirse en relación con la **base jurídica adecuada**, la **minimización de datos**, la **gestión de transferencias y terceros**, y la **implementación de medidas técnicas y organizativas** que aseguren la confidencialidad, la integridad y la trazabilidad del sistema.

Además de RGPD/LOPDGDD, el proyecto deberá alinearse con un **Sistema de Gestión de Seguridad de la Información (SGSI)** conforme a **ISO/IEC 27001**, apoyado en controles de **ISO/IEC 27002**, y con **ISO/IEC 27701** para la gestión de la privacidad.

- **Base jurídica y minimización:** utilizar una **base jurídica adecuada** y la categoría especial correspondiente; en el caso de entrenamiento de algoritmos, normalmente no resulta viable basarse en el consentimiento informado por los tamaños muestrales requeridos. En España, la **Disposición Adicional decimoséptima de la LOPDGDD** permite la reutilización de datos seudonimizados con fines de investigación sin necesidad de consentimiento, siempre que se cumplan las condiciones establecidas. Además, recoger solo los datos necesarios para los objetivos y mantener la separación entre datos de identificación y asistenciales.
- **Seudonimización/anonimización:** Las técnicas de seudonimización y anonimización deben aplicarse conforme al objetivo del tratamiento. La anonimización, en particular, debe **impedir**

la **reidentificación** de forma razonable e irreversible, en línea con lo establecido en el Considerando 26 del RGPD. Además, debe **evaluarse el riesgo** residual de reidentificación y proteger adecuadamente las llaves que permitirían revertir la seudonimización, mediante medidas técnicas y organizativas específicas.

- **Control de acceso y registro:** El acceso a los datos debe estar restringido según el principio de mínimo privilegio. Esto implica implementar mecanismos de autenticación robusta, registrar todas las acciones de acceso y consulta mediante sistemas de **logging**, y conservar dichos registros con garantías de integridad. El cifrado debe aplicarse tanto en tránsito como en reposo, con una gestión segura de las claves criptográficas y una segmentación adecuada de los entornos de ejecución para evitar accesos no autorizados.
- **Cifrado:** los datos deben protegerse mediante **cifrado robusto tanto en tránsito como en reposo**, aplicando algoritmos y longitudes de clave alineados con las recomendaciones internacionales vigentes. La **gestión de claves** debe realizarse de forma segura, con rotación periódica, almacenamiento en módulos de seguridad hardware (HSM) cuando sea posible y control estricto de accesos. Además, debe garantizarse la **segmentación de entornos y redes**, evitando la mezcla de datos sensibles con otros sistemas y aplicando técnicas de **aislamiento lógico y físico** que refuerzen la confidencialidad e integridad de la información.
- **Transferencias y terceros:** cuando el tratamiento implique transferencias internacionales o la participación de terceros, es obligatorio evaluar a los **encargados y corresponsables del tratamiento**, formalizar una **regulación contractual adecuada** y establecer cláusulas que garanticen el **cumplimiento normativo**. Deben aplicarse además **salvaguardas específicas** que aseguren la protección de los datos en todo su ciclo de vida. En el caso de proveedores de servicios en la nube, la evaluación debe contemplar la **ubicación física de los datos**, el cumplimiento del **Esquema Nacional de Seguridad (ENS)** y de la **Directiva NIS2**, así como la revisión detallada de las **cláusulas contractuales relativas a confidencialidad, disponibilidad y trazabilidad**.
- **DPIA/EIPD:** la **evaluación de impacto en protección de datos (DPIA/EIPD)** debe realizarse siempre que el tratamiento pueda implicar un **alto riesgo para los derechos y libertades de las personas**. Este análisis debe identificar riesgos, incluir **medidas de mitigación adecuadas** y actualizarse ante cualquier **cambio significativo** en el tratamiento, la arquitectura del sistema o la finalidad prevista. La DPIA debe estar **firmada por los responsables correspondientes** y formar parte del **expediente técnico del proyecto**, quedando sujeta a revisión periódica durante todo el ciclo de vida.
- El proyecto debe disponer de un **Plan de Respuesta ante Ciberincidentes (PRC)** con **simulacros** al menos anuales, criterios de activación, roles **RACI** (Responsable (quien ejecuta la tarea), Aprobador (Accountable, quien toma la decisión final), Consultado (Consulted, quien debe ser consultado) e Informado (Informed, quien recibe información sobre el progreso)) y contacto con el **CSIRT (Centro de respuesta a incidentes de seguridad) de referencia**. Incluir **gestión de legacy** (parches, segmentación/aislamiento, compensaciones) y **registro/notificación de incidentes** conforme a NIS2/ENS y normativa nacional.

Evidencias: Las evidencias que deben generarse en este apartado incluyen el registro de actividades de tratamiento, el informe de evaluación de impacto en protección de datos, las políticas de seguridad aplicables, los procedimientos normalizados de operación (SOP), los resultados de auditorías o pruebas de penetración cuando proceda, y las actas que acrediten la formación y concienciación del personal implicado en el tratamiento de los datos. Además, ligado al SGSI, declaración de aplicabilidad (SoA), registro de activos, análisis y tratamiento de riesgos del SGSI, controles de acceso por rol, registro/auditoría, cifrado en tránsito y reposo, gestión de copias (RPO/RTO) y pruebas periódicas de restauración.

6.4. Aprendizaje federado y colaboración multicéntrica

El aprendizaje federado es una técnica que permite entrenar y validar modelos de IA sin necesidad de centralizar los datos. A diferencia del enfoque tradicional, en el que los datos se recopilan en un único servidor para su procesamiento, **el aprendizaje federado traslada el modelo a cada institución participante**. En cada nodo, el modelo se entrena localmente con los datos disponibles, y solo se comparten las actualizaciones del modelo, como los parámetros o actualizaciones, con un servidor de agregación o con otros nodos, según la arquitectura definida. Esta estrategia permite **preservar la privacidad, reducir el riesgo de exposición de datos personales y facilitar la colaboración** entre entidades que no pueden compartir sus datos por razones legales, éticas o técnicas.

En el ámbito sanitario, esta metodología resulta especialmente **útil para trabajar con datos sensibles** como los clínicos, genómicos o de imagen médica. Además, se alinea con los principios de protección de datos, como la minimización, la limitación de la finalidad y la responsabilidad proactiva, al garantizar que la información permanece bajo el control del responsable del tratamiento y no se transfiere fuera de su entorno. Sin embargo, **el aprendizaje federado no equivale a anonimización**: las actualizaciones del modelo pueden filtrar información si no se aplican medidas técnicas y organizativas adecuadas. Por ello, debe diseñarse bajo el principio de privacidad desde el diseño y por defecto, con una evaluación de impacto en protección de datos que contemple el ciclo completo, la gobernanza interinstitucional y un modelo de amenazas específico.

Para que esta técnica sea viable, **la arquitectura del sistema federado debe definirse con precisión**. Esto implica describir los nodos donde residen los datos y se ejecuta el entrenamiento local, el agregador que consolida las actualizaciones y el canal de comunicación seguro y auditado que conecta a todos los participantes. Es necesario establecer políticas claras sobre la frecuencia y el tamaño de las actualizaciones, los criterios de participación y los procedimientos en caso de caída de nodos. Además, cada entorno debe estar aislado y reforzado mediante **hardening**, que consiste en aplicar medidas de endurecimiento de la seguridad, como la desactivación de servicios innecesarios, el control estricto de accesos y la configuración segura de los sistemas. La arquitectura debe garantizar tolerancia a fallos, reintentos automáticos y un registro completo de eventos con trazabilidad para auditoría.

La protección de la privacidad y la robustez del sistema federado exige medidas adicionales. Entre ellas se incluyen **técnicas de agregación segura** que eviten exponer contribuciones individuales, opciones de privacidad diferencial (**differential privacy**) para limitar la reconstrucción a partir de actualizaciones y defensas frente a ataques de inferencia y envenenamiento, como la verificación de actualizaciones, el recorte (**clipping**) y la **agregación robusta**. Estas medidas reducen el riesgo de filtración de

información y de manipulación del modelo por participantes maliciosos, pero deben equilibrarse con el impacto en rendimiento y convergencia, documentando las decisiones adoptadas.

La gobernanza del aprendizaje federado debe formalizarse mediante acuerdos claros que definen las responsabilidades de cada entidad en la determinación de fines y medios del tratamiento, en coherencia con los artículos 26 y 28 del RGPD. Es imprescindible establecer acuerdos de transferencia de datos (**Data Transfer Agreement, DTA**) y acuerdos de nivel de servicio (**Service Level Agreement, SLA**) que recojan niveles de seguridad, disponibilidad y trazabilidad. Además, debe constituirse un comité técnico y un comité ético con funciones definidas, incluyendo la validación de resultados, las reglas de actualización del modelo y los criterios de publicación. Cuando el aprendizaje federado soporte un producto sanitario o software como dispositivo médico, la gobernanza debe alinearse con los requisitos del MDR/IVDR y del AI Act para sistemas de alto riesgo.

La evaluación federada debe basarse en **métricas armonizadas**, con análisis por centro y consolidado, e incluir la variabilidad entre nodos, la heterogeneidad de datos y la detección de sesgos. Es necesario justificar el esquema de ponderación de contribuciones y monitorizar la calibración y el rendimiento por subgrupos en cada sitio, en coherencia con los criterios de calidad y equidad. Además, se recomienda realizar pruebas específicas para detectar posibles fugas de información antes de consolidar versiones del modelo global.

Finalmente, las evidencias que debe aportar el proyecto incluyen un **diagrama de la arquitectura**, las **políticas de orquestación** (actualización, caída de nodos, seguridad y *hardening*), las actas del comité técnico y ético con las decisiones adoptadas y un informe de rendimiento federado por sitio y consolidado, con análisis de heterogeneidad y sesgos. Cuando aplique, debe constar el encaje regulatorio y la integración de los requisitos del MDR/IVDR y del AI Act en el expediente.

Alcance: habilitar entrenamiento/validación sin movimiento de datos crudos, preservando privacidad y garantizando gobernanza interinstitucional.

- **Arquitectura y orquestación:** definición de nodos, agregador y canal de comunicación seguro; establecimiento de políticas de actualización y de caída de nodos; aplicación de medidas de **aislamiento y refuerzo de seguridad del sistema (hardening)**, incluyendo la **desactivación de servicios innecesarios**, el **control de accesos** y la **aplicación de configuraciones seguras** en cada nodo.
- **Privacidad y robustez:** técnicas de agregación segura; opciones de privacidad diferencial (**differential privacy**) y defensa frente a ataques de inferencia/envenenamiento cuando sea pertinente.
- **Gobernanza y acuerdos:** acuerdos de corresponsabilidad/encargo; DTA y SLA; establecimiento de un **comité técnico** y un **comité ético** con **funciones definidas**, incluyendo la revisión y validación de resultados, las **reglas de actualización del modelo** y los **criterios de publicación**.
- **Evaluación federada:** métricas armonizadas; evaluación por centro y global; análisis de heterogeneidad y sesgos entre nodos.

Evidencias: diagrama de arquitectura; políticas de orquestación; actas del comité; informe de rendimiento federado por sitio y consolidado.

6.5. Documentación de datos y modelos (*DataSheet* y *ModelCard*)

La **documentación estructurada de los datos y de los modelos** es esencial para garantizar la **transparencia, la trazabilidad y el uso responsable** de los sistemas de IA en salud. Este proceso permite comprender el origen, las características y las limitaciones tanto de los conjuntos de datos como de los modelos, facilitando su evaluación por parte de equipos técnicos, clínicos y reguladores, y asegurando que las decisiones se basen en información verificable.

- ***DataSheet* (dataset):** debe describir el conjunto de datos utilizado, incluyendo su origen y periodo de recogida, los criterios de inclusión y exclusión aplicados, los procesos de limpieza y etiquetado, las variables y codificaciones empleadas, así como la representatividad del *dataset* respecto a la población diana. También debe detallar las limitaciones conocidas y las condiciones de acceso y uso, de forma que cualquier reutilización se realice bajo un marco controlado y conforme a la normativa aplicable.
- ***ModelCard* (modelo):** debe documentar el modelo entrenado, especificando el objetivo clínico y la finalidad prevista (*intended use*), la población diana, la versión y fecha de liberación, los datos y procesos de preprocesamiento utilizados y las métricas de rendimiento (incluyendo calibración). Además, debe reflejar los resultados por subgrupos relevantes, los riesgos identificados y las salvaguardas implementadas, así como los usos permitidos y prohibidos. Finalmente, debe incluir los requisitos de monitorización y mantenimiento, indicando cómo se controlará el rendimiento en producción y los criterios para su actualización.
- **Actualización y control de cambios:** es crítico mantener la coherencia entre datos, modelos y servicios desplegados. Debe garantizarse un **versionado sincronizado** entre *dataset* y modelo, con criterios objetivos para identificar cambios significativos (p. ej., impacto en seguridad, rendimiento o población diana). Además, debe mantenerse un **registro de versiones** que documente la correspondencia entre *dataset*, modelo y servicio desplegado, incluyendo **fecha, responsable, motivo del cambio y huella digital**. Asimismo, debe establecerse un **plan de comunicación** que asegure que usuarios y evaluadores reciban información clara y oportuna sobre cualquier modificación relevante.

Evidencias esperadas: fichas *DataSheet* y *ModelCard* completas, actualizadas, fechadas y firmadas por los responsables; registro de versiones que documente la correspondencia entre *dataset*, modelo y servicio desplegado, incorporando bitácora de cambios y huellas digitales (hashes) que permitan verificar integridad; repositorio controlado con permisos definidos y trazabilidad de publicaciones/commits; y un **procedimiento normalizado de operación (SOP)** que detalle criterios de aceptación, *checklist* de verificación previa a la liberación, procedimiento de *rollback* en caso de incidencias y responsables de cada fase del despliegue.

Salida esperada del apartado 6: *checklist* de verificación cumplimentada; plan de gestión de datos aprobado; informe de calidad y representatividad; evaluación de impacto en protección de datos (cuando proceda); artefactos de aprendizaje federado (si aplica); fichas *DataSheet* y *ModelCard* versionadas; registro de versiones sincronizado *dataset*–modelo–servicio (con bitácora de cambios y *hashes*); enlace a repositorio controlado con trazabilidad de publicaciones; y SOP de liberación y despliegue vigente.

7. Requisitos técnicos del sistema de IA

Definir requisitos verificables para el diseño, la validación, la explicabilidad, la monitorización y la operación segura del sistema de IA en salud, asegurando rendimiento clínicamente relevante, trazabilidad y mantenimiento controlado.

7.1. Diseño y validación del modelo

El **diseño y la validación del modelo** son etapas críticas para demostrar que un sistema de IA cumple su **finalidad prevista (intended use)** y que sus resultados son **fiables, seguros y clínicamente relevantes** en el contexto asistencial donde se aplicará. No basta con que el modelo funcione en un entorno de laboratorio: debe aportar **evidencia aplicable a la práctica clínica**, ser seguro para los pacientes y mostrar capacidad de **generalización** a diferentes poblaciones y organizaciones sanitarias.

Requisitos mínimos:

- **Especificación clínica y técnica:** definición precisa del objetivo clínico, la población diana y los criterios de valoración (*endpoints*), junto con los escenarios reales de uso. Deben explicitarse los supuestos y limitaciones del modelo para clarificar en qué condiciones es fiable y en cuáles no.
- **Diseño experimental sólido:** separación adecuada de conjuntos de entrenamiento, validación y prueba; prevención de fuga de información (*data leakage*); control del sobreajuste (*overfitting*); y uso de protocolos documentados y reproducibles (p. ej., validación cruzada).
- **Métricas relevantes:** además de precisión global, deben evaluarse sensibilidad, especificidad, F1, AUC-ROC, AUPRC, exactitud balanceada y métricas operativas como tiempos de respuesta o tasa de fallos. Siempre que sea posible, acompañadas de intervalos de confianza.
- **Calibración y utilidad clínica:** evaluación de la calibración global y por subgrupos (*Brier score*, ECE, intercepto y pendiente de calibración, *reliability diagrams*), con recalibración cuando proceda. La utilidad clínica debe demostrarse mediante análisis de curvas de decisión (*decision-curve analysis (net benefit)*) y, cuando aplique, con otras evidencias como impacto en decisiones, reducción de eventos evitables o análisis de umbrales y compensaciones (*trade-offs*) beneficio/daño. Umbrales y criterios de aceptación deben estar preespecificados.
- **Validación externa y generalización:** al menos una validación en centros o poblaciones distintas antes de cualquier uso asistencial, analizando transferibilidad por subgrupos relevantes (sexo, edad, comorbilidades, centro).
- **Gestión de sesgos y equidad:** medición del rendimiento diferencial por subgrupos, documentación de posibles sesgos de selección o información y aplicación de medidas de mitigación (reponderación, obtención de nuevos datos, ajustes en el modelo).
- **Robustez y pruebas de estrés:** evaluar y documentar el desempeño con datos faltantes o inconsistentes, ruido y degradación de señal, errores de codificación/mapeo semántico, escenarios fuera de distribución (*out-of-distribution, (OOD)*) y, cuando el riesgo lo justifique,

perturbaciones adversarias (*adversarial disruptions*); establecer umbrales de aceptación, tolerancias y medidas de mitigación/rollback.

- **Interpretabilidad y explicabilidad:** considerar desde el diseño qué grado de explicabilidad se necesita para cada rol (clínicos, equipos de datos, auditores), evitando trabajos adicionales posteriores y facilitando la aceptación clínica.
- **Reproducibilidad y trazabilidad:** control de versiones de datos, código y artefactos (incluyendo semillas aleatorias, dependencias y contenedores), de modo que cualquier auditoría o reevaluación pueda replicar resultados.
- **Plan de actualización y cambios significativos:** definición de criterios objetivos para identificar cambios que requieran nueva validación (p. ej., incorporación de nuevos datos que alteren la población diana, modificaciones en la arquitectura del modelo o pérdida relevante de calibración). Estos cambios deben estar trazados, comunicados a los usuarios y, si afectan a seguridad o finalidad prevista, sometidos a revisión ética o regulatoria.
- **Impacto clínico y eficiencia:** vincular la validación técnica con resultados clínicos intermedios o finales, cambios en decisiones, tiempos de proceso y beneficios potenciales en seguridad del paciente. Cuando corresponda, incluir valoración de coste-efectividad o impacto presupuestario.

Evidencias esperadas: protocolo de validación detallado; informe de rendimiento con intervalos de confianza; análisis por subgrupos y calibración; resultados de pruebas de robustez y estrés; cuaderno de reproducibilidad; SOP de evaluación; plan de actualización con criterios de cambio significativo; y, cuando aplique, plan para medir impacto clínico y eficiencia en fases posteriores.

7.2. Explicabilidad y trazabilidad

La **explicabilidad** y la **trazabilidad** son pilares fundamentales para garantizar que los sistemas de IA en salud sean **comprendibles, auditables y seguros**. La explicabilidad permite que los usuarios entiendan por qué el modelo genera una determinada predicción o recomendación, mientras que la trazabilidad asegura que cada decisión y cada versión del sistema puedan reconstruirse en caso de auditoría, incidente o revisión regulatoria. Un sistema que no puede explicar sus decisiones ni demostrar cómo y cuándo se modificó no debería considerarse apto para su uso en entornos asistenciales.

Requisitos mínimos:

- **Explicabilidad por rol:** debe adaptarse al perfil de usuario.
 - Los profesionales clínicos necesitan visualizar los factores más relevantes que han influido en el resultado de un caso concreto, junto con advertencias sobre limitaciones y, cuando sea posible, indicadores de confianza.
 - Los equipos de ciencia de datos requieren información técnica, como la importancia relativa de las variables, la estabilidad de las explicaciones y la sensibilidad del modelo a cambios en los datos.

- Los evaluadores y auditores necesitan comprender los supuestos, limitaciones y riesgos del modelo para verificar su adecuación al contexto clínico y regulatorio.
- **Técnicas y salvaguardas:** las explicaciones pueden obtenerse mediante modelos intrínsecamente interpretables o mediante métodos *posthoc* (p. ej., SHAP, LIME). En ambos casos es imprescindible validar la **estabilidad de las explicaciones** y advertir claramente sobre su ámbito de validez, evitando usos fuera de contexto.
- **Registro y trazabilidad:** cada inferencia debe quedar registrada con un **identificador único**, la **versión del modelo y del pipeline**, los **parámetros y umbrales** aplicados y la **entrada/salida** en formato anonimizado o representada mediante huellas digitales (hashes) que garanticen integridad sin exponer datos personales. También debe registrarse la **decisión final** y, cuando proceda, la **intervención humana asociada**, en línea con las exigencias de supervisión del AI Act.
- **Libro de cambios (changelog):** debe mantenerse una documentación cronológica de todas las modificaciones del modelo, incluyendo impacto esperado y resultados de verificación antes del despliegue. Este control de versiones debe vincularse a un repositorio seguro con permisos definidos y trazabilidad de publicaciones (*commits*), de manera que pueda saberse en todo momento qué versión estaba activa y bajo qué condiciones.

En conjunto, la explicabilidad y la trazabilidad no son elementos opcionales, sino requisitos esenciales para la confianza, la seguridad del paciente y el cumplimiento normativo. Un sistema que no puede explicar sus decisiones ni demostrar cómo y cuándo se modificó no debería considerarse apto para su uso en entornos asistenciales.

Evidencias esperadas: guía de explicabilidad diferenciada por rol; ejemplos documentados de casos explicados; política de *logging* y conservación de registros; registro de versiones con control de cambios y trazabilidad en repositorio seguro.

7.3. Monitorización, *drift* y recalibración

Una vez desplegado en un entorno asistencial, el sistema de IA debe someterse a una **supervisión continua**. La monitorización no es opcional: es la única forma de garantizar que el **rendimiento** y la **seguridad** se mantienen en condiciones reales, donde los datos, los procesos y la práctica clínica pueden diferir de los escenarios de validación inicial. Sin un seguimiento adecuado, el modelo puede degradarse con el tiempo, comprometiendo la seguridad del paciente y la fiabilidad de las decisiones.

Requisitos mínimos:

- **Plan de monitorización:** debe contemplar métricas clínicas (rendimiento global y por subgrupos como sexo, edad, comorbilidades, centro) y métricas de servicio (latencia, disponibilidad, tasa de errores, consumo de recursos). También debe registrarse la tasa de intervenciones manuales (*overrides* clínicos), que constituyen señales tempranas de falta de confianza en el sistema. Estos indicadores deben visualizarse en un tablero actualizado en tiempo real y revisarse periódicamente por un comité responsable.
- **Detección de *drift*:** uno de los principales riesgos en la fase operativa.

- **Data drift:** cambios en la distribución de las variables de entrada (p. ej., variaciones en protocolos clínicos o en la codificación de datos).
- **Concept drift:** cambios en la relación entre variables y resultados clínicos (p. ej., nuevas variantes de enfermedad que alteran correlaciones). Ambos fenómenos pueden degradar el rendimiento sin ser evidentes para los usuarios. Es necesario implementar mecanismos de detección basados en pruebas estadísticas y en el seguimiento de indicadores clave, con umbrales de alerta predefinidos y procedimientos claros de respuesta.
- **Recalibración y mantenimiento:** cuando se detecta degradación significativa, el sistema debe contar con procedimientos de recalibración estadística y actualización del modelo. Antes de liberar cualquier cambio, debe realizarse una **evaluación previa (pre-release)** en datos recientes y, cuando sea posible, en un entorno de **shadow mode**, donde el modelo actualizado se ejecuta en paralelo sin afectar a la práctica clínica hasta comprobar su fiabilidad.
- **Gestión de incidencias:** cada evento debe clasificarse por severidad, documentarse y resolverse conforme a acuerdos de nivel de servicio (**SLA**) y objetivos de nivel de servicio (**SLO**). Debe existir un procedimiento seguro de reversión (**rollback**) para restaurar la versión anterior en caso de fallo, así como un plan de comunicación para informar a usuarios y órganos de supervisión.
- **Cambios significativos:** deben definirse criterios objetivos para identificar modificaciones que requieren revalidación y, en su caso, nueva revisión ética o regulatoria. Ejemplos: cambios en la población diana, en la arquitectura del modelo o en la finalidad prevista.

Evidencias esperadas: tablero de monitorización en tiempo real; informe periódico de *drift* y rendimiento; actas de recalibración y actualizaciones; registro de incidencias con clasificación de severidad; y procedimiento normalizado de respuesta a incidencias (SOP) que incluya criterios de *rollback* y comunicación.

La monitorización continua, la detección temprana del *drift* y la recalibración no son tareas puntuales, sino procesos permanentes que garantizan la **seguridad del paciente**, la **confianza de los profesionales** y el **cumplimiento normativo**.

7.4. Seguridad del software y ciclo de vida (MLOps)

Garantizar la seguridad del software y un ciclo de vida controlado es clave para que los sistemas de IA en salud sean fiables, auditables y sostenibles. Este objetivo se alcanza aplicando buenas prácticas de ingeniería, ciberseguridad y cumplimiento normativo en todas las fases (desarrollo, pruebas, despliegue y operación), valorando amenazas específicas de IA (ataques adversariales, envenenamiento de datos, fuga o inferencia) y estableciendo defensas efectivas (detección de anomalías, validación de actualizaciones, SBOM, SAST/DAST y respuestas orquestadas desde el SOC). Para ello, se emplean prácticas de MLOps que integran automatización, trazabilidad y gobernanza en la gestión de modelos.

El ciclo de vida del software cumplirá **IEC 81001-5-1** y se coordinará con **IEC 62304** y **ISO 14971** (trazabilidad requisito ↔ riesgo ↔ control ↔ verificación). Se exigirá **SBOM por versión** (formatos **CycloneDX** o **SPDX**), gestión de vulnerabilidades (**CVE** con criterios **CVSS**), y una **política de parches y comunicación** de actualizaciones.

Requisitos mínimos:

- **Desarrollo seguro:** seguridad desde el diseño; control estricto de versiones; revisión sistemática de código; pruebas unitarias e integradas por liberación; pinning de dependencias; escaneos automáticos de vulnerabilidades; inventario de componentes (SBOM) actualizado.
- **Gestión del ciclo de vida de modelos (MLOps):** versionado de datos y modelos (dataset y model registry) con hashes y metadatos; trazabilidad de experimentos; **validación antes de promoción**; estrategias de despliegue **canary/shadow** con rollback automático; políticas de actualización del modelo con criterios de aceptación y reversión.
- **Entornos y despliegue:** segregación de entornos (desarrollo, pruebas, producción); contenedores inmutables para reproducibilidad; infraestructura como código; CI/CD con aprobaciones explícitas, controles automáticos y registro de auditoría de cada cambio.
- **Secretos y datos:** gestión segura de credenciales y claves; mínimo privilegio; cifrado robusto en tránsito y en reposo; políticas de copias y **pruebas periódicas de restauración** con **RPO/RTO** definidos y evidenciados.
- **Observabilidad:** métricas, logs y trazas centralizadas; detección de anomalías y alertas automáticas; **monitorización continua de rendimiento, calibración y deriva de datos/modelo** con umbrales y acciones definidas.
- **Ciberseguridad y robustez:** hardening del host y del runtime; desactivación de servicios innecesarios; actualizaciones de seguridad; pruebas de penetración proporcionales al riesgo; cuando proceda, **pruebas de robustez frente a manipulación de datos o modelos** (adversarial/poisoning).
- **Gestión y divulgación de vulnerabilidades:** proceso de gestión basado en **ISO 30111** y divulgación responsable conforme a **ISO 29147**; canal público (VDP); SLA de respuesta (**MTTD/MTTR**); criterios de notificación y comunicación a centros/usuarios; **coordinated vulnerability disclosure** y registro de incidencias integrado con el cuadro de mando de ciberresiliencia.
- **Documentación operativa:** SOP de despliegue y rollback; plan de continuidad de negocio; **matriz RACI** de roles y responsabilidades; programas de formación para usuarios y administradores; **model card** del sistema con limitaciones y escenarios no recomendados.

Evidencias esperadas

Paquete de despliegue versionado; **SBOM vigente**; resultados de pruebas unitarias, de integración y de seguridad (**SAST/DAST**); informes de ciberseguridad y pruebas de penetración; **registro de experimentos y validaciones** del modelo; evidencias de **canary/shadow** y rollback; informe de **escaneo de dependencias**; plan y evidencias de **hardening**; **monitorización de deriva/calibración** con umbrales;

acta de revisión de cambios conforme a **IEC 62304**; plan de continuidad; manual de operación; evidencias de copias y **pruebas de restauración**; registro VDP (casos y tiempos).

En conjunto, la seguridad del software y la gestión del ciclo de vida mediante MLOps no son tareas aisladas, sino procesos continuos que garantizan la fiabilidad, trazabilidad y resiliencia del sistema en entornos clínicos reales.

Salida esperada del apartado 7: protocolo de validación y su informe, guía de explicabilidad y política de **logging**, plan de monitorización con umbrales y procedimientos de recalibración, registro de cambios, **SOP de MLOps**, **SBOM** y evidencias de seguridad/ pruebas.

8. Interoperabilidad e integración en el entorno clínico

Garantizar que el sistema de IA se integra de forma segura, eficiente y mantenible con la **HCE**, sistemas de imagen (**PACS/RIS**), laboratorio, farmacia y otros componentes del entorno asistencial, usando estándares abiertos y patrones de integración robustos.

8.1. Requisitos funcionales de integración

Para que un sistema de IA aporte valor real en la práctica clínica, no basta con que funcione en entornos aislados: debe integrarse de forma **segura, eficiente y mantenible** con la historia clínica electrónica (HCE) y con el resto de sistemas asistenciales (PACS/RIS para imagen médica, LIS de laboratorio, farmacia, admisión, etc.). Esta integración debe apoyarse en **estándares abiertos** y patrones robustos que garanticen la **interoperabilidad, la trazabilidad y la sostenibilidad operativa**.

Requisitos mínimos:

- **Fuentes y consumo de datos:** identificación clara de los sistemas origen (HCE, LIS, RIS/PACS, farmacia, admisión) y de las entidades necesarias (episodios, problemas, diagnósticos, procedimientos, medicación, resultados analíticos, informes, imágenes y señales fisiológicas). Esta definición previene dependencias ocultas y asegura un flujo de datos completo y coherente.
- **Publicación de resultados:** las salidas deben generarse en formatos estructurados o semiestructurados (como JSON) y, cuando sea necesario, acompañarse de informes legibles para los profesionales. Deben vincularse al episodio/encuentro clínico correspondiente e incluir de forma explícita la **versión del modelo** y los **umbrales aplicados** en la inferencia, garantizando trazabilidad y seguridad del paciente.
- **Patrones de integración:** combinación de mecanismos síncronos y asíncronos.
 - Consultas bajo demanda: servicios **REST**.
 - Actualizaciones automáticas: mensajería/eventos (p. ej., **FHIR Subscriptions**, colas).
 - Soporte a la decisión clínica: **CDS Hooks** y **SMART on FHIR** para lanzar aplicaciones directamente desde la HCE, manteniendo la experiencia integrada en el flujo asistencial.

- **Latencia y disponibilidad:** definición de objetivos de nivel de servicio (**Service Level Objectives (SLO)**) por flujo (p. ej., inferencia < X s, disponibilidad \geq Y %). Deben contemplarse **modos degradados** que permitan continuar la atención clínica en caso de indisponibilidad del sistema, evitando interrupciones asistenciales.

Evidencias esperadas: diagramas de arquitectura lógica y de flujo; catálogo de interfaces, *endpoints* y eventos; especificaciones de tiempos objetivo y documentación técnica de integración.

8.2. Estándares y modelado semántico

La **interoperabilidad técnica y semántica** debe garantizar que los datos puedan **intercambiarse, comprenderse y reutilizarse de forma consistente** entre centros, proveedores y países, tanto para el **uso primario asistencial** como para el **uso secundario** en investigación, analítica avanzada, evaluación y espacios de datos federados. Este requisito es esencial en el marco del **Espacio Europeo de Datos Sanitarios (EHDS)** y su despliegue nacional, y debe cubrir datos clínicos, imagen, biobancos y genómica, junto con mecanismos de catalogación y gobernanza.

Requisitos mínimos:

- **Mensajería clínica y modelos de persistencia:**
 - **HL7 FHIR:** intercambio estructurado de recursos como demografía, episodios, observaciones, procedimientos, medicación y documentos clínicos; perfiles nacionales cuando existan.
 - **DICOM:** ingestión y publicación de imágenes y metadatos, incluyendo reportes estructurados (DICOM SR); referencia a series y estudios desde la HCE.
 - **IHE (XDS, XCA, MHD):** integración de sistemas heterogéneos mediante perfiles probados de intercambio documental.
 - **openEHR / ISO 13606:** modelos para persistencia clínica estructurada y trazabilidad semántica, con arquetipos reutilizables y separación clara entre modelo clínico y plantillas.
- **Terminologías y vocabularios clínicos:**
 - **SNOMED CT** para problemas y procedimientos.
 - **LOINC** para observaciones y laboratorio.
 - **ICD-10-ES** para clasificación administrativa.
 - **ATC / RxNorm** para fármacos, según el entorno.
 - **UCUM** para unidades de medida.
 - **Orphanet/ORDO** para enfermedades raras, **ICF** para funcionalidad y discapacidad.

Todas estas terminologías requieren mapeos consistentes, mantenidos y versionados, con validadores automáticos de conformidad.
- **Modelos para uso secundario y federación:**

- **OMOP Common Data Model (OHDSI):** armonización analítica en entornos multicéntricos, con ecosistema de cohorts y herramientas reutilizables; se recomienda documentar el proceso ETL y los puentes FHIR ↔ OMOP.
- **HealthDCAT-AP:** catalogación y descubrimiento de activos de datos, alineado con DCAT-AP europeo y principios FAIR.
- **MIABIS (BBMRI-ERIC):** descripción homogénea de colecciones, muestras y procesos en biobancos.
- **Genómica y datos ómicos:**
 - **GA4GH frameworks:** Phenopackets para fenotipos, htsget/refget para secuenciación, Beacon v2 para descubrimiento federado, DUO (Data Use Ontology) para condiciones de uso.
 - **HL7 FHIR Genomics:** integración clínica de hallazgos genéticos.
 - **VCF / HGVS:** representación estandarizada de variantes.

Estas especificaciones permiten interoperabilidad en entornos distribuidos y alinean genómica clínica con práctica asistencial.
- **Gobernanza de mapeos y versiones:** mantener tablas de equivalencia documentadas y versionadas, con criterios de compatibilidad y trazabilidad; registrar qué versión de terminología o perfil estaba en vigor en cada intercambio o análisis.

Evidencias esperadas: guía de perfiles FHIR/DICOM implementados; diccionarios terminológicos y mapeos documentados; modelos de datos aplicados (openEHR, OMOP, MIABIS, *FHIR Genomics*); validaciones de conformidad; y documentación de compatibilidad con el Reglamento del EHDS, TEHDAS y marcos europeos (*Beyond 1 Million Genomes*, GDI). Todo ello debe estar accesible en un **repositorio controlado**, con trazabilidad de versiones y *commits*.

8.3. Seguridad, identidad y control de acceso

La **protección de la identidad y la gestión segura de los accesos** son esenciales para garantizar la **confidencialidad, integridad y trazabilidad** en el uso de sistemas de IA en salud. Estos mecanismos no solo previenen accesos no autorizados, sino que también permiten demostrar cumplimiento normativo y reducir riesgos clínicos y legales.

Requisitos mínimos:

- **Autenticación y SSO:** integración con la infraestructura corporativa de gestión de identidades, de modo que los usuarios se autentiquen mediante mecanismos robustos alineados con las políticas del centro. Se recomienda el uso de protocolos estándar como **OpenID Connect (OIDC)** y **OAuth2**, que permiten autenticación federada y control granular de permisos. Las credenciales deben gestionarse con **tokens de acceso** de alcance mínimo necesario, sujetos a políticas de **rotación y expiración controladas**.

- **Autorización:** aplicar principios de segregación de funciones mediante modelos como **Role-Based Access Control (RBAC)** o **Attribute-Based Access Control (ABAC)**, asignando permisos según rol (clínico, investigador, técnico) o atributos contextuales. En situaciones excepcionales, puede habilitarse acceso mediante mecanismos de “**break-glass**”, siempre con registro y justificación de la acción.
- **Cifrado:** toda la información debe protegerse en **tránsito (TLS)** y en **reposo** mediante algoritmos robustos. La gestión de claves debe realizarse en entornos seguros, con rotación periódica y controles estrictos de acceso. La **segmentación de redes** y la **microsegmentación de servicios críticos** refuerzan la seguridad, reducen la superficie de ataque y limitan el impacto de incidentes.
- **Trazabilidad y auditoría:** cada acceso, consulta o inferencia debe registrarse en un sistema de **logging** que recoja quién realizó la acción, cuándo, sobre qué paciente o episodio y con qué versión del modelo. Los registros deben conservarse íntegros y confidenciales, protegidos frente a manipulación, y retenidos conforme a la normativa y políticas internas, garantizando disponibilidad para auditorías, revisiones regulatorias o investigaciones de incidentes.

Segmentación, identidades y telemetría de seguridad

El despliegue aplicará **segmentación de red por zonas/roles**, **autenticación fuerte** e **identidad federada** cuando proceda, junto con **telemetría centralizada** (logs, métricas y trazas) y **retención probatoria**. La configuración y operación deberán ser coherentes con el **SGSI (ISO/IEC 27001)** adoptado en el proyecto, y reportarán periódicamente al **cuadro de mando de ciberresiliencia (10.7)**.

Evidencias esperadas: matriz de permisos por rol; política de gestión de identidades (incluida MFA cuando aplique); evidencias de cifrado en tránsito y en reposo; configuración de segmentación/microsegmentación; ejemplos de registros de auditoría con versión de modelo y referencia a paciente/episodio; documentación de **políticas de retención y protección de logs**; y trazabilidad de revisiones de accesos y eventos **break-glass**.

En conjunto, la **seguridad, identidad y control de acceso** no son solo requisitos técnicos, sino condiciones indispensables para la **confianza, la protección de los derechos de los pacientes** y la **sostenibilidad de los proyectos de IA en entornos clínicos**.

8.4. Flujo clínico, usabilidad y seguridad del paciente

La integración de un sistema de IA en la práctica clínica no debe alterar el flujo asistencial ni aumentar la carga de trabajo de los profesionales. Por el contrario, debe insertarse de manera **natural en los procesos existentes**, aportando valor sin generar riesgos adicionales. El diseño debe garantizar que la IA se utilice en el **momento adecuado del recorrido asistencial**, apoyando la toma de decisiones clínicas sin interrumpir la dinámica del equipo.

Requisitos mínimos:

- **Análisis de flujo:** elaboración de diagramas BPMN u otras herramientas equivalentes para identificar puntos de entrada y salida de la IA, responsabilidades de cada actor y tiempos asociados. Esto permite anticipar cuellos de botella, definir responsabilidades y asegurar que el sistema se usa en el momento clínico adecuado.

- **Presentación de resultados:** los resultados deben mostrarse de forma clara, comprensible y contextualizada, incluyendo mensajes con unidades correctas, advertencias sobre limitaciones e indicadores de incertidumbre cuando proceda. Deben acompañarse de **enlaces a evidencia científica o guías clínicas**. El sistema debe permitir **override por parte del clínico**, con registro y justificación, garantizando la **supervisión humana efectiva** exigida por la normativa.
- **Prevención de fatiga de alertas:** priorización de notificaciones relevantes y mecanismos seguros de silenciamiento. Los umbrales de activación deben ser **configurables bajo gobernanza**, evitando tanto la sobrecarga de avisos como la omisión de alertas críticas.
- **Usabilidad y factores humanos:** el diseño debe ser **consistente con la HCE**, accesible y validado mediante pruebas con usuarios representativos antes del despliegue. Estas pruebas deben evaluar comprensión, facilidad de uso, impacto en la carga cognitiva y confianza en las recomendaciones.
- **Seguridad del paciente:** identificación de peligros potenciales asociados al uso del sistema y establecimiento de controles preventivos, como advertencias visibles, doble verificación en decisiones críticas y mecanismos de **fallback manual** en caso de fallo. Debe existir un procedimiento claro de **notificación y gestión de incidentes** para actuar con rapidez ante eventos adversos.

Evidencias esperadas: prototipos y capturas integradas en HCE, informe de pruebas de usabilidad, análisis de riesgos y plan de mitigación documentado, junto con registros de incidentes y acciones correctivas.

En conjunto, la integración de la IA en el flujo clínico debe **mejorar la calidad asistencial sin comprometer la seguridad ni la eficiencia**. Un sistema que no se adapta al contexto real o que introduce riesgos adicionales no debería considerarse apto para su uso en entornos sanitarios.

8.5. Entorno de ejecución y despliegue

Definir **dónde se ejecuta el sistema y cómo se mantiene operativo** es esencial para garantizar la **seguridad, eficiencia y continuidad del servicio**. El entorno de ejecución debe minimizar riesgos, optimizar el rendimiento y permitir una gestión controlada de cambios y actualizaciones, asegurando que el sistema cumple las exigencias clínicas y regulatorias.

Requisitos mínimos:

- **Topología:** especificar si el sistema se ejecutará en **on-premises**, en la **nube** o en un **modelo híbrido**, ya que esta decisión condiciona la latencia, la seguridad y la escalabilidad. Las **rutas de datos** deben documentarse con diagramas claros, identificando zonas de seguridad (DMZ, VPC, VPN) que protegen los flujos de información. En el caso de imágenes médicas o señales en tiempo real, la proximidad a **PACS** y **HCE** es clave para reducir latencia. Cuando sea necesario, puede recurrirse a **inferencias en el borde (edge computing)**, ejecutando el modelo cerca del origen de los datos para evitar transferencias de gran volumen y reducir tiempos de respuesta.
- **Contenerización y CI/CD:** los artefactos deben ser **inmutables**, firmados digitalmente y acompañados de un **SBOM (Software Bill of Materials)** actualizado. Los procesos de

integración y despliegue continuo (CI/CD) deben incluir pruebas automáticas, escaneo de seguridad, aprobaciones explícitas y despliegues seguros (**blue/green o canary releases**) que permitan validar nuevas versiones sin interrumpir el servicio. Debe existir un procedimiento probado de **rollback** para restaurar la versión anterior en caso de fallo.

- **Observabilidad:** el sistema debe contar con **métricas, logs y trazas centralizadas**, con identificadores de correlación para seguir cada transacción. Los paneles de monitorización deben mostrar el estado de colas, tiempos de respuesta y errores, y activar **alertas automáticas** vinculadas a los **SLO (Service Level Objectives)** definidos por flujo.
- **Capacidad y rendimiento:** anticipar escenarios de alta demanda mediante un **dimensionamiento adecuado de recursos**, colas con mecanismos de **backpressure, autoscaling** para CPU/GPU y **pruebas de carga y resiliencia** que incluyan métricas **p95/p99**, esenciales en entornos clínicos para garantizar tiempos de respuesta incluso en condiciones extremas. Asimismo, deben realizarse pruebas de **failover** para comprobar la continuidad del servicio en caso de fallo.

Evidencias esperadas: diagrama de despliegue con rutas y zonas de seguridad; pipelines CI/CD documentados; resultados de pruebas de carga, resiliencia y recuperación ante fallos; y **SOP de operación** actualizado con procedimientos de despliegue, *rollback* y continuidad de negocio.

En conjunto, un **entorno de ejecución bien diseñado** asegura la **estabilidad, seguridad y escalabilidad** del sistema, facilita su mantenimiento y refuerza su **alineación con los requisitos regulatorios y de calidad** en salud.

8.6. Pruebas de interoperabilidad y verificación en entorno clínico

Antes de que un sistema de IA se utilice con pacientes, es imprescindible comprobar que la **integración funciona exactamente como se ha diseñado**. Estas pruebas no son un trámite, sino una **garantía de seguridad y fiabilidad**, y deben repetirse de forma periódica durante la operación para asegurar que los cambios o actualizaciones no introducen riesgos.

Requisitos mínimos:

- **Datos sintéticos y de prueba:** uso de conjuntos realistas para pruebas unitarias e integración. Cuando se empleen datos reales, deben anonimizarse para proteger la privacidad. Estas pruebas iniciales permiten detectar errores en la lógica sin comprometer información sensible.
- **Pruebas de extremo a extremo:** validación de escenarios clínicos representativos, que incluyan tanto casos habituales como situaciones límite y errores deliberados. El objetivo es verificar que el sistema responde correctamente en todas las condiciones previstas y que los fallos se gestionan de forma segura, sin interrumpir la atención asistencial.
- **Entorno shadow mode:** ejecución del sistema en paralelo al flujo clínico real, sin impacto en las decisiones médicas, antes de la activación. Esto permite evaluar rendimiento y estabilidad en condiciones reales, detectar posibles problemas y afinar la configuración sin poner en riesgo la seguridad del paciente.

- **Homologación por versiones:** cada nueva versión del modelo o servicio debe someterse a un proceso formal de homologación, con **criterios claros de aceptación**, verificación de **invariantes funcionales** y documentación de resultados. Este control evita que cambios menores introduzcan errores clínicamente relevantes.

Evidencias esperadas: plan y resultados de pruebas documentados, *checklist* de homologación por versión y actas de decisiones **go/no-go** que respalden la autorización del despliegue.

En conjunto, las pruebas de interoperabilidad y verificación establecen un **marco de confianza** para la evolución del sistema: cada paso hacia el uso asistencial debe estar respaldado por **evidencias documentadas y revisiones formales**, garantizando un despliegue seguro y sostenible en entornos clínicos.

8.7. Operación multicéntrica y portabilidad

Para que un sistema de IA pueda escalar más allá del centro donde se desarrolló, es necesario garantizar que su despliegue en otros hospitales sea **seguro, eficiente y trazable**. La portabilidad no consiste en copiar un modelo, sino en asegurar que las **interfaces, los flujos clínicos y las dependencias técnicas** se adapten a entornos heterogéneos sin comprometer la interoperabilidad ni la seguridad del paciente.

Requisitos mínimos:

- **Contratos de interoperabilidad:** documentos que especifiquen cómo se comunican los sistemas (interfaces, formatos de datos, estándares y reglas semánticas). Deben ser **estables y versionados**, con **compatibilidad hacia atrás** para evitar que una actualización rompa integraciones en hospitales que aún no han migrado.
- **Portabilidad de modelos:** empaquetado en formatos estándar (**ONNX, contenedores con servicio de inferencia**) que permitan ejecución en distintos entornos sin modificar el código. Debe incluirse la documentación completa de **dependencias** (librerías, versiones, *frameworks*) y un **Software Bill of Materials (SBOM)** actualizado para auditorías de seguridad, alineado con normativas como **NIS2** y el **ENS**.
- **Configurabilidad local:** mecanismos que permitan ajustar parámetros específicos de cada centro (terminologías, umbrales clínicos, flujos asistenciales) **sin necesidad de reentrenar** el modelo, siempre que sea posible. Esto reduce costes de implantación y asegura coherencia con la finalidad prevista.
- **Catálogo de sitios:** registro de los despliegues en cada hospital, documentando perfiles implementados, mapeos terminológicos, excepciones aplicadas, trazabilidad de versiones y resultados de validación local. Este catálogo facilita auditorías y reconstrucción del historial de despliegue.
- **Validación distribuida:** en escenarios de **aprendizaje federado** o validaciones multicéntricas, se recomienda realizar pruebas específicas por nodo, analizando la heterogeneidad de datos y el rendimiento por subgrupos, para demostrar que el sistema mantiene seguridad y eficacia en entornos diversos.

Evidencias esperadas: paquete portable del modelo, guía de despliegue multicentro, catálogo de sitios con registros de validación y diferencias documentadas, SBOM actualizado y actas de homologación local.

En conjunto, la **operación multicéntrica y la portabilidad** no son tareas accesorias, sino condiciones necesarias para la **sostenibilidad y escalabilidad** de los sistemas de IA en salud. Un modelo que no puede trasladarse de forma controlada y segura difícilmente alcanzará impacto real en el **Sistema Nacional de Salud**.

8.8. Gobierno de la integración y acuerdos

La integración de un sistema de IA en el ecosistema clínico no termina con su despliegue inicial. Para que sea **sostenible en el tiempo**, es necesario establecer un **marco de gobierno** que asegure la **continuidad del servicio, la gestión ordenada de cambios y la coordinación entre todas las partes implicadas**. Sin este marco, cualquier actualización técnica, cambio organizativo o incidencia puede convertirse en un riesgo para la **seguridad del paciente** y para la **interoperabilidad**.

Requisitos mínimos:

- **Comité técnico de integración:** órgano responsable de supervisar la operación del sistema, priorizar mejoras y resolver incidencias. Debe incluir perfiles clínicos, técnicos y de seguridad, reunirse con periodicidad definida y hacer seguimiento de los **SLA (Service Level Agreements)** y **SLO (Service Level Objectives)** relativos a disponibilidad, tiempos de respuesta y calidad del servicio.
- **Contratos y anexos:** formalización de acuerdos que definan las responsabilidades de cada parte. Deben incluir **DTA (Data Transfer Agreements)**, contratos de **corresponsabilidad o encargo del tratamiento**, y anexos específicos sobre **seguridad, disponibilidad y plan de continuidad**. Este marco contractual garantiza cumplimiento normativo, trazabilidad y protección institucional.
- **Gestión de cambios:** cada modificación (técnica, funcional u organizativa) debe planificarse en **ventanas de mantenimiento** que minimicen el impacto asistencial. Es necesario comunicar los cambios a los usuarios, ofrecer formación cuando afecten a la práctica y actualizar la documentación técnica y clínica.

Evidencias esperadas: actas del comité técnico de integración, acuerdos firmados, plan de continuidad aprobado, registros de cambios comunicados y documentación actualizada tras cada modificación.

En conjunto, el **gobierno de la integración** no es un trámite administrativo, sino un **mecanismo esencial para mantener la seguridad, la interoperabilidad y la confianza** en el sistema a lo largo de todo su ciclo de vida. Un proyecto que carece de este marco corre el riesgo de volverse insostenible y de comprometer su adopción.

Salida esperada del apartado 8: catálogo de interfaces y perfiles estándar, mapeos terminológicos, política de seguridad y auditoría, prototipos integrados en HCE, plan de pruebas y resultados, diagrama de despliegue y SOP de operación, así como acuerdos firmados y un marco de gobierno de la integración activo y trazable.

9. Evaluación y madurez: criterios y baremación

La evaluación de proyectos de IA en salud debe ser **objetiva, verificable y proporcional al nivel de madurez tecnológica**. Por ello, esta guía establece criterios diferenciados para dos tipologías:

- **PI (proyectos de investigación y desarrollo, TRL 1–3).**
- **ICPS (investigación clínica con producto sanitario, estudios regulados, TRL ≥4).**

La evaluación combina **cinco dimensiones clave**:

1. **Calidad científico-técnica**, que asegura la solidez metodológica y la relevancia clínica.
2. **Cumplimiento ético y regulatorio**, imprescindible para la protección de derechos y la seguridad del paciente.
3. **Interoperabilidad e integración**, que determina la viabilidad técnica en entornos asistenciales reales.
4. **Operación segura y ciberresiliencia (ciberseguridad)**, que incluye monitorización, gestión de riesgos y trazabilidad, protección frente a amenazas, gestión de vulnerabilidades, control de accesos, hardening y configuración segura, así como continuidad de negocio y recuperación ante incidentes.
5. **Valor clínico y económico**, que orienta la sostenibilidad y la adopción en el Sistema Nacional de Salud.

La **puntuación total** es de **36 puntos para PI y 48 puntos para ICPS**, con **criterios eliminatorios** que impiden la aprobación cuando faltan elementos esenciales (p. ej., base jurídica conforme al RGPD/LOPDGDD, mitigación de riesgos para la seguridad del paciente o, en ICPS, ausencia de controles básicos de ciberseguridad/ciberresiliencia y validación externa previa al uso asistencial). Solo se puntuán **evidencias documentadas**; las afirmaciones sin soporte no son evaluables.

La categoría **“Viable y financiable”** se reserva a aquellos proyectos que, además de alcanzar la máxima puntuación relativa, demuestran un encaje realista con los recursos disponibles. El resto de los proyectos que superen los umbrales mínimos podrán ser considerados **“Viables”**, manteniendo un valor científico y clínico significativo, aunque sin prioridad de financiación.

Umbrales orientativos:

- **Viable:** ≥ 24 puntos en PI y ≥ 28 puntos en ICPS.
- **Viable y financiable:** ≥ 30 puntos en PI y ≥ 40 puntos en ICPS.

Estos umbrales son orientativos y pueden **ajustarse según la convocatoria y la disponibilidad de recursos**.

9.1. Metodología de evaluación (común PI/ICPS)

Escala por ítem (0–3): 0 = no evidencia; 1 = insuficiente o especulativo; 2 = adecuado con mejoras pendientes; 3 = sólido y verificable.

Evidencias exigibles: documentos, protocolos, resultados, tablas y anexos (no se puntuán afirmaciones sin soporte). Para ciberresiliencia (ciberseguridad), se valorarán, cuando proceda, políticas y controles, registro y auditoría de eventos, gestión de vulnerabilidades, pruebas de resistencia, y plan de respuesta y recuperación ante incidentes.

Cálculo por bloques: media de ítems del bloque × peso del bloque (puntos).

Criterios eliminatorios (knock-out):

- Ausencia de **base jurídica** y medidas de **RGPD/LOPDGDD** cuando proceda.
- Riesgos no mitigados para **seguridad del paciente**.
- En **ICPS**: falta de **validación externa** previa al uso asistencial o inexistencia de **plan de monitorización**.
- En proyectos que califican como **PS**: ausencia de **gestión de riesgos** y documentación regulatoria mínima.
- **Ciberresiliencia insuficiente en relación con el riesgo/TRL** (p. ej., ausencia de controles básicos, falta de registro y auditoría, inexistencia de plan de respuesta a incidentes o de gestión de vulnerabilidades).

9.2. Baremación para PI (TRL 1–3) — 36 puntos

Bloque	Puntos	Contenido evaluable (ejemplos)
1. Diseño científico y relevancia clínica	8	Problema y endpoints clínicos; población diana; justificación y estado del arte; hipótesis y beneficios esperados.
2. Datos: calidad y representatividad	10	Procedencia y linaje; diccionario y metadatos; limpieza; missingness ; representatividad; plan de sesgos y subgrupos; controles de acceso a datasets y segregación de entornos (desarrollo/prueba) proporcionales al riesgo.
3. Metodología y rendimiento	8	Particiones y control de fuga; métricas (incl. calibración); validación interna; reproducibilidad (versionado/semillas); robustez y pruebas de estrés (datos faltantes, OOD, degradación de señal y, si procede, perturbaciones adversarias).
4. Ética y protección de datos	4	Base jurídica; DPIA/EIPD (si aplica); minimización; ciberresiliencia proporcional al riesgo/TRL: control de accesos y privilegios mínimos, registro y auditoría, gestión básica de vulnerabilidades, cifrado en tránsito/en reposo cuando proceda.
5. Plan de madurez y transferencia	4	Hoja de ruta a TRL 4–6 ; validación externa planificada; criterios de “cambio significativo”; plan inicial de ciberresiliencia para la transición a ICPS (hardening, pruebas de resistencia/penetración programadas y continuidad/recuperación)

6. Innovación y alineación estratégica	2	Novedad, transferencia potencial, encaje con estrategias del SNS/ISCIII.
--	---	---

Umbrales orientativos (PI, 36 pts):

- **No viable:** < 24 puntos.
- **Viable:** 24–29 puntos.
- **Viable y financiable:** ≥ 30 puntos.

Observaciones PI: para pasar a **ICPS** se espera al menos **validación externa** inicial, análisis de sesgos por subgrupos y **DPIA** actualizada. Además, evidencias de **ciberresiliencia** básicas (políticas y controles, logging/auditoría y plan de respuesta a incidentes) acordes al riesgo.

9.3. Baremación para ICPS (TRL ≥4) — 48 puntos

Bloque	Puntos	Contenido evaluable (ejemplos)
1. Valor clínico y seguridad del paciente en investigación	10	Relevancia clínica y <i>endpoints</i> del estudio; salvaguardas del protocolo; procedimientos de <i>override</i> ; gestión de incidentes y reporte; medidas de ciberresiliencia orientadas a seguridad del paciente (detección, contención y recuperación ante incidentes; registro y auditoría)
2. Regulación y conformidad (MDR/IVDR + AI Act + CEIm/AEMPS)	10	Cualificación y clase (si PS); GSPR y gestión de riesgos; plan de evaluación clínica/desempeño; autorizaciones CEIm y, cuando proceda, AEMPS; encaje AI Act (alto riesgo); controles de ciberseguridad conforme a MDR Anexo I (seguridad/performancia), gestión de vulnerabilidades y referencias aplicables (p. ej., IEC 62304/81001-5-1, ISO/IEC 27001) cuando proceda
3. Interoperabilidad e integración para el estudio	8	Integración con HCE/FHIR/DICOM necesaria para el protocolo; trazabilidad por sujeto/episodio; pruebas <i>end-to-end</i> en entorno de investigación; usabilidad para investigadores/clinicians; seguridad en la integración (autenticación/autorización, cifrado en tránsito, segregación de entornos y hardening), con pruebas <i>end-to-end</i> que incluyan casos de seguridad.
4. Monitorización, MLOps y control de cambios	8	Métricas del estudio y del servicio; <i>drift</i> y recalibración (si aplica durante el estudio); <i>changelog</i> y bloqueo de versiones para protocolo; CI/CD controlado; observabilidad y telemetría de seguridad (logging/auditoría centralizada), gestión de vulnerabilidades y dependencias (incl. SBOM), y checks de seguridad en el pipeline CI/CD.
5. Datos y equidad	6	Calidad/representatividad del <i>dataset</i> del estudio; análisis por subgrupos; mitigación de sesgos; <i>DataSheet/ModelCard</i> ;

		controles de acceso y privilegios mínimos, cifrado en reposo/cuando proceda y segregación de datos sensibles.
6. Sostenibilidad y escalabilidad	6	Recursos del estudio (soporte, seguridad, formación); plan de transición a posestudio (PMS/PMCF o despliegue asistencial); gobernanza multicéntrica si procede; plan de ciberresiliencia y continuidad (backup/restore, ejercicio de respuesta a incidentes, responsabilidades y escalado)

Umbrales orientativos (ICPS, 48 pts):

- **No viable:** < 28 puntos.
- **Viable:** 28–39 puntos.
- **Viable y financiable:** ≥ 40 puntos.

Los umbrales pueden ajustarse por convocatoria y disponibilidad de recursos.

9.4. Procedimiento de evaluación y actas

La evaluación de proyectos debe realizarse con **criterios homogéneos, trazabilidad completa y participación multidisciplinar**, garantizando que las decisiones sean **objetivas, reproducibles y transparentes**.

Requisitos mínimos:

- **Panel y roles:** el comité de evaluación debe incluir, como mínimo, un **clínico del área, un especialista en datos/IA, un perfil experto en interoperabilidad y un perfil regulatorio o de ética**. Se recomienda añadir un **perfil de ciberseguridad/ciberresiliencia**. Todos los miembros deben firmar una **declaración de conflictos de interés** para asegurar la imparcialidad.
- **Ejecución:** el proceso se basa en la **revisión documental** y, cuando proceda, en una **defensa breve** del equipo solicitante. Cada evaluador puntuá de forma independiente utilizando la escala definida en la guía; posteriormente, los resultados se discuten en **sesión conjunta** para alcanzar consenso. La revisión deberá cubrir explícitamente aspectos de ciberresiliencia (controles y políticas, gestión de vulnerabilidades, logging/auditoría y plan de respuesta/recuperación ante incidentes), con hallazgos y riesgos clasificados por criticidad.
- **Checklists y trazabilidad:** se emplean **listas de verificación** (anexos) y se registra la evidencia que respalda cada puntuación. Las actas deben conservar las **versiones revisadas y la hoja de cálculo con el detalle de puntuaciones**, asegurando que cualquier auditoría pueda reconstruir el proceso. Se mantendrá un registro de hallazgos de ciberseguridad y su trazabilidad (riesgo, responsable, fecha objetivo y estado), así como de pruebas de resistencia cuando proceda.
- **Condicionalidades:** el panel puede proponer **condiciones de financiación** vinculadas a plazos y responsables, como completar una validación externa, cerrar hallazgos de seguridad o aportar un acuerdo multicéntrico. En materia de ciberresiliencia, podrán exigirse acciones como hardening, corrección de vulnerabilidades críticas, ejercicio de respuesta a incidentes o

evidencia de pruebas de penetración/estrés proporcionales al riesgo. Estas condiciones deben documentarse en el acta.

- **Desempate:** en caso de empate, se priorizan proyectos con **mayor TRL, mejor impacto clínico documentado, mejor relación coste-impacto y menor riesgo regulatorio**, maximizando así el valor y la viabilidad de la inversión pública. A igualdad del resto, se priorizará menor riesgo operativo/ciberseguridad y mayor madurez de controles

Evidencias esperadas: actas firmadas del comité de evaluación, listas de verificación cumplimentadas, hoja de puntuación final desglosada por bloques y registro de condiciones de financiación (cuando existan), informe de revisión de ciberresiliencia con hallazgos y plan de remediación, y registro de auditoría/logs de evaluación.

En conjunto, este procedimiento no solo define cómo puntuar, sino que establece un **marco de gobernanza** que refuerza la **transparencia, la rendición de cuentas y la confianza** en las decisiones de evaluación y financiación, **incluida la ciberresiliencia**.

9.5. Resultados y documentación de la evaluación

Al finalizar el proceso de evaluación, es fundamental generar una **documentación clara, completa y trazable** que refleje las decisiones adoptadas, las condiciones establecidas y los elementos necesarios para el seguimiento. Esta documentación no solo sirve como registro interno, sino que también aporta **transparencia y auditabilidad** en revisiones posteriores.

Requisitos mínimos:

- **Informe resumido:** debe incluir la **puntuación obtenida por bloques, el umbral alcanzado** (no viable / viable / viable y financiable) y un análisis de las **principales fortalezas y debilidades** del proyecto. Incluirá de forma explícita el estado de **ciberresiliencia** (controles aplicados, hallazgos y riesgos abiertos) cuando proceda.
- **Plan de acciones:** cuando el proyecto se sitúe en el umbral de viabilidad o presente aspectos críticos, debe elaborarse un plan que defina **hitos concretos, responsables asignados y plazos definidos** para elevar la madurez del proyecto y facilitar su avance hacia fases posteriores. En materia de ciberresiliencia, el plan recogerá remediaciones priorizadas (p. ej., hardening, gestión de vulnerabilidades, logging/auditoría, pruebas de resistencia/recuperación) con fechas objetivo y criterios de aceptación.
- **Registro de seguimiento:** debe mantenerse un registro actualizado para **reevaluaciones periódicas**, que permita comprobar el cierre de acciones pendientes y evaluar el impacto de cambios significativos. El registro deberá trazar también las acciones de ciberresiliencia y su estado (abierta/en curso/cerrada), conservando evidencia de verificación. Este registro garantiza que las condiciones impuestas se cumplen antes de autorizar nuevas fases o financiación definitiva.

Evidencias esperadas: informe de evaluación firmado, plan de acciones (cuando proceda), registro actualizado de seguimiento y reevaluación, anexo de ciberresiliencia con hallazgos, plan de remediación y evidencias de cierre cuando aplique.

En conjunto, estos resultados y documentos son esenciales para demostrar la **trazabilidad, equidad y transparencia** en la toma de decisiones, reforzando la confianza en el proceso de evaluación y en la priorización de recursos públicos, incluida la ciberresiliencia

Salida esperada del apartado 9: Acta firmada del comité de evaluación con decisión y observaciones (PI o ICPS). Hoja de puntuación por bloques (PI: 36; ICPS: 48) con criterios eliminatorios aplicados y justificación de las puntuaciones. Condiciones de financiación y/o requisitos de subsanación previos al avance (plazos, responsables y evidencias requeridas). Registro de seguimiento: plan de acciones, hitos y trazabilidad, con referencia a los **formularios de evaluación PI/ICPS** y a las evidencias anexas.

10. Implementación práctica y gestión del cambio

Asegurar una implantación ordenada, segura y medible del sistema de **IA** en el entorno asistencial, minimizando riesgos clínicos y organizativos y maximizando la adopción efectiva.

10.1. Gobernanza y liderazgo del cambio

La implantación de un sistema de IA en el entorno asistencial no puede abordarse como un simple despliegue técnico. Requiere una **estructura clara de gobernanza** que defina quién toma las decisiones, cómo se gestionan los riesgos y qué mecanismos aseguran la coordinación entre los distintos actores. Sin este marco, la implantación puede generar incertidumbre, resistencias y riesgos para la **seguridad del paciente**.

Requisitos mínimos:

- **Patrocinio y roles:** contar con un **patrocinio clínico y directivo** que respalde el proyecto y lo alinee con los objetivos estratégicos del centro. Deben designarse roles clave: un **product owner** como responsable funcional, un **responsable de seguridad del paciente**, el **Delegado de Protección de Datos (DPD)**, referentes de **TI e interoperabilidad**, y perfiles de **calidad y regulación**. Esta estructura asegura decisiones equilibradas entre criterios clínicos, técnicos y normativos.
- **Comité de implantación:** órgano de supervisión que revisa hitos, riesgos y dependencias, y adopta decisiones críticas de avance o retroceso (**go/no-go**). Debe reunirse con periodicidad definida y mantener **actas documentadas** de acuerdos y acciones.
- **Matriz RACI:** formalización de responsabilidades para cada tarea, especificando quién es **responsable (Responsible)**, quién aprueba (**Accountable**), quién debe ser consultado (**Consulted**) y quién informado (**Informed**). Esta herramienta es esencial para áreas críticas como formación, pruebas, soporte, comunicación y monitorización.
- Definir escalado de incidentes: **CSIRT autonómico (si existe)** → **INCIBE-CERT** (privado/sanitario no-AGE) o **CCN-CERT** (sector público/AGE) → **punto de contacto único** nacional → **ENISA/Red CSIRTs y Reserva UE** cuando proceda.

Seguridad y enlace corporativo:

El proyecto designará un **responsable de seguridad del proyecto** con autoridad operativa y enlace funcional con el director de Seguridad de la Información (**Chief Information Security Officer (CISO)**), corporativo. Sus funciones incluyen: asegurar el cumplimiento de políticas y marcos aplicables, coordinar evaluaciones de riesgo y planes de tratamiento, validar controles técnicos/organizativos antes de cada hito, supervisar la gestión de incidentes y el reporte a los **CSIRT** de referencia cuando proceda, y mantener el **RACI** de seguridad actualizado (propietarios, responsables, consultados e informados). se documentará la cadena de escalado y la sustitución en caso de ausencia.

Evidencias esperadas: actas del comité de implantación, **matriz RACI** actualizada, registro de riesgos con responsables asignados y un **plan de comunicación** que detalle cómo se informará a los distintos grupos implicados.

En conjunto, la **gobernanza y el liderazgo del cambio** no son accesorios, sino factores determinantes para una implantación **segura, ordenada y aceptada por los profesionales**. Un proyecto sin estructura clara corre el riesgo de fracasar, incluso si la tecnología es sólida.

10.2. Análisis de preparación e impacto organizativo

Antes de implantar un sistema de IA en un entorno asistencial, es imprescindible comprender cómo afectará a los **procesos, las cargas de trabajo y los roles profesionales**. Este análisis permite anticipar riesgos, planificar medidas de mitigación y asegurar que la tecnología se integra de forma **ordenada, eficiente y segura**.

Requisitos mínimos:

- **Mapeo de procesos (BPMN):** elaboración de diagramas que identifiquen los **puntos de entrada y salida** de la IA, los **traspasos de responsabilidad (hand-offs)** y los **tiempos asociados** a cada tarea. Esto permite detectar cuellos de botella y rediseñar flujos para mantener la continuidad asistencial.
- **Carga asistencial y tiempos:** estimación de las **variaciones en la duración de tareas** derivadas del uso del sistema. Un modelo que aumente los tiempos sin aportar valor añadido puede generar rechazo; por ello, deben definirse estrategias que preserven la eficiencia.
- **Impacto en roles y resistencias:** análisis de cómo la introducción de la IA afecta a las responsabilidades de cada perfil profesional. Es fundamental identificar **adoptadores clave**, detectar **barreras culturales u organizativas** y diseñar medidas de acompañamiento (formación, sensibilización, participación en la toma de decisiones) que favorezcan la aceptación.
- **Políticas clínicas:** definición de reglas claras sobre **responsabilidades, override con justificación** y procedimientos para la **documentación de decisiones en la HCE**. Estas políticas son necesarias para garantizar **supervisión humana efectiva** y cumplir con los requisitos normativos de trazabilidad.

Evidencias esperadas: diagrama actualizado de procesos asistenciales, informe de impacto en tiempos y cargas de trabajo, y plan de mitigación que incluya medidas organizativas, formativas y de comunicación para facilitar la adopción segura del sistema.

En conjunto, el análisis de preparación e impacto organizativo constituye una herramienta clave para **minimizar resistencias, asegurar la eficiencia y reforzar la confianza** de los profesionales en la implantación de sistemas de IA en salud.

10.3. Formación, acreditación y soporte a usuarios

La implantación de un sistema de IA en el entorno asistencial solo será efectiva si los profesionales comprenden su **funcionamiento, sus limitaciones y las salvaguardas** que lo acompañan. La formación no es un complemento, sino un requisito esencial para garantizar la **seguridad del paciente** y la **confianza en la herramienta**.

Requisitos mínimos:

- **Plan por perfiles:** se establece **formación obligatoria, acreditable y periódica** en ciberseguridad para todo el personal clínico-técnico del proyecto, con registro de completado. El plan incluirá, como mínimo: higiene digital, gestión segura de credenciales, **phishing/deepfakes y su verificación**, protección de datos en entornos asistenciales e investigación, **gestión y comunicación de incidentes** (incluidos simulacros), y buenas prácticas en IA/MLOPS (SBOM, control de versiones, manejo de datos y modelos). se impartirán **sesiones de reciclaje anuales** y formación *ad-hoc* tras cambios relevantes o incidentes, con **evaluación final y trazabilidad** de resultados.
- **Modalidades:** combinación de talleres prácticos, simulaciones con casos reales, ayudas en contexto y manuales rápidos que sirvan como referencia. Este enfoque mixto facilita tanto la curva de aprendizaje inicial como la consulta durante la práctica asistencial.
- **Acreditación y refresco:** registro formal de la **superación de la formación**, con **sesiones de reciclaje periódicas** para mantener competencias actualizadas, especialmente cuando se produzcan cambios en el sistema.
- **Soporte:** establecimiento de un canal operativo para **incidencias y dudas**, con tiempos de respuesta definidos, complementado por un **catálogo de FAQs** que resuelva consultas frecuentes y reduzca la incertidumbre en el uso clínico.

Evidencias esperadas: plan de formación aprobado, materiales utilizados, registros de asistencia, resultados de evaluaciones, acreditaciones individuales y documentación del canal de soporte activo. En conjunto, la **formación, acreditación y soporte a usuarios** son piezas críticas para garantizar una adopción **segura, eficaz y sostenible** de los sistemas de IA en salud.

10.4. Pilotos y despliegue por fases

Antes de extender un sistema de IA a todo un entorno asistencial, es fundamental **reducir riesgos mediante una validación incremental**. Los pilotos permiten comprobar el comportamiento del sistema en condiciones reales, identificar problemas tempranos y ajustar procesos antes de un despliegue completo.

Requisitos mínimos:

- **Selección de unidad piloto:** elección basada en criterios objetivos como la **relevancia del caso de uso**, el **volumen de actividad**, el **compromiso del equipo clínico** y la capacidad para garantizar la **trazabilidad de datos y resultados**. Una unidad bien elegida facilita la obtención de evidencias sólidas y la detección temprana de barreras.
- **Shadow mode (TRL 6):** ejecución inicial en paralelo al flujo clínico, sin influir en decisiones médicas. Esta modalidad permite evaluar **métricas de rendimiento, seguridad y usabilidad** sin comprometer la atención al paciente, validando la integración técnica y la experiencia de usuario.
- **Criterios de avance:** definición de umbrales mínimos de **rendimiento, seguridad y aceptación profesional**. El avance a fases posteriores debe documentarse mediante un acta formal de decisión **go/no-go**, asegurando trazabilidad y transparencia.
- **Escalado controlado:** despliegue progresivo por **servicio, turno o centro**, acompañado de **retroalimentación estructurada** y de un **plan de mitigación** que resuelva incidencias antes de continuar con la expansión.

Evidencias esperadas: protocolo de piloto aprobado, informe de resultados, acta de decisión de avance y plan de escalado con cronograma, responsables y criterios de éxito.

En conjunto, los **pilotos y el despliegue por fases** constituyen un mecanismo clave para garantizar una **implantación segura, eficiente y aceptada** por los profesionales antes del escalado completo del sistema de IA en salud.

10.5. Gestión de cambios técnicos y versiones

Una vez que el sistema de IA está en operación, es inevitable que se produzcan **cambios**: actualizaciones del modelo, ajustes de umbrales, mejoras funcionales o correcciones de errores. Sin embargo, cualquier modificación no controlada puede introducir riesgos para la **seguridad del paciente, la interoperabilidad o la trazabilidad**. Por ello, la gestión de cambios debe ser un **proceso estructurado, documentado y auditible**.

Requisitos mínimos:

- **Versionado visible en la HCE:** cada inferencia debe registrar de forma explícita la **versión del modelo y los umbrales aplicados**, garantizando trazabilidad para auditorías, revisiones regulatorias y análisis de incidentes.
- **Cambio significativo:** definir criterios objetivos que determinen qué modificaciones requieren revalidación (p. ej., incorporación de nuevos datos, alteración de parámetros críticos, cambios de arquitectura del modelo). En estos casos, debe aplicarse un **proceso formal de revalidación** previo al despliegue, dentro de una **ventana de mantenimiento** planificada.
- **Rollback seguro:** contar con un **procedimiento probado de reversión**, que permita restaurar la versión anterior de forma rápida y controlada si la actualización genera fallos o resultados inesperados. Este *rollback* debe ir acompañado de **comunicación a los usuarios** y registro de la incidencia.

- **Sandbox y homologación:** todas las actualizaciones deben probarse previamente en un entorno de *sandbox/homologación*, con una *checklist de aceptación* que verifique funcionalidad, seguridad y compatibilidad con flujos clínicos. Solo tras superar estas pruebas se autoriza el despliegue en producción.

Evidencias esperadas: registro de cambios (*changelog*) que documente cada modificación; informes de verificación por versión; y registros completos de **despliegue y rollback**.

En conjunto, la **gestión estructurada de cambios técnicos y versiones** asegura que la evolución del sistema se realice de forma **segura, transparente y conforme a la normativa**, preservando la confianza de profesionales y pacientes.

10.6. Seguridad del paciente y ética en operación

La implantación de un sistema de IA en entornos asistenciales debe regirse por un principio fundamental: **prevenir daños y garantizar un uso responsable**. La tecnología no sustituye la responsabilidad clínica, sino que la complementa. Por ello, es necesario establecer controles que aseguren la **seguridad del paciente y el respeto a los principios éticos**.

Requisitos mínimos:

- **Análisis de peligros y controles:** identificación de riesgos potenciales y definición de medidas preventivas, incluyendo advertencias claras sobre limitaciones, definición de **límites de uso** y mecanismos de **doble verificación** en decisiones críticas.
- **Gestión de incidentes:** establecimiento de un procedimiento para **clasificar eventos por severidad**, investigar su **causa raíz**, implementar **acciones correctivas** y comunicar los incidentes a los órganos competentes y, cuando proceda, a los usuarios afectados, siguiendo la normativa aplicable.
- **Supervisión humana efectiva:** el sistema debe permitir que el profesional pueda realizar un **override** de cualquier recomendación, con trazabilidad y justificación documentada. Este control asegura que la **decisión final siempre recae en el equipo clínico**.
- **Transparencia:** los profesionales deben recibir información clara sobre el **funcionamiento, limitaciones y criterios** del sistema. Cuando proceda, también debe informarse al paciente sobre el uso de IA en su atención, en línea con los principios de **autonomía y consentimiento informado**.

Evidencias esperadas: plan de seguridad aprobado; registro actualizado de incidentes con análisis y resolución; y **auditorías periódicas** que verifiquen la aplicación de controles y la efectividad de las medidas correctivas.

En conjunto, la seguridad del paciente y la ética en operación constituyen la **base de confianza** necesaria para la adopción sostenible de sistemas de IA en salud.

10.7. Indicadores, resultados y cuadro de mando

La implantación de un sistema de IA no termina con su despliegue. Para garantizar su **valor, seguridad y sostenibilidad**, es necesario medir de forma continua su **adopción**, su **rendimiento clínico** y su **impacto organizativo**. Este seguimiento permite identificar desviaciones, corregir problemas y demostrar el **retorno de la inversión**.

Requisitos mínimos:

- **Adopción y uso:** medir cobertura del sistema, frecuencia de utilización, tasa y motivos de **override**, así como la **satisfacción de los usuarios**. Estos indicadores reflejan la aceptación del sistema y su integración en la práctica clínica.
- **Calidad y rendimiento clínico:** monitorizar métricas operativas y clínicas, incluida la **calibración del modelo**, con análisis desagregado por subgrupos. Esto permite asegurar que el sistema mantiene su fiabilidad en distintos contextos.
- **Equidad:** evaluar diferencias de rendimiento y uso entre **sexo, edad, centro, comorbilidad** u otros subgrupos relevantes. Detectar y corregir desigualdades es esencial para cumplir con principios éticos y regulatorios.
- **Eficiencia:** medir tiempos de proceso, **trabajos adicionales evitados** y mejoras en **productividad**, datos clave para justificar la inversión y planificar la escalabilidad.

Evidencias esperadas: cuadro de mando actualizado con indicadores clave; informes periódicos (p. ej., mensuales); y un **plan de acciones derivado de las métricas**, que permita corregir desviaciones y mejorar el desempeño del sistema.

En conjunto, el uso de indicadores y cuadros de mando refuerza la **transparencia**, la **mejora continua** y la **confianza** en los sistemas de IA en salud, asegurando que aportan un valor tangible y equitativo al sistema asistencial.

10.7.1. Cuadro de mando de Ciberresiliencia

El proyecto medirá y hará seguimiento, como mínimo, de los siguientes indicadores: **MTTD** (tiempo medio de detección), **MTTR** (tiempo medio de recuperación), **RTO/RPO** efectivos, **% de restauraciones exitosas**, **n.º de simulacros/año**, **n.º de incidentes notificados** y **% de endpoints legacy segmentados**.

Definiciones operativas:

- **MTTD:** Tiempo medio desde la ocurrencia del incidente hasta su detección (fuente: SOC/SIEM).
- **MTTR:** Tiempo medio desde la detección hasta la completa recuperación del servicio (fuente: gestor de incidentes/OTRS/Jira).
- **RTO/RPO:** Objetivo de tiempo de recuperación y de punto de recuperación realmente alcanzados en pruebas y en incidentes (fuente: runbooks y registros de recuperación).
- **% de restauraciones exitosas:** Restauraciones verificadas/total de restauraciones ensayadas en el periodo.
- **N.º de simulacros/año:** Ejercicios documentados de respuesta ante ciberincidentes.

- **N.º de incidentes notificados:** Incidentes reportados al CSIRT/autoridad competente conforme a procedimiento.
- **% de endpoints legacy segmentados:** Endpoints sin soporte o con soporte limitado situados en segmentos/zonas aisladas sobre el total de endpoints legacy.

Gobernanza del KPI:

- **Periodicidad:** Al menos mensual (trimestral para la revisión ejecutiva).
- **Fuentes de datos:** SOC/SIEM, gestor de incidencias, plataforma de copias, CMDB/inventario.
- **Responsable del dato:** Responsable de seguridad del proyecto (enlace con CISO).
- **Umbral y tendencia:** Definir objetivo, umbrales de alerta y acciones de mejora cuando se superen.

Con este marco de indicadores y gobernanza, el proyecto asegura **evidencia objetiva, trazabilidad y capacidad de reacción** para sostener el valor clínico, la seguridad y la mejora continua del sistema de IA a lo largo de todo su ciclo de vida.

10.8. Sostenibilidad operativa y continuidad

Un sistema de IA no debe considerarse implantado con éxito hasta que se garantice su **operación estable y segura a lo largo del tiempo**. La sostenibilidad implica disponer de recursos, procesos y herramientas que aseguren la **continuidad del servicio** incluso ante incidencias, cambios tecnológicos o caídas de proveedores.

Requisitos mínimos:

- **Soporte y mantenimiento:** definición de **acuerdos de nivel de servicio (SLA)** y **objetivos de nivel de servicio (SLO)**, así como mecanismos de **guardia y escalado** para resolver incidencias críticas en plazos predefinidos. Estos acuerdos son esenciales para mantener la confianza de los usuarios y la seguridad del paciente.
- **Continuidad de negocio:** implementación de **copias de seguridad periódicas**, definición de objetivos de **punto de recuperación (RPO)** y **tiempo de recuperación (RTO)**, y realización de **pruebas regulares de restauración**. Debe existir un plan específico para **escenarios de caída de proveedores**, que contemple alternativas técnicas y procedimientos de contingencia.
- **MLOps y monitorización:** integración de pipelines de **CI/CD**, sistemas de **observabilidad** (métricas, logs, trazas) y **alarmas automáticas** que permitan detectar anomalías de forma temprana y asegurar la resiliencia operativa.
- **Gestión del conocimiento:** mantenimiento de un **repositorio actualizado** que incluya **procedimientos normalizados (SOP)**, manuales operativos y un registro de **lecciones aprendidas**, accesible para todos los roles implicados en la operación y mantenimiento del sistema.

Evidencias esperadas: plan de continuidad aprobado, resultados de pruebas de recuperación, **bitácora operativa** con registro de incidencias y acciones correctivas, y repositorio actualizado con documentación técnica y procedimental.

En conjunto, la **sostenibilidad operativa y la continuidad** son condiciones indispensables para asegurar que los sistemas de IA en salud mantienen su **valor clínico, seguridad y fiabilidad** a lo largo del tiempo.

10.9. Participación de pacientes y comunicación

La **participación del paciente** no debe limitarse a recibir información, sino que debe convertirse en un **elemento activo del ciclo de vida** del sistema de IA. Involucrar a los pacientes desde fases tempranas mejora la **usabilidad**, refuerza la **confianza** y otorga legitimidad ética al proyecto.

La implicación de pacientes en los proyectos de IA y Big Data en salud debe trascender la mera consulta y orientarse hacia un verdadero **codiseño**. Para ello, se recomienda incorporar metodologías participativas estructuradas (p. ej., *design thinking*, talleres colaborativos, *world café*) en las fases iniciales de definición de objetivos, asegurando que las soluciones desarrolladas respondan a necesidades reales y socialmente aceptadas.

Asimismo, se sugiere la utilización de **herramientas estandarizadas de evaluación de impacto percibido**, como cuestionarios de experiencias (PREMs) y resultados reportados por pacientes (PROMs), adaptados a tecnologías de IA. Estos instrumentos permiten valorar aspectos clave como la confianza en el sistema, la comprensibilidad de los resultados o la percepción de equidad y no discriminación.

Finalmente, se recomienda la creación de **mecanismos de gobernanza inclusiva**, como comités asesores de pacientes integrados en los proyectos, con funciones de revisión y aportación de recomendaciones. Esta participación estructurada fortalece la legitimidad ética y social de los desarrollos, y contribuye a anticipar y mitigar riesgos relacionados con la aceptación y el impacto real en la calidad de vida de las personas.

Requisitos mínimos:

- **Codiseño y participación temprana:** inclusión de representantes de pacientes en **fases de diseño, pruebas piloto y comités de seguimiento**, promoviendo un enfoque de codiseño que integre su perspectiva en la toma de decisiones y facilite la identificación temprana de barreras.
- **Transparencia reforzada:** elaboración de **materiales informativos claros, accesibles y comprensibles** que expliquen la finalidad, beneficios y limitaciones del sistema, los **datos que utiliza**, los **riesgos potenciales** y las **salvaguardas implementadas**. Estos materiales deben estar alineados con los principios de **explicabilidad del AI Act**.
- **Accesibilidad y equidad:** adaptación de los materiales a diferentes niveles de **alfabetización digital y lingüística**, en formatos inclusivos (lectura fácil, audio, braille, traducciones), garantizando la equidad y evitando discriminación, en línea con el **EHDS** y las recomendaciones internacionales.
- **Ética y consentimiento informado:** cuando el uso de IA pueda influir en decisiones clínicas, debe proporcionarse información suficiente y documentar el consentimiento del paciente, incluyendo la **posibilidad de revocación**. Incluso cuando no sea obligatorio, informar sobre el uso de IA refuerza la autonomía y la confianza del paciente.

- **Feedback bidireccional y trazabilidad:** habilitar mecanismos de retroalimentación más allá de encuestas puntuales, como PROMs, PREMs y canales de sugerencias. Debe existir un proceso documentado para **analizar, priorizar y responder** a estas aportaciones, con trazabilidad de las acciones derivadas para demostrar mejoras concretas.

Evidencias esperadas: materiales informativos accesibles; actas que acrediten la participación de pacientes en comités o pilotos; resultados de encuestas y otros mecanismos de feedback; y un **registro de acciones de mejora** implementadas a partir de sus aportaciones.

En conjunto, la participación activa y la comunicación transparente con los pacientes fortalecen la **dimensión ética, la inclusión y la confianza social**, y aseguran que la implantación de la IA en salud esté alineada con las **recomendaciones internacionales** y los **marcos regulatorios europeos**.

10.10. Cierre, transferencia y retirada

El **final del ciclo de vida** de un sistema de IA en salud no debe abordarse como un simple trámite técnico, sino como una fase crítica para garantizar la **seguridad, la trazabilidad y la preservación del valor generado**. Una retirada desordenada puede comprometer la continuidad asistencial, la protección de datos y la reutilización de evidencias.

Requisitos mínimos:

- **Criterios de retirada:** definición de criterios objetivos como pérdida de rendimiento clínico, aparición de riesgos para la seguridad del paciente, obsolescencia tecnológica o decisiones estratégicas de la organización. Estos criterios deben estar documentados y aprobados por el comité de gobernanza del sistema.
- **Plan de salida:** elaboración de un plan estructurado que contemple la **migración o archivo seguro** de datos, modelos y artefactos asociados, garantizando la **trazabilidad completa**. El plan debe incluir la preservación de documentación técnica, versiones del modelo, registros de inferencias y la **eliminación segura de entornos y credenciales** para prevenir accesos no autorizados.
- **Evaluación final:** análisis de los **resultados clínicos y económicos** alcanzados, comparándolos con los objetivos iniciales, y recopilación de **lecciones aprendidas** que puedan compartirse en foros científicos y técnicos. Esta evaluación aporta transparencia, mejora continua y sostenibilidad al ecosistema de IA en salud.
- **Aspectos regulatorios:** en estudios regulados (**ICPS**), el cierre debe incluir la **revisión final por el CEIm** y el **envío del informe de cierre**, conforme al **RD 192/2023** y al **MDR**. Este proceso confirma el cumplimiento del protocolo y la revisión de incidentes. En proyectos no regulados no es obligatorio, pero se recomienda como buena práctica para reforzar la transparencia y la trazabilidad.
- **Comunicación:** informar del cierre a todos los actores implicados (profesionales, pacientes, órganos de supervisión), explicando motivos, impacto y medidas adoptadas para garantizar la **seguridad y la continuidad asistencial**.

Evidencias esperadas: plan de retirada aprobado; acta de cierre firmada por el comité responsable; documentación de migración o archivo; registro de eliminación segura; e **informe de evaluación post-implantación** con resultados, lecciones aprendidas y recomendaciones para futuros proyectos.

Salida esperada del apartado 10: plan de implantación aprobado, protocolo de piloto y criterios de avance, materiales de formación, matriz **RACI**, tablero de indicadores, plan de continuidad y seguridad, y procedimientos de gestión de cambios y retirada. Además, **plan de participación y comunicación con pacientes**, con evidencias de codiseño, materiales accesibles y reportes **PREMs/PROMs**.

11. Sostenibilidad, escalabilidad y modelo económico

Asegurar la viabilidad económica y operativa del sistema de IA a lo largo del tiempo, planificando costes, recursos, contratos y métricas que permitan su mantenimiento, escalado multicéntrico y creación de valor clínico y organizativo.

11.1. Modelo económico y financiación

La sostenibilidad de un sistema de IA en salud no depende solo de su rendimiento técnico, sino también de su **viabilidad económica**. Un modelo sin un **caso de negocio sólido** corre el riesgo de quedar inoperativo tras la fase piloto, incluso si aporta valor clínico. Por ello, es imprescindible definir desde el inicio cómo se financiará, qué retorno se espera y qué riesgos pueden comprometer su continuidad.

Requisitos mínimos:

- **Caso de negocio:** descripción del **problema clínico y organizativo** que se aborda, los **beneficios esperados**, las **alternativas existentes** y los **riesgos potenciales** que pueden afectar a la adopción.
- **Horizonte y perspectiva:** definición del **periodo de análisis** y de la perspectiva adoptada (hospital, servicio de salud o SNS), considerando distintos **escenarios de adopción progresiva**.
- **TCO y ROI/ICER:** cálculo del **Total Cost of Ownership (TCO)** para estimar el coste total del ciclo de vida, del **Return on Investment (ROI)** para medir el retorno esperado y, cuando proceda, del **Incremental Cost-Effectiveness Ratio (ICER)** junto con un análisis de impacto presupuestario.
- **Análisis de sensibilidad:** construcción de escenarios **optimista, central y pesimista**, variando parámetros críticos como volumen de uso, costes de infraestructura, licencias o recursos humanos, con el fin de anticipar riesgos y márgenes de maniobra.
- **Ciberseguridad:** el proyecto incluirá una **partida presupuestaria específica y trazable** para ciberseguridad. El porcentaje se justificará según el **análisis de riesgos** y la **exposición asistencial** (uso clínico, criticidad de procesos y datos, interconexiones). Se desglosará en **CAPEX/OPEX** y se garantizará la **trazabilidad del gasto** con hitos, evidencias y responsable económico. Los informes de **IBM** sitúan el **coste medio por brecha en salud** en torno a **9,8–10,9 M USD** por incidente, por lo que **presupuestar mitigación y respuesta es crítico**.
- **Financiación y compra:** identificación de las **fuentes de financiación** (presupuesto propio, convocatorias públicas, compra pública de innovación, cofinanciación) y de las **modalidades de**

pago (licencia, suscripción, por uso, por resultado). Este plan debe alinearse con un **cronograma** que vincule hitos técnicos y financieros.

Evidencias esperadas: documento de caso de negocio; hoja de cálculo con supuestos y análisis de sensibilidad; plan de financiación detallado; y cronograma que conecte fases técnicas y necesidades económicas.

En conjunto, el modelo económico y de financiación debe garantizar la **viabilidad, sostenibilidad y escalabilidad** del sistema de IA, permitiendo su integración estable en la práctica clínica y en la estrategia del SNS.

11.2. Costes y recursos (CAPEX/OPEX)

Un sistema de IA en salud no es un gasto puntual, sino una inversión a lo largo de todo su ciclo de vida que combina costes iniciales (**CAPEX**) y costes recurrentes (**OPEX**). Identificar y desglosar estos costes es esencial para garantizar la sostenibilidad, la escalabilidad y la continuidad operativa, evitando interrupciones por falta de recursos.

Requisitos mínimos:

- **CAPEX (costes de capital):** incluyen la integración técnica, la infraestructura necesaria (servidores, GPU, almacenamiento), el etiquetado y la curación de datos, el desarrollo o la parametrización del sistema y las pruebas y certificaciones. Aunque se concentran en fases iniciales, condicionan la capacidad de escalado futuro.

Ciberseguridad (CAPEX):

- Auditorías técnicas y pruebas de intrusión previas a producción, con remediación verificada;
- Hardening inicial y configuración segura de entornos on-prem/nube (IAM, cifrado, microsegmentación, gestión de secretos);
- Puesta en marcha de SOC/SIEM (licencias/activación) y de la observabilidad (métricas, logs, trazas);
- Establecimiento del pipeline CI/CD con controles de seguridad (SAST/DAST), SBOM y análisis de dependencias (SCA);
- Definición y pruebas iniciales de RTO/RPO y del plan de recuperación.

- **OPEX (costes operativos):** abarcan las licencias, el uso de servicios en la nube, el soporte y el mantenimiento, las prácticas de MLOps y la monitorización, la ciberseguridad, la formación continua y el refresco de competencias, los seguros y las auditorías periódicas. Son recurrentes y deben contemplarse para todo el ciclo de vida del sistema.

Ciberseguridad (OPEX):

- Operación del SOC/SIEM y respuesta orquestada a incidentes.
- Monitorización continua y gestión de vulnerabilidades (parcheo, seguimiento de SBOM, SCA).
- Copias de seguridad, verificación periódica de restauraciones y revisión de RTO/RPO.

- Auditorías periódicas y pruebas de intrusión regulares con cierre de hallazgos.
- Formación obligatoria y reciclajes con registro de completado.
- Servicios de respuesta a incidentes (peritaje, preservación de evidencias, notificación y comunicación).
- **Costes de datos:** comprenden la extracción, la normalización, los mapeos terminológicos, la anonimización/seudonimización y la gobernanza de los datos. Estos procesos son críticos para garantizar la calidad, la reproducibilidad y el cumplimiento normativo.
- **Recursos humanos:** planificación de dedicaciones de perfiles clínicos, de ciencia de datos, TI, seguridad, calidad/regulación y gestión del cambio. La disponibilidad de personal cualificado es un factor crítico para la operación sostenida y la mejora continua.

Evidencias esperadas: presupuesto detallado **CAPEX/OPEX** (con línea específica de ciberseguridad), matriz de responsables asignados y plan de aprovisionamiento que vincule recursos a las fases del ciclo de vida.

En conjunto, la correcta planificación de costes y recursos asegura que la inversión en IA se traduzca en impacto clínico real, continuidad operativa y sostenibilidad económica en el tiempo.

11.3. Escalabilidad y despliegue multicéntrico

Para que un sistema de IA genere un **impacto real en el Sistema Nacional de Salud**, debe poder **extenderse a otros servicios y centros con el mínimo trabajo adicional posible**. Este objetivo exige planificar la **portabilidad técnica**, gestionar la **heterogeneidad de datos** y establecer una **gobernanza sólida** entre instituciones.

Requisitos mínimos:

- **Portabilidad técnica:** el modelo y los servicios asociados deben entregarse en **formatos estándar** (contenedores, ONNX cuando aplique), acompañados de scripts y documentación que faciliten el despliegue en distintos entornos. El sistema debe ofrecer **parametrización local** para adaptar terminologías, umbrales clínicos y flujos asistenciales sin necesidad de reentrenar. Además, se requiere la documentación detallada de los **perfiles FHIR y DICOM** utilizados para garantizar la interoperabilidad.
- **Heterogeneidad de datos:** cada hospital puede presentar variaciones en **codificaciones, procesos y calidad de datos**. Es necesario establecer una **estrategia de adaptación** mediante mapeos y validaciones de transferencia que aseguren que el modelo mantiene su **rendimiento y calibración** en entornos diversos. Estas pruebas deben incluir análisis por subgrupos y validación local antes del despliegue operativo.
- **Gobernanza y acuerdos:** la operación multicéntrica debe apoyarse en **acuerdos formales** que definan responsabilidades, **niveles de servicio (SLA/SLO)**, mecanismos de soporte y un **comité técnico multicéntrico**. Estos acuerdos deben contemplar la protección de datos mediante **DTA** y, cuando corresponda, contratos de **corresponsabilidad o encargo del tratamiento**, así como planes de continuidad.

- **Aprendizaje federado:** cuando se emplee, la arquitectura y la gobernanza deben permitir **entrenar y validar modelos sin mover datos crudos**, preservando la privacidad y cumpliendo el RGPD. Esto requiere definir nodos, protocolos de actualización, medidas de seguridad y métricas armonizadas para evaluar el rendimiento global y por centro.

Evidencias esperadas: guía de despliegue multicentro; contrato tipo con cláusulas de gobernanza y protección de datos; registro de centros participantes con sus perfiles y mapeos; y resultados de validación local y consolidada.

En conjunto, la **escalabilidad y el despliegue multicéntrico** son condiciones necesarias para que los sistemas de IA en salud alcancen un **impacto sostenible y equitativo** en el SNS, evitando duplicidades y asegurando la interoperabilidad entre instituciones.

11.4. Sostenibilidad operativa y continuidad

Un sistema de IA no debe considerarse implantado con éxito hasta que se garantice su **operación estable, segura y trazable a lo largo del tiempo**. La sostenibilidad implica disponer de **recursos, procesos y herramientas** que aseguren la **continuidad del servicio**, incluso ante incidencias, cambios tecnológicos o caídas de proveedores.

Requisitos mínimos:

- **Plan de servicio:** definición de **niveles de servicio (SLA)** y **objetivos de nivel de servicio (SLO)**, así como de las **ventanas de mantenimiento**, procedimientos de **soporte** y mecanismos de **escalado de incidencias**. Estos acuerdos son esenciales para mantener la confianza de los usuarios y la seguridad del paciente.
- **Continuidad de negocio:** respaldo mediante **copias de seguridad periódicas**, definición de objetivos claros de **punto de recuperación (RPO)** y **tiempo de recuperación (RTO)**, y realización de **pruebas regulares de restauración**. Debe incluirse un **plan específico para escenarios de caída de proveedores**, con alternativas técnicas y procedimientos de contingencia.
- **Evolución controlada:** establecimiento de una **política de versiones**, criterios objetivos para identificar **cambios significativos**, procedimientos de **homologación por versión** y mecanismos de **rollback seguro** en caso de fallo. Cada actualización debe probarse en un **entorno controlado** antes de su despliegue en producción.
- **Gestión del conocimiento:** mantenimiento de un **repositorio documental actualizado** que incluya **procedimientos normalizados (SOP)**, manuales operativos, registros de cambios y **lecciones aprendidas**, accesible a todos los roles implicados en la operación y el mantenimiento del sistema. Esto garantiza la continuidad incluso ante rotación de personal o relevo en roles críticos.

Evidencias esperadas: plan de continuidad aprobado; resultados de pruebas de recuperación; **registro de cambios (changelog)** documentado; y repositorio actualizado con manuales, SOP y documentación técnica y procedimental.

En conjunto, la sostenibilidad operativa a largo plazo es una condición indispensable para que los sistemas de IA en salud aporten un **valor clínico y organizativo real**, manteniendo la **confianza de usuarios y pacientes** en el tiempo.

11.5. Evaluación económica y resultados en salud

La adopción de un sistema de IA en salud debe justificarse no solo por su **viabilidad técnica**, sino también por el **valor clínico y económico** que aporta. Evaluar estos aspectos de forma rigurosa es clave para fundamentar decisiones de **financiación, escalado y sostenibilidad**.

Requisitos mínimos:

- **Métricas clínicas y de uso:** inclusión de **resultados clínicos relevantes** (p. ej., reducción de eventos adversos, mejora diagnóstica), indicadores de **seguridad del paciente**, utilización de recursos y tiempos de proceso. Debe monitorizarse también la **adopción real del sistema**, la **tasa de anulaciones (override)** y la experiencia de profesionales y pacientes mediante **PROMs** y **PREMs**.
- **Evaluación económica:** análisis de **coste-efectividad** o **coste-utilidad** (p. ej., **ICER**) y, cuando proceda, **impacto presupuestario** para la adopción a escala. El análisis debe considerar **costes directos** (infraestructura, licencias, mantenimiento) e **indirectos** (tiempo clínico, formación, gestión del cambio), en comparación con beneficios esperados en salud y eficiencia operativa. Se recomienda realizar **análisis de sensibilidad** con escenarios optimista, central y pesimista.
- **Equidad:** evaluación de la distribución de beneficios y resultados entre **subgrupos relevantes** (sexo, edad, centro, comorbilidad), identificando posibles desigualdades en el acceso o en el rendimiento. Este análisis responde tanto a principios éticos como a requisitos emergentes en marcos regulatorios y de evaluación de tecnologías sanitarias.
- **Benchmarking y transparencia:** siempre que sea posible, comparar resultados con **estándares clínicos** o con otros sistemas, y publicar hallazgos para favorecer la **transparencia** y la **reutilización de evidencias**.

Evidencias esperadas: protocolo de evaluación aprobado; informes periódicos con resultados clínicos, económicos y de adopción; análisis de sensibilidad documentado; métricas de experiencia de usuario y paciente; y anexos que recojan supuestos, cálculos y resultados por subgrupos.

En conjunto, la **evaluación económica y de resultados en salud** es esencial para demostrar que un sistema de IA no solo funciona técnicamente, sino que **aporta valor real, equitativo y sostenible** al sistema sanitario.

11.6. Contratación, propiedad intelectual y licencias

Cuando un proyecto entra en fase de **operación o escalado**, es imprescindible formalizar contratos con proveedores tecnológicos (software, nube, mantenimiento). El objetivo es garantizar la **continuidad del servicio, la protección de datos y el cumplimiento normativo** (RGPD/LOPDGDD, MDR/IVDR, AI Act),

evitando riesgos como la dependencia excesiva de un proveedor (*vendor lock-in*) o la falta de trazabilidad.

Un **marco contractual claro** es esencial para reducir riesgos legales, garantizar la sostenibilidad y proteger los derechos de todas las partes. La falta de definición en aspectos como la **titularidad de los datos**, el **acceso a evidencias** o las **condiciones de salida** puede comprometer la operación y la seguridad del paciente.

Requisitos mínimos:

- **Derechos y titularidad:** el contrato debe definir con precisión la titularidad y condiciones de uso de los **datos, modelos, código, documentación y mejoras derivadas**. Debe incluir cláusulas que garanticen el **acceso a logs y trazas** para auditorías regulatorias y de seguridad.
- **Licencias y uso:** especificar las **condiciones de uso** (número de usuarios, instancias, territorios), los **límites de uso** y las **obligaciones asociadas**. Deben incorporarse **acuerdos de nivel de servicio (SLA/SLO)** y penalizaciones en caso de incumplimiento, para asegurar calidad, disponibilidad y continuidad.
- **Portabilidad y salida:** garantizar el **derecho de uso continuado** en caso de rescisión, la **entrega de artefactos y datos derivados**, y un **plan de transición** con asistencia técnica. Se recomienda incluir mecanismos como **escrow de código y modelos** para escenarios de quiebra o cese de actividad del proveedor.
- **Cumplimiento:** reflejar las obligaciones derivadas de **RGPD/LOPDGDD** (protección de datos, seguridad, notificación de incidentes), y, cuando aplique, el **alineamiento con MDR/IVDR y AI Act** para sistemas de alto riesgo. El contrato debe contemplar también la **subcontratación y transferencias internacionales**, exigiendo autorización previa y garantías adecuadas.
- **Compatibilidad y mejoras:** definir la propiedad intelectual sobre **mejoras y derivados**, e incluir la obligación de mantener **compatibilidad con estándares abiertos** que faciliten la portabilidad técnica y reduzcan riesgos de bloqueo tecnológico.
- **Conflictos de interés con proveedores:** cuando el proveedor haya participado previamente en actividades de investigación, co-creación o pruebas de concepto relacionadas con el sistema de IA, deben **establecerse procedimientos específicos para la identificación y gestión** de posibles conflictos de interés (declaración formal de vínculos, trazabilidad de las interacciones previas y evaluación independiente de las ofertas), de modo que la transición desde la colaboración científica hasta la contratación quede documentada y separada funcionalmente.

Evidencias esperadas: contrato y anexos tipo; **matriz de riesgos contractuales**; **checklist de cumplimiento normativo**; y, cuando proceda, copia del **acuerdo de escrow**.

En conjunto, un marco contractual bien diseñado asegura que los sistemas de IA en salud operen bajo **condiciones claras, seguras y sostenibles**, protegiendo tanto a los proveedores como a las instituciones sanitarias y, en última instancia, a los pacientes.

11.7. Riesgos estratégicos y mitigación

La sostenibilidad de un sistema de IA en salud no depende solo de su diseño técnico, sino también de la capacidad para **anticipar y gestionar riesgos estratégicos** que pueden comprometer su valor y continuidad. Estos riesgos no siempre son evidentes en la fase inicial, pero su impacto puede ser crítico si no se planifican medidas preventivas y mecanismos de seguimiento.

Requisitos mínimos:

- **Dependencia de proveedor (vendor lock-in):** exigir el uso de **formatos abiertos, APIs documentadas** y, cuando proceda, establecer mecanismos de **escrow para código y modelos** que garanticen la continuidad en caso de rescisión contractual o quiebra del proveedor.
- **Obsolescencia tecnológica y cambios regulatorios:** definir una **hoja de ruta tecnológica** que contemple actualizaciones planificadas y **revisiones anuales de cumplimiento normativo**, especialmente en relación con **MDR/IVDR, AI Act y RGPD/LOPDGDD**.
- **Drift y degradación de rendimiento:** disponer de **planes de recalibración, reentrenamiento y reevaluación clínica**, con criterios objetivos que determinen cuándo es necesario actuar para evitar riesgos clínicos derivados de la pérdida de precisión o validez del modelo.
- **Variabilidad de demanda y costes en la nube:** implementar **límites de consumo**, sistemas de **alerta temprana** y estrategias de optimización de recursos (reservas, planes de ahorro) para evitar desviaciones presupuestarias y garantizar la sostenibilidad económica.

Evidencias esperadas: registro de riesgos que incluya impacto y probabilidad; planes de mitigación documentados; y evidencias de **seguimiento periódico** de los riesgos identificados y de la efectividad de las medidas adoptadas.

En conjunto, la gestión proactiva de riesgos estratégicos asegura que los sistemas de IA en salud puedan mantenerse **seguros, viables y sostenibles en el tiempo**, incluso frente a incertidumbres tecnológicas, regulatorias o de mercado.

11.8. Indicadores clave (KPIs) de sostenibilidad

La sostenibilidad de un sistema de IA en salud no se garantiza con un plan inicial: requiere **monitorización continua** para anticipar riesgos, corregir desviaciones y demostrar valor clínico, económico y organizativo. Los **indicadores clave (KPIs)** permiten evaluar la viabilidad financiera, operativa y clínica del sistema, así como su capacidad de **escalado multicéntrico**.

Indicadores mínimos:

- **Financieros:** seguimiento del **TCO (Total Cost of Ownership)** anual, el **ROI acumulado**, el **coste por episodio/uso** y la **desviación presupuestaria** respecto a lo planificado. Estos datos son esenciales para justificar la continuidad y priorizar inversiones.
- **Operativos:** monitorización de la **disponibilidad del servicio, la latencia, la tasa de errores, las incidencias clasificadas por severidad, el tiempo medio de resolución y la tasa de rollback**. Estos indicadores reflejan la estabilidad, resiliencia y calidad del servicio.

- **Clínicos y de equidad:** evaluación de las **métricas de rendimiento y seguridad por subgrupos** (sexo, edad, centro, comorbilidad), junto con la **tasa de override** por parte de profesionales y la **satisfacción de usuarios y pacientes**. Estos indicadores garantizan que el sistema mantiene su fiabilidad y evita desigualdades.
- **Escalado:** número de **centros o unidades activas, tiempo medio de despliegue y esfuerzo de parametrización** requerido. Estos indicadores son clave para planificar la extensión multicéntrica y optimizar recursos.
- **Sostenibilidad ambiental:** seguimiento del **consumo energético** y de la **huella de carbono** asociados a la operación y, cuando proceda, al entrenamiento de los modelos, así como de la eficiencia de los algoritmos y de la infraestructura utilizada. Estos indicadores permiten alinear el despliegue del sistema de IA con las políticas institucionales de **sostenibilidad** y facilitar la rendición de cuentas en términos de **impacto ambiental**.
- **Ciberresiliencia:** véase el **cuadro de mando de ciberresiliencia** del apartado 10.7 para la definición, el método de medida y las evidencias asociadas (**MTTD, MTTR, RTO/RPO, % de restauraciones exitosas, n.º de simulacros/año, n.º de incidentes notificados, % de endpoints legacy segmentados**).

Evidencias esperadas: cuadro de mando actualizado con metas y umbrales; informes periódicos (p. ej., mensuales); y un **plan de acciones derivado de las métricas**, que permita corregir desviaciones y mejorar el desempeño del sistema.

En conjunto, los KPIs de sostenibilidad constituyen la **herramienta central de gestión y mejora continua**, asegurando que el sistema de IA aporta un valor **tangible, equitativo y sostenible** en el tiempo.

Salida esperada del apartado 11:

- **Caso de negocio** completo, incluyendo **TCO, ROI/ICER** y análisis de sensibilidad (escenarios optimista, central y pesimista).
- **Presupuesto detallado CAPEX/OPEX** con matriz de responsables y plan de aprovisionamiento.
- **Plan de financiación y modalidades de compra**, con cronograma vinculado a hitos técnicos y económicos.
- **Guía de despliegue multicentro** con perfiles/mapeos, acuerdos de gobernanza y evidencias de validación local.
- **Plan de continuidad y sostenibilidad operativa**, incluyendo SLA/SLO, RPO/RTO, pruebas de recuperación y repositorio de conocimiento.
- **Contratos y anexos de cumplimiento** (DTA, SLA/SLO, cláusulas de portabilidad, *escrow* cuando proceda) y *checklist* de obligaciones normativas (RGPD, MDR/IVDR, AI Act).
- **Registro de riesgos estratégicos**, con impacto/probabilidad, medidas de mitigación y evidencias de seguimiento.
- **Cuadro de mando de KPIs de sostenibilidad**, con metas, umbrales y plan de acciones derivadas.



12. Referencias

12.1. Normativa europea y española

- **Reglamento (UE) 2017/745 del Parlamento Europeo y del Consejo, de 5 de abril de 2017, sobre los productos sanitarios (MDR).** DOUE L 117, 5.5.2017.
Enlace: <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32017R0745>.
- **Reglamento (UE) 2017/746 del Parlamento Europeo y del Consejo, de 5 de abril de 2017, sobre los productos sanitarios para diagnóstico in vitro (IVDR).** DOUE L 117, 5.5.2017.
Enlace: <https://www.boe.es/doue/2017/117/L00176-00332.pdf>
- **Reglamento (UE) 2016/679 (RGPD), de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales.** DOUE L 119, 4.5.2016.
Enlace: <https://www.boe.es/doue/2016/119/L00001-00088.pdf>
- **Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales (LOPDGDD).** BOE-A-2018-16673.
Enlace: <https://www.boe.es/buscar/act.php?id=BOE-A-2018-16673>
- **Real Decreto 192/2023, de 21 de marzo, por el que se regulan los productos sanitarios.** BOE-A-2023-7416.
Enlace: <https://www.boe.es/buscar/doc.php?id=BOE-A-2023-7416>
- **Reglamento (UE) 2024/1689 (Ley de Inteligencia Artificial – AI Act).** DOUE L, 12.7.2024.
Enlace: <https://www.boe.es/buscar/doc.php?id=DOUE-L-2024-81079>

12.2. Guías reguladoras (MDCG/AEMPS)

- **MDCG 2019-11 rev.1 (2025): qualification and classification of software as medical device (MDR/IVDR).** Comisión Europea.
Enlace: https://health.ec.europa.eu/latest-updates/update-mdcg-2019-11-rev1-qualification-and-classification-software-regulation-eu-2017745-and-2025-06-17_en
- **MDCG 2020-1 (2020): clinical evaluation/performance evaluation of medical device software.** Comisión Europea.
Enlace: https://health.ec.europa.eu/system/files/2020-09/md_mdcg_2020_1_guidance_clinic_eva_md_software_en_0.pdf
- **MDCG – índice de documentos y otras guías.** Comisión Europea.
Enlace: https://health.ec.europa.eu/medical-devices-sector/new-regulations/guidance-mdcg-endorsed-documents-and-other-guidance_en
- **AEMPS – legislación y guías sobre productos sanitarios (incluye RD 192/2023).**
Enlace: <https://www.aemps.gob.es/la-aemps/legislacion/legislacion-sobre-productos-sanitarios/>

12.3. Buenas prácticas clínicas e investigación

- **ICH E6(R3) Guideline for Good Clinical Practice (GCP).** ICH, Step 4/Final (06.01.2025).
Enlace: <https://www.ema.europa.eu/en/ich-e6-good-clinical-practice-scientific-guideline#ich-e6r3-principles-and-annex-1-current-version-effective-from-23-july-2025-8264>
- **ISO 14155:2020 – investigación clínica de productos sanitarios en sujetos humanos (GCP).** Organización Internacional de Normalización.
Enlace (ficha oficial): <https://www.iso.org/standard/71690.html>

12.4. Normas técnicas (calidad, software y riesgo)

- **IEC 62304 – procesos del ciclo de vida del software de producto sanitario.** ISO/IEC Webstore.
Enlace: <https://www.iso.org/standard/38421.html>
- **ISO 14971:2019 – gestión de riesgos para productos sanitarios.** ISO.
Enlace: <https://www.iso.org/obp/ui/es/#iso:std:72704:es>

12.5. Ciberseguridad y resiliencia

- **Reglamento (UE) 2025/38** del Parlamento Europeo y del Consejo, de 19 de diciembre de 2024, por el que se establecen medidas destinadas a reforzar la solidaridad y las capacidades en la Unión a fin de detectar ciberamenazas e incidentes, prepararse y responder a ellos.
Enlace: <https://www.boe.es/buscar/doc.php?id=DOUE-L-2025-80049>
- **Ley de Ciberresiliencia (CRA, Cyber Resilience Act):** en vigor **10/12/2024**; obligaciones principales desde **11/12/2027**.
Enlace: <https://digital-strategy.ec.europa.eu/en/library/cyber-resilience-act>
- **Directiva (UE) 2022/2555** del Parlamento Europeo y del Consejo de 14 de diciembre de 2022 relativa a las medidas destinadas a garantizar un elevado nivel común de ciberseguridad en toda la Unión. Conocida como **NIS2**.
Enlace: <https://www.boe.es/buscar/doc.php?id=DOUE-L-2022-81963>
- **Real Decreto 311/2022, de 3 de mayo, por el que se regula el Esquema Nacional de Seguridad. (ENS):**
Enlace: <https://www.boe.es/buscar/act.php?id=BOE-A-2022-7191>
- **CSIRT de referencia: INCIBE-CERT** (privado) y **CCN-CERT** (sector público/AGE); coordinación y notificación definidas en **RD-ley 12/2018** y **RD 43/2021**.
- **Dictamen del Comité Europeo de las Regiones Ciberseguridad de los hospitales y los prestadores de asistencia sanitaria.**
Enlace: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=OJ:C_202504415
- **ISO/IEC 27001** es un estándar para la seguridad de la información (*Information security, cybersecurity and privacy protection - Information security management systems - Requirements*)

Enlace: <https://www.iso.org/es/norma/27001>

- **ISO/IEC 27002** es una norma internacional que brinda orientación a las organizaciones que desean **establecer, implantar y mejorar un sistema de gestión de seguridad de la información** (SGSI) centrado en la **ciberseguridad**.

Enlace: <https://www.iso.org/es/contents/data/standard/07/56/75652.html>

- **ISO/IEC 27701** es una norma internacional que especifica los requisitos y aporta orientación para **establecer, implementar, mantener y mejorar continuamente un sistema de gestión de información de privacidad** (PIMS).

Enlace: <https://www.iso.org/es/contents/data/standard/08/58/85819.html>

- **IEC 81001-5-1:2022** establece los requisitos de seguridad para el software sanitario y los HITS.

Enlace: <https://www.iso.org/standard/76097.html>

- **ISO/IEC 30111** es una norma internacional que describe el manejo adecuado de la información de posibles vulnerabilidades en productos.

Enlace: <https://www.iso.org/standard/69725.html>

- **ISO/IEC 29147** es una norma internacional que proporciona directrices para los procesos de divulgación de vulnerabilidades..

Enlace: <https://www.iso.org/es/contents/data/standard/07/23/72311.html>

12.6. Datos, ética y gobernanza de la IA en salud

- **Ethics guidelines for trustworthy AI.** High-Level Expert Group on AI, Comisión Europea (08.04.2019).

Enlace: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

- **Ethics and governance of artificial intelligence for health.** OMS (2021).

Enlace: <https://www.who.int/publications/i/item/9789240029200>

12.7. Interoperabilidad y terminologías

- **HL7 FHIR – overview y especificación.** HL7.

Enlace: <https://hl7.org/fhir/overview.html>

- **DICOM – estándar y edición vigente.** NEMA/MITA.

Enlaces: <https://dicomstandard.org> y <https://dicom.nema.org/medical/dicom/current>

- **openEHR – especificaciones.**

Enlace: <https://specifications.openehr.org/>

- **OMOP Common Data Model – documentación.** OHDSI.

Enlace: <https://ohdsi.github.io/CommonDataModel/>

13. Anexos operativos

13.1. Checklists transversales (verificación rápida)

Proyecto / Sistema de IA: _____

Versión del documento: _____

Fecha de revisión: ____ / ____ / ____

Revisor(es): _____

Observaciones generales: _____

Uso: marcar **SI** (cumplido), **No** (pendiente), **N/A** (no aplica). Añadir enlace a evidencia.

A) Datos y gobernanza

- Existe **DMP** aprobado y versionado.
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)
- Diccionario de variables y linaje/documentación **ETL/ELT**.
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)
- Informe de calidad y representatividad (subgrupos).
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)
- Plan de retención/eliminación y copias (**RPO/RTO**).
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)

B) Protección de datos (RGPD/LOPDGDD)

- **Base jurídica** definida y Registro de actividades de tratamiento.
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)
- **DPIA/EIPD** realizada y firmada.
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)
- Seudonimización/anonimización y control de acceso por rol.
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)
- Contratos de encargo/corresponsabilidad y transferencias internacionales, si procede.
 - Sí No N/A

- Evidencia: (añadir enlace a evidencia)

C) Regulación/ética

- Cualificación del software (¿PS? clase y justificación).
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)
- Gestión de riesgos (ISO 14971) y plan de evaluación clínica/desempeño.
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)
- Aprobaciones **CEIm/organizativas** y, cuando proceda, interacción con **AEMPS**.
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)
- Mapeo de requisitos **AI Act** (alto riesgo) y cronograma.
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)

D) Diseño/validación

- Protocolo de validación con métricas (incluida **calibración**). Evidencia: _____
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)
- **Validación externa** (al menos una) y análisis por subgrupos.
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)
- Reproducibilidad (versionado de datos/código/artefactos).
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)

E) Explicabilidad y trazabilidad

- Guía de explicabilidad por rol (clínico, datos, auditor). Evidencia: _____
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)
- Política de **logging** de inferencias y conservación.
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)

F) Monitorización y MLOps

- Plan de métricas (servicio y clínicas), **drift** y recalibración.

- Sí No N/A
- Evidencia: (añadir enlace a evidencia)
- **Changelog**, criterios de **cambio significativo** y **CI/CD**.
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)

G) Interoperabilidad e integración

- Perfiles **HL7 FHIR** y **DICOM** definidos y validados.
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)
- Catálogo de interfaces, pruebas **end-to-end** y modo **shadow**.
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)

H) Seguridad del paciente y operación

- Análisis de peligros, salvaguardas y **override** clínico.
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)
- Plan de incidentes y notificación; **SLA/SLO**.
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)

I) Verificación rápida de ciberseguridad (SBOM/VDP/backups)

- **SBOM** por versión:
 - Sí No N/A
 - referencia de versión: _____
- **VDP publicado** (divulgación de vulnerabilidades):
 - Sí No N/A
 - url/correo: _____ — sla: _____
- **Parchoeo de CVES críticas < 30 días**:
 - Sí No N/A
 - evidencias/fecha: _____
- **Prueba de restauración (< 12 meses)**:
 - Sí No N/A
 - fecha última prueba: _____
 - resultado: _____

- OK
- con incidencias

- Gestión de claves (kms/hsm) y rotación:
 - Sí No N/A
 - Periodicidad: _____
- Registro/auditoría de accesos e inferencias:
 - Sí No N/A
 - Retención: _____ meses

J) Sostenibilidad y contratación

- TCO/ROI y plan de financiación/compra.
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)
- Contratos/licencias (propiedad, portabilidad y salida).
 - Sí No N/A
 - Evidencia: (añadir enlace a evidencia)

13.2. Formulario de evaluación PI (TRL 1–3)

Identificación del proyecto: título, IP, centro, línea clínica, TRL objetivo.

Resumen no técnico (≤ 200 palabras): _____

Bloques y puntuación (0–3 por ítem):

1. Diseño científico y relevancia clínica (8 pts) → /8
2. Datos: calidad y representatividad (10 pts) → /10
3. Metodología y rendimiento (8 pts) → /8
4. Ética y protección de datos (4 pts) → /4
5. Plan de madurez y transferencia (4 pts) → /4
6. Innovación y encaje estratégico (2 pts) → /2

Total PI (36): ___/36 → **No viable / Viable / Viable y financiable**

Evidencias adjuntas (enlaces): protocolo, DMP, informe de calidad, resultados, DPIA, etc.

Acciones recomendadas y plazos: _____

13.3. Formulario de evaluación ICPS (TRL ≥4)

Proyecto / Sistema de IA: _____

TRL actual: _____

TRL objetivo: _____

Fecha de revisión: ___ / ___ / ___

Revisor(es): _____

Identificación del estudio: título, promotor/fabricante legal, centro/es, clase de PS, intended purpose, TRL. (≤ 200 palabras): _____

Autorizaciones/Registros: CEIm (fecha y referencia), AEMPS (notificación/autorización, si aplica), EUDAMED (si aplica).

Diseño del estudio: tipo (exploratorio/confirmatorio; pre/post-market), población, endpoints, tamaño muestral y análisis.

Integración y flujo del estudio: datos requeridos, integración con HCE/FHIR/DICOM si procede, procedimientos y roles.

Riesgos y salvaguardas: gestión de riesgos, GSPR, plan de incidentes y supervisión humana.

Datos/Equidad: calidad y representatividad, subgrupos, DataSheet/ModelCard.

Operación y seguridad: bloqueo de versiones para protocolo, logging y auditoría, protección de datos (RGPD/LOPDGDD, DPIA).

Transición posestudio: plan de PMS/PMCF o despliegue asistencial, condiciones para adopción.

Bloques y puntuación (0–3 por ítem):

1. Valor clínico y seguridad del paciente (10 pts) → /10

2. Regulación y conformidad (10 pts) → **/10**
3. Interoperabilidad e integración (8 pts) → **/8**
4. Monitorización y MLOps (8 pts) → **/8**
5. Datos y equidad (6 pts) → **/6**
6. Sostenibilidad y escalabilidad (6 pts) → **/6**

Total ICPS (48): /48 → **No viable / Viable / Viable y financiable**

Evidencias adjuntas: validación externa, CEIm/AEMPS, perfiles FHIR/DICOM, plan de monitorización, contrato/SLAs, evaluación económica.

Condiciones de despliegue (si aplica):

13.4. Plantilla DataSheet (dataset)

- **Origen y periodo:**
- **Población y criterios:** inclusión/exclusión, centros.
- **Variables y terminologías:** diccionario, codificaciones (SNOMED CT, LOINC, ICD-10-ES, etc.).
- **Procesos de limpieza/etiquetado:** reglas y validaciones.
- **Representatividad y sesgos conocidos:** análisis por subgrupos.
- **Limitaciones y calidad:** *missingness, outliers*, coherencia temporal.
- **Acceso y licencias:** restricciones y procedimiento de solicitud.
- **Versionado y linaje:** identificadores y cambios respecto a versiones previas.

13.5. Plantilla ModelCard (modelo)

- **Objetivo clínico e finalidad prevista (intended use):**
- **Población diana y contexto de uso:**
- **Datos y preprocesado:** versiones y filtros.
- **Arquitectura y entrenamiento:** resumen reproducible.
- **Métricas clave:** rendimiento y **calibración** global y por subgrupos.
- **Riesgos, salvaguardas y supervisión humana:**
- **Usos permitidos/prohibidos:** límites explícitos.

- **Monitorización y mantenimiento:** métricas en producción, umbrales y criterios de **cambio significativo**.
- **Limitaciones conocidas y escenarios no recomendados:**
 - Limitaciones técnicas: _____
 - Condiciones clínicas o contextos donde no se recomienda su uso: _____
- **Versionado:** identificador de modelo y fecha.

13.6. Plan de monitorización posdespliegue (plantilla)

- **Objetivos y responsables:** comité, roles y **RACI**.
- **Métricas clínicas y de servicio:** disponibilidad, latencia, fallos, rendimiento por subgrupos, tasa de **override**.
- **Detección de *drift*:** tests y ventanas; umbrales y tiempos de reacción.
- **Recalibración/actualización:** procedimiento, entorno **shadow**, criterios de aceptación.
- **Reporte y periodicidad:** tablero, informes y comunicación a usuarios/órganos de supervisión.

13.7. Gestión de cambios y homologación por versión (plantilla)

- **Clasificación del cambio:** menor / mayor / **significativo**.
- **Riesgo e impacto:** clínica, seguridad, rendimiento, integración.
- **Pruebas requeridas:** unitarias, integración, **end-to-end**, usabilidad.
- **Aprobaciones:** comité técnico/ético; ventana de mantenimiento.
- **Despliegue y *rollback*:** plan y validación post-despliegue.
- **Documentación:** **changelog**, versiones de datos/modelo/servicio.

13.8. Dossier de interoperabilidad (plantilla)

- **Perfiles y recursos:** **HL7 FHIR** (recursos y perfiles nacionales), **DICOM/DICOM SR**.
- **Mapeos terminológicos:** SNOMED CT, LOINC, ICD-10-ES, ATC/RxNorm (según entorno).
- **Interfaces y eventos:** endpoints, autenticación (**OIDC/OAuth2**), **CDS Hooks/SMART** si aplica.
- **Pruebas y conformidad:** validadores, casos de prueba y resultados.

- **Seguridad y auditoría:** cifrado, trazas y retención de logs.

13.9. Expediente regulatorio mínimo (si cualifica como PS)

- **Intended purpose y clasificación** (MDR/IVDR) con justificación.
- **Gestión de riesgos** (ISO 14971) y **IEC 62304** (ciclo de vida software).
- **Evaluación clínica/desempeño y plan post-market.**
- **Cumplimiento AI Act** para alto riesgo (gobernanza, datos, trazabilidad, supervisión humana).
- **Vigilancia y seguridad:** incidentes, tendencias y acciones correctivas.

13.10. DPIA/EIPD (esqueleto)

- **Descripción del tratamiento y finalidad:** _____
- **Bases jurídicas y necesidad/proporcionalidad:** _____
- **Riesgos para derechos/libertades:** catálogo y evaluación.
- **Medidas de mitigación:** técnicas y organizativas (incluye seudonimización/anonimización).
- **Conclusión y firma:** DPD, responsables y fecha.

13.11. Caso de negocio y evaluación económica (plantilla)

- **Problema y alternativas:** _____
- **Costes: CAPEX/OPEX**, datos e integración.
- **Beneficios:** clínicos, operativos y económicos.
- **Análisis: TCO, ROI, ICER** (si aplica) y sensibilidad.
- **Plan de financiación/compra:** convocatoria, **CPI**, licencias/uso/resultado.

13.12. Reportes tipo

- **Informe a CEIm:** sinopsis, población, datos, riesgos/beneficios, consentimiento/ información y salvaguardas.

- **Comunicación a AEMPS (cuando proceda):** cualificación/clase, riesgos, evaluación clínica/desempeño, plan poscomercialización.
- **Informe a financiadores (p. ej., ISCIII):** objetivos, TRL, evidencias, plan de madurez, presupuesto y KPIs.
- **Informe operativo interno:** estado, incidencias, **drift**, recalibraciones y acciones.

13.13. Política SBOM y boletín de vulnerabilidades (plantilla)

- **Alcance y formato SBOM** (CycloneDX/SPDX), responsable, periodicidad de emisión.
- **Flujo de revisión de CVEs** (umbral CVSS), parcheo y comunicación.
- **Boletín de vulnerabilidades:** versión afectada, CVE, impacto, workaround, fecha de parche.

Salida: SBOM firmada por versión + boletín vigente.

13.14. Procedimiento de divulgación de vulnerabilidades (VDP)

- Canal público (URL/correo), PGP, alcance, no-legal-threats clause.
- SLA de acuse (≤ 7 días) y de corrección (según severidad).
- Coordinación con CSIRTs y notificación regulatoria cuando proceda.

Salida: VDP publicado + registro de casos y tiempos.

13.15. Plantilla de modelado de amenazas (STRIDE) y matriz 14971

- Catálogo inicial STRIDE (spoofing, tampering, repudiation, information disclosure, denial of service, elevation of privilege) por componente.
- Matriz **requisito \leftrightarrow amenaza \leftrightarrow control \leftrightarrow evidencia** enlazada a **ISO 14971** y verificación (7.4).

Salida: diagrama de datos/flujo + matriz firmada y versionada.

Salida esperada del apartado 12: checklists cumplimentados, formularios **PI/ICPS** con puntuación, **DataSheet/ModelCard** actualizados, planes de monitorización y cambios aprobados, dossier de interoperabilidad, expediente regulatorio (si aplica), **DPIA**, caso de negocio y reportes tipo.