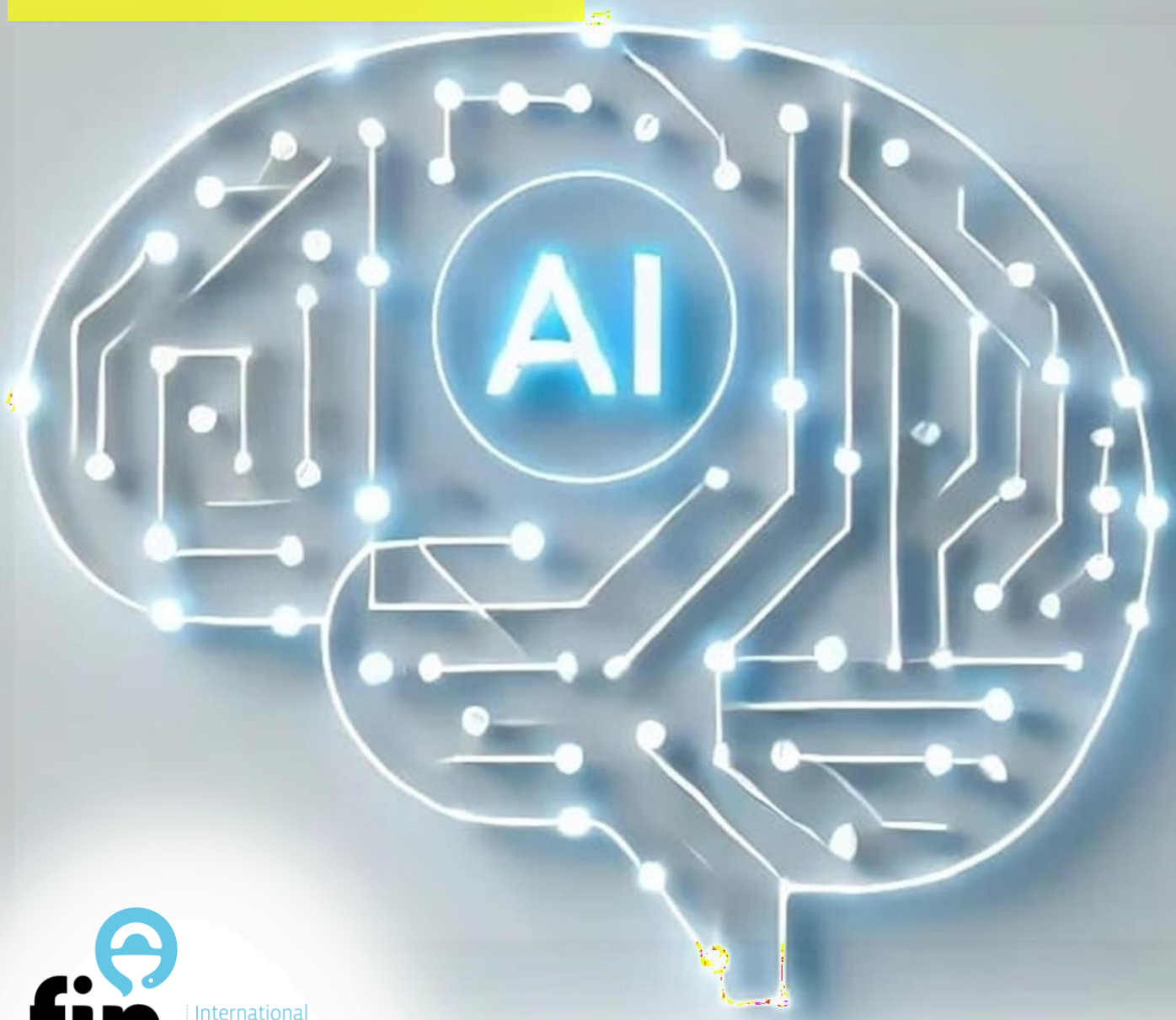
Kit de herramientas de inteligencia artificial para farmacias.

Introducción y guía de recursos para farmacéuticos

Marzo de 2025



fip International Pharmaceutical Federation

Federación Farmacéutica Internacional (FIP)
Andries Bickerweg 5
2517 JP La Haya
Países Bajos
www.fip.org

Redactor:

Dr. Whitley Yi, Presidente del Grupo de Trabajo sobre Inteligencia Artificial del Grupo Asesor Tecnológico (GAT) de la FIP.

Coeditores:

Dr. Ardalan Mirzaei, Representante para Asia y el Pacífico en la Sección de Información sobre Salud y Medicamentos de la FIP.
Sr. Brent Sin Hidge, Grupo de Trabajo sobre Inteligencia Artificial del GAT de la FIP.
Dra. Mariana Guia, Grupo de Trabajo sobre Inteligencia Artificial del GAT de la FIP.
Dr. Paul Voigt, Grupo de Trabajo sobre Inteligencia Artificial del GAT de la FIP.

Traducción al español:

Farm. Jorge Schlottke - Miembro Ejecutivo de la Sección de Farmacia Comunitaria de FIP - Argentina.
Farm. Florencia Gómez – Confederación Farmacéutica Argentina - Argentina.
Dr. José Luis Nájera García – Miembro Ejecutivo de la Sección de Farmacia Comunitaria de FIP - España
Farm. Gabriel Molina – Pasante en el Foro Farmacéutico de las Américas - Argentina.
Farm. Gonzalo Miguel Adsuar Meseguer – Enlace entre la Sección de Farmacia Comunitaria y el Grupo de Nuevos Farmacéuticos de FIP - España.

Cita recomendada:

Federación Internacional Farmacéutica (FIP). Kit de herramientas de inteligencia artificial para farmacias. Introducción y guía de recursos para farmacéuticos. La Haya: Federación Farmacéutica Internacional; 2025

Contenido

1 Agradecimientos

La FIP expresa su gratitud y reconocimiento a todos los autores que han colaborado en la elaboración de este kit de herramientas. Los autores se enumeran a continuación:

- Whitley Yi, PharmD, BCPS, Presidente del Grupo de Trabajo sobre Inteligencia Artificial del Grupo Asesor Tecnológico (GAT) de la FIP, Director de Farmacia y Servicios para Miembros, Well, Profesora adjunta, Facultad de Farmacia y Ciencias Farmacéuticas Skaggs de la Universidad de Colorado, EE. UU.
- Brent Sin Hidge, BPharm (NMU), Grupo de Trabajo de Inteligencia Artificial del GAT de la FIP; Miembro del Comité Ejecutivo Nacional de la Sociedad Farmacéutica de Sudáfrica; Presidente de la Asociación Sudafricana de Farmacéuticos Hospitalarios e Institucionales de Cabo Occidental; Farmacéutico Hospitalario, Hospital Netcare Blaauwberg, Sudáfrica.
- Bruno Macedo, Grupo de Trabajo de Inteligencia Artificial del GAT de la FIP, Director General y Fundador de MedFacts, Director de Programas, Fundación Calouste Gulbenkian, Portugal.
- Claudia Rijcken, Grupo de Trabajo de Inteligencia Artificial del GAT de la FIP, Miembro del Grupo Asesor Tecnológico de la FIP, CSO y Fundadora de Pharmi, Profesora de la Facultad de Farmacia de la Universidad de Utrecht, Países Bajos.
- Florencia Ojeda, Ingeniera de Sistemas, Máster en Inteligencia Artificial, Grupo de Trabajo de Inteligencia Artificial del GAT de la FIP, Uruguay
- Joanna Klopotowska, Miembro del Grupo de Trabajo de Inteligencia Artificial del GAT de la FIP, Profesora Adjunta e Investigadora Principal, Amsterdam UMC Centro Médico Académico, Países Bajos.
- Mariana Guia, PharmD; Grupo de Trabajo de Inteligencia Artificial del GAT de la FIP; Especialista en Soluciones Sanitarias, Asociación Nacional Portuguesa de Farmacias, Portugal
- Markus Manner, Grupo de Trabajo de Inteligencia Artificial del GAT de la FIP, Director de Desarrollo, Suomen Apteekkariliitto, Asociación de Farmacias Finlandesas, Finlandia.
- Martin Kondža, MPharm, PhD, Grupo de Trabajo de Inteligencia Artificial del GAT de la FIP, Profesor Asistente, Universidad de Mostar, Bosnia and Herzegovina / Cámara Farmacéutica de la Federación de Bosnia y Herzegovina, Bosnia y Herzegovina.
- Mauro Tschanz, Grupo de Trabajo de Inteligencia Artificial del GAT de la FIP, Experto en Digitalización de la Asociación Suiza de Farmacia (PharmaSuisse), Suiza
- Mohd Syamir Mohamad Shukeri, B.Pharm., M. Comm. Health, R.Ph. MMPS, Grupo de Trabajo de Inteligencia Artificial del GAT de la FIP, Sociedad de Farmacéuticos de Malasia, Malasia
- Paul Voigt, BPharm, MSc, DBiotech, ADHSML, Grupo de Trabajo de Inteligencia Artificial del GAT de la FIP, Sociedad Farmacéutica de Sudáfrica, Director de Operaciones de Inventario, Mediclinic Southern Africa
- Régis Vaillancourt, B.Pharm., PharmD, FFIP, FOPQ, FCSHP, Grupo de Trabajo de Inteligencia Artificial del GAT de la FIP, Vicepresidente de asuntos farmacéuticos BCE Pharma Inc, Canadá
- Sangeetha Ramdave, Grupo de Trabajo sobre Inteligencia Artificial del GAT de la FIP, Académica de sesión, Universidad de Monash, Australia

FIP agradece al presidente del Grupo Asesor de Tecnología de la FIP, Lars-Åke Söderlund, y a todos los miembros del Grupo Asesor de Tecnología de la FIP su apoyo, aportes y comentarios sobre este kit de herramientas.

1 Resumen ejecutivo

El "Kit de herramientas de inteligencia artificial para la farmacia" es un recurso diseñado para ayudar a los farmacéuticos a integrar las tecnologías de inteligencia artificial (IA) en su práctica. Desarrollado por la Federación Internacional Farmacéutica (FIP), este conjunto de herramientas pretende acortar la distancia entre el campo de la IA, en rápida evolución, y las necesidades prácticas de los farmacéuticos.

Objetivo e importancia de la IA en la farmacia: La IA está revolucionando la atención sanitaria al permitir el análisis de grandes cantidades de datos médicos, ayudar en la toma de decisiones clínicas, personalizar los tratamientos de los pacientes, predecir brotes de enfermedades y optimizar los flujos de trabajo operativos. Estrechamente alineada con el Objetivo de Desarrollo 20 de FIP (Salud Digital), la IA ayuda a crear una fuerza de trabajo farmacéutica digitalmente competente, mejorando la atención clínica, la detección de enfermedades, la gestión de los sistemas de salud y la investigación farmacéutica. Sin embargo, los farmacéuticos se enfrentan a retos críticos, como la privacidad de los datos, las amenazas a la ciberseguridad, los posibles sesgos de los algoritmos y las preocupaciones éticas.

Compromiso e iniciativas de la FIP: La FIP se compromete a orientar y apoyar a los farmacéuticos en el uso eficaz de las tecnologías de IA. En 2023, la FIP creó un grupo de trabajo sobre IA para identificar las necesidades de sus circunscripciones en torno a la IA y mejorar la colaboración intersectorial. Este grupo pretende colmar las lagunas de conocimiento y promover la comprensión y el uso de las tecnologías de IA entre farmacéuticos y científicos farmacéuticos.

Objetivos del kit de herramientas: Este kit de herramientas proporciona una guía de alto nivel para farmacéuticos, ofreciendo una visión general de las consideraciones de implementación de la IA, aplicaciones prácticas e innovación inspiradora. Su objetivo es capacitar a los farmacéuticos para ofrecer una atención al paciente más segura, eficaz y personalizada sin afectar su pensamiento crítico ni su juicio profesional.

IA y los Objetivos de Desarrollo de la FIP: El kit de herramientas apoya la consecución de los Objetivos de Desarrollo de la FIP y se armoniza con los Objetivos de Desarrollo Sostenible de la ONU, centrándose en la educación y el desarrollo profesional continuo (DPC), la alfabetización digital y la innovación en la práctica farmacéutica. Entre los objetivos específicos se incluyen el desarrollo de recursos educativos para la competencia digital, la integración de la formación en IA en la educación profesional y la provisión de orientación estratégica sobre políticas de personal sanitario digital.

Retos y consideraciones: La integración efectiva de la IA en la farmacia requiere que los farmacéuticos comprendan las capacidades y limitaciones de la IA, seleccionando herramientas adecuadas adaptadas a desafíos específicos. Las consideraciones clave incluyen garantizar la calidad de los datos, el cumplimiento normativo, los marcos éticos y la inversión en infraestructura. Además, abordar el acceso equitativo es vital para evitar exacerbar las disparidades y desigualdades sanitarias existentes.

Conclusiones: Este conjunto de herramientas tiene como objetivo servir como un recurso integral, permitiendo a los farmacéuticos navegar por las complejidades de la IA, capitalizar su potencial y comprender claramente sus limitaciones. Al fomentar la innovación, la colaboración y un enfoque proactivo, la FIP busca mejorar el papel de la IA en la práctica farmacéutica, impulsando en última instancia mejoras en la atención al paciente y en los resultados sanitarios globales.

2 Prefacio

2.1 Finalidad del kit de herramientas

La inteligencia artificial (IA) es una rama de la informática dedicada a crear sistemas capaces de realizar tareas que normalmente requieren inteligencia humana, como el aprendizaje, el razonamiento, la resolución de problemas y el reconocimiento de patrones. Las aplicaciones de IA se utilizan cada vez más para analizar grandes volúmenes de datos médicos, ayudar en la toma de decisiones clínicas, personalizar planes de tratamiento, predecir brotes de enfermedades y optimizar los flujos de trabajo operativos en entornos sanitarios.

La salud digital es una prioridad para la FIP, que orienta a los farmacéuticos sobre la mejor manera de utilizar los avances digitales para apoyar la prestación de asistencia sanitaria, la práctica farmacéutica y la investigación farmacéutica. La IA es un área emergente dentro de la salud digital que tiene un inmenso potencial para mejorar la eficiencia en la atención clínica, la detección y vigilancia de enfermedades, la optimización de la gestión de los sistemas de salud y la facilitación de la investigación y el desarrollo farmacéutico avanzado. Además, la IA tiene el potencial de promover la equidad en el acceso a la asistencia sanitaria en todo el mundo, capacitando a los pacientes para gestionar activamente sus necesidades sanitarias. A pesar de la proliferación de soluciones digitales basadas en IA que están siendo introducidas en el ámbito farmacéutico, la IA presenta sus propios retos y plantea riesgos en materia de privacidad de datos, seguridad, sesgos codificados y amenazas de ciberseguridad.

La FIP es consciente de su responsabilidad en colaborar, orientar y educar a los farmacéuticos de todo el mundo en la utilización eficaz de la IA para lograr resultados sanitarios óptimos. La FIP se compromete a fomentar un entorno en el que los farmacéuticos puedan adoptar con confianza estas tecnologías. En 2023, con estos objetivos en mente, la FIP estableció un grupo de trabajo de IA como subgrupo del Grupo Asesor de Tecnología de la FIP, encargado de identificar las necesidades relacionadas con la IA de los miembros de la FIP y de proporcionar recursos prácticos y orientación.

Este conjunto de herramientas, desarrollado por el Grupo de Trabajo sobre IA, pretende tender un puente entre el campo de la IA, en rápida evolución, y las necesidades prácticas y cotidianas de los farmacéuticos. Sirve como guía de alto nivel para los farmacéuticos a la hora de navegar por las complejidades de las aplicaciones de la IA, proporcionando una visión general para la implementación de la IA. Su objetivo es ilustrar las aplicaciones prácticas e inspirar la innovación entre los miembros de la FIP, garantizando que estén bien equipados para navegar por el cambiante panorama de la IA y aprovechar sus beneficios, al tiempo que comprenden sus limitaciones.

En última instancia, este kit de herramientas tiene como objetivo capacitar a los farmacéuticos y equipos de farmacia para integrar con confianza la IA en su práctica de una manera que complemente su experiencia, lo que les permite ofrecer una atención al paciente más segura, eficaz, eficiente y personalizada, preservando al mismo tiempo el pensamiento crítico y el juicio profesional de los farmacéuticos.

2.2 La IA y los Objetivos de Desarrollo de la FIP

Lanzados en septiembre de 2020, los Objetivos de Desarrollo de la FIP pretenden dirigir la transformación de la profesión farmacéutica a nivel mundial hasta 2030 (visite https://developmentgoals.fip.org/). En consonancia con los Objetivos de Desarrollo Sostenible de las Naciones Unidas, una iniciativa mundial más amplia, los Objetivos de Desarrollo de la FIP se centran específicamente en mejorar la práctica, la educación y las ciencias farmacéuticas. Los mismos permiten la identificación de puntos en común y la colaboración intersectorial dentro de un marco transformador para la profesión farmacéutica.

El Objetivo de Desarrollo de la FIP sobre Salud Digital (Objetivo de Desarrollo 20) se estructura en torno a tres elementos: educación y fuerza laboral, práctica y ciencia.

Educación y fuerza laboral: ´Facilitadores de la transformación digital dentro del personal de farmacia y procesos eficaces para facilitar el desarrollo de personal de farmacia con conocimientos digitales.´

En el contexto de la IA, este conjunto de herramientas pretende apoyar los siguientes mecanismos de los Objetivos de Desarrollo:

- Desarrollar cursos y capacitaciones para una fuerza laboral alfabetizada digitalmente (Objetivos de Desarrollo 1 y 2 de la FIP).
- Integrar las competencias en materia de alfabetización y salud digital en los marcos avanzados y especializados de FIP (Objetivos de Desarrollo 4 y 5 de la FIP).
- Proporcionar herramientas de aprendizaje multidisciplinar para mejorar la alfabetización sanitaria digital (Objetivo de Desarrollo 8 de la FIP).
- Promover el uso e interpretación de la IA en la formación y educación de farmacéuticos y científicos farmacéuticos. Crear oportunidades de formación y desarrollo continuos para garantizar una práctica actualizada con los avances tecnológicos (Objetivo de Desarrollo 9 de la FIP).
- Proporcionar orientación sobre la incorporación de la IA en las políticas de desarrollo del personal sanitario digital, incluido el empleo (Objetivo de Desarrollo 13 de la FIP).

Práctica: ´Sistemas y estructuras implementados para desarrollar y prestar servicios sanitarios digitales de calidad mediante la alfabetización digital y la utilización de la tecnología y los habilitadores digitales, configuración de servicios digitales receptivos para ampliar el acceso y la equidad.´

En el contexto de la IA, este conjunto de herramientas pretende apoyar los siguientes mecanismos de los Objetivos de Desarrollo:

- Proporcionar información sobre los habilitadores de salud digital y las tecnologías impulsadas por IA que apoyan la prestación de servicios de vanguardia y la toma de decisiones clínicas (Objetivo de Desarrollo 20 de la FIP).
- Debatir la alfabetización digital y la gobernanza en relación con la propiedad, la ética, la privacidad, las implicancias operativas y la calidad de la información. Orientar las políticas para apoyar y mejorar la gestión de los datos sanitarios y la rendición de cuentas sobre los resultados de los pacientes (Objetivo de Desarrollo 20 de la FIP).
- Promover iniciativas de salud digital impulsadas por la IA que mejoren el acceso equitativo a la atención farmacéutica (Objetivo de Desarrollo 20 de la FIP).

Ciencia: ´Aplicación de la tecnología digital a la prestación de asistencia sanitaria y al desarrollo de productos médicos innovadores.´

En el contexto de la IA, este conjunto de herramientas pretende apoyar el siguiente mecanismo de los Objetivos de Desarrollo:

- Facilitar la integración de la IA como una solución de ciencia de datos, permitiendo a los farmacéuticos ofrecer una mejor atención al paciente a través de tecnologías sanitarias innovadoras (Objetivo de Desarrollo 20 de la FIP).

3 Contexto

El principal problema que motivó la necesidad de este kit de herramientas de IA es la integración de la IA en la práctica farmacéutica diaria en entornos industriales, hospitalarios y comunitarios, con el objetivo de optimizar los procesos de trabajo de los farmacéuticos, mejorar la atención al paciente, estimular la interacción multidisciplinar y formar a los farmacéuticos en el uso eficiente de las mismas.

La integración de las tecnologías de aprendizaje automático (o ML del inglés “machine learning") e IA está reconfigurando las funciones y responsabilidades de los profesionales sanitarios, incluidos los farmacéuticos. La IA puede aumentar las capacidades de los farmacéuticos en una amplia gama de tareas, como la gestión de la medicación (tanto en el contexto logístico como farmacéutico), el asesoramiento a los pacientes, el chequeo de las interacciones entre medicamentos, la medicina personalizada y la investigación farmacéutica, entre otras.

Los recientes avances en IA generativa han acelerado aún más su adopción, lo que plantea la necesidad de un despliegue reflexivo y responsable. Aunque esta es muy prometedora para la atención farmacéutica, también presenta importantes limitaciones y retos que deben abordarse para garantizar una integración eficaz y ética.

La calidad y la integridad de los datos utilizados y producidos por los sistemas de IA constituyen una preocupación importante. Dado que sus resultados son tan fiables como los datos de los que aprenden. Las imprecisiones, sesgos o incoherencias en los conjuntos de datos de entrenamiento pueden llevar a conclusiones o predicciones incorrectas. Esto es especialmente crítico en la atención farmacéutica, donde la seguridad del paciente y la eficacia de los medicamentos están en juego.

Otro problema crítico es la "alucinación de la IA", cuando los sistemas generativos de IA producen resultados incorrectos, sin sentido o inventados. Esto plantea retos importantes en aplicaciones críticas como la atención farmacéutica.[1] Por ejemplo, la IA puede generar consejos médicos o recomendaciones de tratamiento que parezcan plausibles, pero que sean falsos, debido a limitaciones en los datos de entrenamiento, errores de reconocimiento de patrones o errores de inferencia. Para hacer frente a estas alucinaciones es necesario mejorar la calidad de los datos de entrenamiento, aplicar medidas de verificación y garantizar la supervisión humana para mantener la precisión y la fiabilidad. A medida que los proveedores de atención farmacéutica integran cada vez más herramientas de IA, comprender y mitigar las alucinaciones de la IA es crucial para garantizar una atención al paciente segura y eficaz.

Además de las salvaguardias técnicas y la supervisión humana, las consideraciones normativas desempeñan un papel fundamental en el éxito de la integración de la IA. La atención farmacéutica está muy regulada, y las aplicaciones de IA deben sortear complejos requisitos normativos relativos al desarrollo de fármacos, los ensayos clínicos y la privacidad de los pacientes. El cumplimiento de estas normativas, tanto a nivel local como internacional, es esencial para garantizar la seguridad de los pacientes y la confianza en las nuevas tecnologías que se están implantando.

Además, la implantación de la IA en la atención farmacéutica requiere una inversión sustancial en términos de infraestructura, formación y gestión. Los proveedores y las organizaciones de atención sanitaria deben estar preparados para invertir en sistemas informáticos sólidos y en la formación del personal para manejar estas tecnologías avanzadas con eficacia.

Las consideraciones éticas también desempeñan un papel fundamental. El uso de la IA debe ajustarse a las normas éticas relativas a la confidencialidad y seguridad del paciente, el consentimiento y la prevención de sesgos en las recomendaciones de tratamiento. Garantizar que los sistemas de IA sean transparentes y su funcionamiento comprensible para los reguladores y los profesionales es crucial para su aceptación y fiabilidad.

Por último, aunque la IA puede mejorar la personalización y la eficiencia de la atención, existe el riesgo de agravar las disparidades sanitarias existentes si no se gestiona con cuidado. Garantizar un acceso equitativo a los beneficios de la tecnología de IA en la atención farmacéutica es esencial para evitar exacerbar las desigualdades sanitarias.

Este kit de herramientas profundiza en estos temas y proporciona una lista de consideraciones a tener en cuenta a la hora de explorar formas de aprovechar e implementar las herramientas de IA.

3.1 Fundamentos de la IA

Los modelos de IA se dividen en dos grandes categorías: predictivos y generativos. Los modelos predictivos aprenden patrones en los datos y los aplican a datos nuevos que nunca han visto antes, es decir, predicen resultados potenciales. La IA predictiva

ya ha encontrado varios casos de uso en la farmacia y la sanidad en general, como el reconocimiento de comprimidos, el diseño de fármacos, el diagnóstico del cáncer y el análisis de la salud de la población. Algunos tipos de modelos predictivos son:

1. Clasificación y etiquetado
2. Predicción y prevención de riesgos
3. Recomendación

Los modelos generativos aprenden patrones de los datos para generar otros nuevos. Los resultados de la IA generativa pueden ir desde texto a hojas de cálculo, pasando por imágenes o vídeos. ChatGPT es un ejemplo popular de IA generativa. Aunque la IA generativa tiene un gran potencial y ya se está utilizando en múltiples sectores, también plantea importantes limitaciones y riesgos que deben abordarse para garantizar que se utilice de forma ética y responsable.

3.1.1 Estilos de aprendizaje para la IA

Tanto los modelos predictivos como los generativos se basan en ML. Los modelos que utilizan redes neuronales se consideran modelos de "aprendizaje profundo". Por su propia naturaleza, los modelos de aprendizaje profundo funcionan como ''cajas negras'', lo que significa que la forma en que el modelo llega a sus conclusiones es inexplicable para la propia IA o su usuario humano. Esto puede plantear obstáculos conceptuales a los organismos reguladores.

El método de aprendizaje aplicado por un determinado modelo o herramienta influirá en su funcionalidad y utilidad. No todos los métodos de aprendizaje son apropiados para un uso determinado. Estos métodos de aprendizaje son:

1. Supervisado
2. Sin supervisión
3. Refuerzo

El aprendizaje supervisado consiste en entrenar el modelo con datos previamente etiquetados. Los datos etiquetados implican que la máquina recibe una descripción de los datos que se le presentan, por ejemplo, una imagen de un gato etiquetada como "GATO". Esto forma una biblioteca de material de referencia que la máquina utiliza cuando predice resultados o genera nuevos datos basados en entradas de datos nuevos sin etiquetar. El aprendizaje supervisado se ha utilizado en la atención sanitaria para desarrollar modelos que predicen la presencia de una enfermedad, crear puntuaciones de riesgo o predecir el pronóstico de una enfermedad.[2] Un motor de IA predictiva que utiliza el aprendizaje supervisado también se ha aplicado al descubrimiento de fármacos, lo que ha dado lugar a un nuevo antibiótico de amplio espectro conocido como Halicina.[3]

Uno de los retos habituales del aprendizaje supervisado en el ámbito sanitario es disponer de suficientes datos etiquetados ya que lleva mucho tiempo crear una base de los mismos. Suele ser más fácil etiquetar cosas como imágenes médicas (por ejemplo, radiografías o resonancias magnéticas). Sin embargo, es más difícil etiquetar un acontecimiento o resultado médico, como el agravamiento de una enfermedad crónica, que requiere conocimientos especializados y la toma de decisiones. No todo el mundo puede interpretar los datos de la misma manera, lo que a su vez repercute en el resultado final del modelo. La privacidad y la seguridad también son un problema a la hora de crear y compartir conjuntos de datos etiquetados.

El aprendizaje no supervisado consiste en entrenar un modelo a partir de datos no etiquetados. Así, se utiliza el aprendizaje profundo para que la máquina identifique patrones, relaciones o similitudes en los datos. A menudo, estos patrones no pueden ser identificados por humanos que revisen los mismos datos. Un ejemplo sería realizar un análisis de conglomerados que agrupe a los pacientes en diferentes grupos de fenotipos basándose en similitudes en muchas variables o puntos de datos disponibles para cada paciente. Sin embargo, hay que tener cuidado si se utiliza este método de aprendizaje para hacer un modelo predictivo; los modelos de aprendizaje no supervisado también pueden ser propensos a resultados sin sentido o sesgados, ya que no han sido entrenados en lo que sería un resultado aceptable.

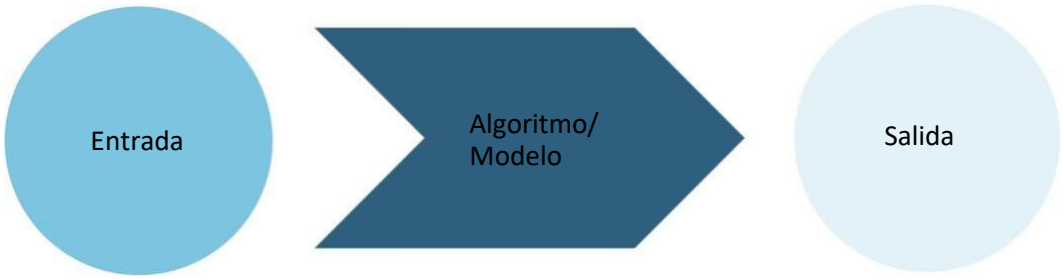
El aprendizaje supervisado y no supervisado es muy potente para identificar patrones, como textos, estructuras químicas, secuencias de ADN, imágenes y vídeos, y utilizarlos para hacer predicciones. Sin embargo, es menos potente en entornos dinámicos, donde la secuencia de una serie de acontecimientos o decisiones es importante para el resultado. Para tales fines, el aprendizaje por refuerzo puede ser más aplicable.

El aprendizaje por refuerzo "recompensa" al modelo por tomar decisiones correctas. El modelo aprende a predecir la siguiente acción o conjunto de acciones que aumentan la probabilidad de una recompensa diferida.[4] El proceso de recompensa debe programarse en el modelo, de modo que después de cada decisión que tome, reciba una recompensa positiva o una recompensa negativa.
Esto permite al modelo saber si las decisiones se toman en la dirección correcta. Se realizaron dos experimentos significativos utilizando un modelo de IA conocido como AlphaZero para jugar a los juegos de Go y ajedrez.[5]Utilizando el aprendizaje por refuerzo, el modelo aprendió qué series de movimientos del juego tenían más probabilidades de dar como resultado la recompensa retardada de ganar la partida. El modelo desarrolló estrategias novedosas que nunca se le habían ocurrido a los humanos en la historia del juego del ajedrez. En el momento de escribir estas líneas, ningún humano ha vencido a AlphaZero. El concepto de aprendizaje por refuerzo puede aplicarse al campo de la medicina, donde se toman múltiples decisiones para obtener una recompensa lejana en el futuro, como mejorar la tasa de supervivencia al alta hospitalaria o ajustar la medicación para alcanzar un objetivo determinado de A1C (hemoglobina glicosilada).

3.1.2 Tipos de herramientas de IA

La IA es una categoría tecnológica muy amplia; sin embargo, todas las herramientas o aplicaciones de IA pueden desglosarse en tres componentes principales: la entrada, el algoritmo/modelo y la salida, como se indica en la Figura 1.

Figura 1: Componentes del modelo de IA



Comprender cada uno de estos componentes es importante para entender cómo puede aprovecharse esta tecnología en cada situación. En el Cuadro 1 se describen los distintos tipos de herramientas de IA disponibles en función de los datos de entrada que utiliza el modelo.

Cuadro 1: Tipos de herramientas de IA en función de las entradas

Modelo de entrada	Descripción de la herramienta AI
Imagen	Se trata de modelos que utilizan imágenes como datos de entrada. Se trata de herramientas de diagnóstico para examinar imágenes médicas como radiografías de tórax o tomografías computarizadas. Un ejemplo es un modelo de identificación de comprimidos capaz de identificar cápsulas y comprimidos sin etiquetar con una precisión del 85%.[6]
Texto	Los modelos que utilizan texto como entrada significan que pueden interactuar con voz natural o texto. Puede tratarse de un chatbot (similar a herramientas como ChatGPT) o de un modelo de clasificación de documentos.
Voz	Estos tipos de modelos toman como entrada audio de voz sin procesar. Algunos ejemplos son los programas de dictado, los parlantes o dispositivos inteligentes, o los asistentes inteligentes en dispositivos móviles. También hay aplicaciones clínicas con IA que utilizan marcadores de voz para diagnosticar o medir los niveles de estrés de un paciente.
Tablas	Los datos tabulares incluyen cualquier tipo de dato que pueda organizarse o almacenarse en una hoja de cálculo. Los modelos que utilizan este tipo de datos incluyen algoritmos de predicción del riesgo.
Multimodal	Un modelo multimodal indica que el modelo puede aceptar más de un tipo de entrada. Por ejemplo, un chatbot puede responder a un texto, pero también a una imagen cargada.

3.2 Evaluación del rendimiento de la IA

Evaluar el rendimiento de los modelos de ML es crucial para validar su solidez técnica y su precisión, pero también para garantizar la transparencia, la responsabilidad y generar confianza entre los usuarios.

Las métricas de rendimiento son las herramientas básicas que guían a los desarrolladores y a las partes interesadas a la hora de comprender el rendimiento de un modelo en distintos escenarios, identificar sus puntos fuertes y débiles y tomar decisiones informadas sobre las mejoras y su aplicabilidad.

En la siguiente sección se resumen las metodologías y herramientas habituales utilizadas para evaluar los modelos de IA y garantizar su interpretabilidad en aras de la transparencia y la comprensión.

3.2.1 ¿Por qué medir el rendimiento?

Medir el rendimiento de los modelos de IA no es una mera necesidad técnica, sino una práctica fundamental para establecer la transparencia, la confianza y la responsabilidad. La transparencia implica proporcionar información clara y accesible sobre el funcionamiento de los modelos y su rendimiento, garantizando que los usuarios puedan entender y confiar en los conocimientos de la IA para apoyar la toma de decisiones. La medición del rendimiento garantiza que los desarrolladores y las organizaciones sean responsables del rendimiento de sus modelos, las implicancias éticas y el impacto potencial en los pacientes. Las fichas técnicas de modelo (también llamadas tarjeta de modelo o model cards) son una herramienta importante para lograrlo.

3.2.2 Cómo ser transparente, rendir cuentas y generar confianza

Las fichas técnicas de modelo sirven como guías completas y estructuradas con detalles esenciales sobre un modelo de ML.[7] Proporcionan un marco estandarizado para la información sobre las características, el rendimiento, las limitaciones y los casos de uso previstos y permiten un análisis comparativo entre diferentes modelos. Google Research desarrolló el Model Card Toolkit (MCT) para aumentar la transparencia en el ML. El MCT ayuda a los desarrolladores a crear tarjetas de modelo, que proporcionan información esencial sobre los orígenes, el uso previsto y las consideraciones éticas de un modelo. Al utilizar el MCT, los desarrolladores pueden documentar fácilmente sus modelos y asegurarse de que se utilicen de forma responsable.[8] Ver una tarjeta de modelo debería ayudar a determinar si este es apropiado para una tarea o uso específico. Una ficha de modelo típica puede incluir varios componentes, como se describe en el Cuadro 2.

Cuadro 2: Componentes de una ficha técnica de modelo

Componente	Detalle
Descripción	Breve descripción e identificación única del modelo, incluida la fecha de producción y la versión del modelo, si corresponde.
Finalidad, usuarios y contexto	Articulación clara de la finalidad prevista del modelo, respondiendo a preguntas como: para qué finalidad se desarrolló el modelo; para qué finalidad no es adecuado utilizarlo; los usuarios potenciales que pueden beneficiarse; el contexto en el que debe o no debe aplicarse; y las posibles limitaciones generales para diferentes finalidades, usuarios y contextos.
Cómo utilizarlo	Directrices comprensibles escritas en un lenguaje sencillo para garantizar que los usuarios puedan desplegar el modelo con eficacia. Si es posible, deben proporcionarse diagramas de flujo y esquemas para facilitar la comprensión del usuario.
Métricas de rendimiento	Las métricas varían según el tipo de modelo y su finalidad (por ejemplo, generación, predicción, clasificación, etc.). Es importante que las métricas de rendimiento vayan acompañadas de una interpretación de lo que significan en el contexto de uso. Siempre que sea posible, deben proporcionarse ejemplos para facilitar la comprensión del usuario, sobre todo en relación con la incertidumbre en torno a los resultados del modelo. Cuando corresponda, se debe enumerar el rendimiento del modelo en diferentes subgrupos de población para proporcionar información sobre los posibles sesgos que puedan existir.

Datos de entrenamiento	La transparencia de los datos de entrenamiento es esencial. Esto incluye la fuente, la calidad, cualquier paso de preprocesamiento aplicado y cualquier grupo o subgrupo específico que pueda estar sobre o subrepresentado. Comprender el conjunto de datos ayuda a los usuarios a calibrar la aplicabilidad del modelo y sus posibles sesgos.
Ética	Las consideraciones éticas son primordiales en la IA. Esto incluye abordar los potenciales sesgos, la imparcialidad y la privacidad. Garantizar que los modelos no perpetúen los prejuicios y que los datos de los usuarios se manejen de forma responsable es fundamental para mantener las normas éticas.
Incertidumbre	Todos los modelos tienen limitaciones e incertidumbres. Documentarlas ayuda a gestionar las expectativas de los usuarios, informarles de los riesgos potenciales y de las áreas en las que el modelo podría rendir por debajo de lo esperado.
Autor, código, licencia y propiedad	Indicar claramente la autoría, la disponibilidad del código, las condiciones de la licencia y los derechos de propiedad aumenta la transparencia y fomenta las contribuciones y mejoras de la comunidad.
Contactos y recursos	Facilitar información de contacto y recursos adicionales ayuda a los usuarios a comprender y utilizar eficazmente el modelo. También abre canales de retroalimentación y mejora.

Para comprender mejor las métricas específicas utilizadas para evaluar el rendimiento de un modelo, el Cuadro 3 ofrece una visión general de las métricas que pueden verse en la ficha de un modelo.

Cuadro 3: Métricas del modelo, su uso y ejemplos

Tipo de modelo	Métricas para medir el rendimiento	Descripción	Ejemplo
IA PREDICTIVA			
Clasificación	Exactitud Precisión Recall Puntuación F1 Sensibilidad Especificidad ROC-AUC Matriz de confusión *Para las definiciones de estas métricas, véase la siguiente sección.*	Evalúa los modelos que etiquetan o clasifican las entradas en categorías. Evalúan la exactitud de las predicciones.	**Ejemplo**: Un modelo clasifica los historiales de los pacientes en categorías como las solicitudes de reposición de recetas o las consultas sobre programación de citas. **Aplicar métricas de rendimiento:** Una precisión y exactitud elevadas indican una categorización correcta y minimizan los errores. Una alta recuperación garantiza que no se pasen por alto categorías importantes. Una elevada área bajo la curva de características operativas del receptor (ROC-AUC) combinada con la matriz de confusión (véase el cuadro 4) indica una distinción eficaz entre categorías.
Regresión	Error cuadrático medio (MSE) R^2 Error medio absoluto (MAE) Error cuadrático medio (RMSE)	Evalúa modelos que predicen valores continuos. Miden la desviación entre los valores previstos y los reales.	**Ejemplo**: Un modelo de predicción de dosis para la medicina personalizada. **Aplicación de métricas de rendimiento:** Un MSE y RMSE bajos indican una estrecha alineación entre las dosis previstas y las reales, lo que reduce los riesgos de subdosificación o sobredosificación. Un R^2 alto sugiere que el modelo efectivamente capta la relación entre las características del paciente y la dosis necesaria.

Series temporales	Error porcentual absoluto medio (MAPE) Error cuadrático medio (RMSE)	Evalúa la precisión y el error en las predicciones de valores futuros basadas en datos pasados.	**Ejemplo**: Un modelo de previsión predice la futura demanda de medicamentos. **Aplicación de métricas de rendimiento**: Un MAPE y un RMSE bajos indican previsiones precisas, lo que ayuda a mantener niveles óptimos de inventario.
IA GENERATIVA			
Generación	**Grandes modelos lingüísticos (LLM):** Perplejidad (Perplexity o PPL) [9] Subestudio de Evaluación Bilingüe (BLEU)[10, 11] Evaluación general de la comprensión lingüística (GLUE)[12] Subestudio para la evaluación de la esencia, orientado en la sensibilidad (ROUGE)[13]	Mide la similitud entre el contenido generado y el de referencia.	**Ejemplo:** Un modelo genera guiones de asesoramiento sobre medicación adaptados a los pacientes. **Aplicación de métricas de rendimiento:** Una puntuación BLEU alta indica que los guiones coinciden estrechamente con referencias de alta calidad preparadas por farmacéuticos clínicos, lo que garantiza que los pacientes reciban consejos precisos y comprensibles.

3.3 Métricas de rendimiento de los modelos de clasificación

Está fuera del alcance de este kit de herramientas proporcionar una visión exhaustiva de todas las métricas de rendimiento, por lo que esta sección entra en mayor detalle en torno a algunas de las métricas de rendimiento de modelos más comunes con las que uno se puede encontrar. La precisión, el recall, la puntuación F1, la especificidad y la sensibilidad son métricas utilizadas con frecuencia para evaluar la eficacia de un modelo de predicción mediante clasificación.[14] La curva ROC-AUC y la matriz de confusión son herramientas útiles para comprender la capacidad de un modelo para distinguir entre diferentes clases en la predicción. A continuación se explican los fundamentos teóricos de estas métricas.

- **La precisión**, también conocida como valor predictivo positivo, mide la exactitud de las predicciones positivas realizadas por el modelo. Resulta especialmente útil en escenarios en los que el costo de los falsos positivos es elevado. La precisión responde a la pregunta "De todas las instancias previstas como positivas, ¿cuántas son realmente positivas?".

  La fórmula de la precisión viene dada por:

$$Precisión = \frac{TP}{TP + FP}$$

  TP (Verdaderos positivos) son las instancias correctamente previstas como positivas.
  FP (Falsos positivos) son las instancias incorrectamente previstas como positivas.

- **El recall**, también conocido como sensibilidad o tasa de verdaderos positivos, mide la capacidad del modelo para identificar todos los casos relevantes. Es crucial en situaciones en las que omitir un caso positivo es costoso. El recall responde a la pregunta: "De todos los casos positivos reales, ¿cuántos identificó correctamente el modelo?".

  La fórmula para el recall es:

$$Sensivilidad\ (Recall) = \frac{TP}{TP + FN}$$

  TP (Verdaderos positivos) son las instancias correctamente previstas como positivas.
  FN (Falsos negativos) son las instancias incorrectamente previstas como negativas.

- **La puntuación F1 (F1 Score)** es la media de la precisión y la sensibilidad (o recall), lo que proporciona un equilibrio entre ambas métricas. Resulta especialmente útil cuando hay una distribución desigual de las clases (es decir, una de las clases es poco frecuente) o cuando hay que minimizar tanto los falsos positivos como los falsos negativos.

  La fórmula para la puntuación F1 es:

$$Puntuación\ F1 = 2\ x\ \frac{Precision\ x\ Recall}{Precision\ +\ Recall}$$

- **La especificidad**, también conocida como tasa de verdaderos negativos, mide la capacidad del modelo para identificar correctamente los casos negativos. Es importante en situaciones en las que el costo de los falsos negativos es elevado. La especificidad responde a la pregunta "De todos los casos negativos reales, ¿cuántos identificó correctamente el modelo?".

  La fórmula de la especificidad es:

$$Especificidad = \frac{TN}{TN + FP}$$

- **La curva ROC AUC** es una representación gráfica del rendimiento de un clasificador. La curva ROC (Receiver Operating Characteristic) representa la tasa de verdaderos positivos (sensibilidad) frente a la tasa de falsos positivos (1 - especificidad) en varios umbrales. El área bajo la curva (AUC) proporciona una medida agregada del rendimiento del modelo en todos los umbrales. Un modelo con un AUC de 1 es perfecto, mientras que un AUC de 0,5 indica que no tiene capacidad de discriminación.

- **La matriz de confusión** es una tabla utilizada para describir el rendimiento de un modelo de clasificación en un conjunto de datos de prueba cuyos valores verdaderos se conocen. Permite visualizar el rendimiento de un algoritmo. La matriz es N x N, donde N es el número de clases. Proporciona información sobre los tipos de errores cometidos por el clasificador.

La matriz de confusión del Cuadro 4 tiene la siguiente estructura:

Cuadro 4: Matriz de confusión

	Predicción positiva	**Predicción negativa**
Positivo real	Verdadero positivo (TP)	Falso negativo (FN)
Actual negativo	Falso positivo (FP)	Verdadero negativo (VN)

Evaluar el rendimiento de los modelos de IA mediante métodos estructurados y transparentes es crucial para fomentar la confianza y la responsabilidad. Al dar prioridad a la transparencia, las consideraciones éticas y la participación de los usuarios finales en el proceso de evaluación, los desarrolladores de IA pueden crear modelos que no solo sean técnicamente sólidos, sino también fiables y acordes con las necesidades de los usuarios.

4 Consideraciones éticas

4.1 Crear una IA digna de confianza

Los recientes avances en ML y la IA han transformado varios sectores, entre ellos la sanidad. Aunque estas tecnologías ofrecen enormes beneficios potenciales, también plantean dilemas éticos que deben abordarse para garantizar un uso responsable y equitativo. Las "Directrices éticas para una IA digna de confianza" del Grupo de Expertos de Alto Nivel sobre Inteligencia Artificial (HLEG sobre IA), encargadas por la Unión Europea, UE, en 2019, proporcionan directrices integrales para garantizar un desarrollo y despliegue éticos y dignos de confianza de la IA[15] Según las Directrices, la IA digna de confianza debe ser:

1. **Legal**: respetando todas las leyes y normativas aplicables;

2. **Ética**: respeto de los principios y valores éticos.

3. **Robustez**: desde el punto de vista técnico y teniendo en cuenta su entorno social.

Las directrices esbozan siete requisitos clave que deben cumplir los sistemas de IA para ser considerados dignos de confianza (véase el cuadro 5).

Cuadro 5: Siete requisitos clave para una IA fiable

Requisito clave	Detalles
1. Intervención humana y supervisión	Los sistemas de IA deben apoyar la autonomía humana y la toma de decisiones, siendo los humanos los responsables últimos de los resultados de la IA.
2. Solidez técnica y seguridad	Los sistemas de IA deben ser seguros, fiables y resistentes durante todo su ciclo de vida para evitar daños no deseados.
3. Privacidad y gobernanza de datos	Los sistemas de IA deben respetar la privacidad, garantizando la protección de los datos personales y el cumplimiento de la normativa sobre protección de datos.
4. Transparencia	Los procesos y decisiones tomados por los sistemas de IA deben ser explicables, comprensibles y accesibles para los usuarios.
5. Diversidad, no discriminación y equidad	Los sistemas de IA deben ser inclusivos y justos, evitando los prejuicios y la discriminación basados en diversos atributos.
6. Bienestar social y medioambiental	El desarrollo y el despliegue de la IA deben beneficiar a la sociedad y al medio ambiente, promoviendo la sostenibilidad y el bienestar social.
7. Responsabilidad	Las partes implicadas en los sistemas de IA (desarrolladores, implementadores, etc.) deben ser responsables de sus decisiones y acciones relacionadas con la IA.

4.2 Consideraciones éticas sobre la aplicación de la IA a la asistencia sanitaria

Las implicancias éticas del uso del ML y la IA en la asistencia sanitaria son complejas y polifacéticas. Aunque estas tecnologías ofrecen un potencial transformador, como la mejora de la seguridad del paciente, también plantean problemas éticos, como la parcialidad en la atención al paciente cuando influyen en la toma de decisiones clínicas.[16]

Un ejemplo de preocupación ética importante son las "pruebas no concluyentes", que ponen de relieve que los algoritmos son probabilísticos y nunca infalibles.[17] Por ejemplo, un reloj inteligente podría diagnosticar incorrectamente un latido irregular del corazón debido a imprecisiones en las mediciones de la frecuencia cardiaca. Estos errores pueden surgir cuando los algoritmos se calibran en función de las normas generales de la población -como tonos de piel específicos- en lugar de las características individuales de los pacientes. Esto demuestra cómo los sesgos en el desarrollo de la IA pueden dar lugar a disparidades en la atención al paciente.

Además, el despliegue responsable de la IA en la atención sanitaria debe tener en cuenta la seguridad del paciente, la privacidad, la transparencia, la equidad, el consentimiento informado y las implicaciones para el personal, especialmente en la práctica farmacéutica.[18] Para garantizar la integración ética de la IA es necesario considerar cuidadosamente los cuatro principios fundamentales[19] de la ética médica:

- Autonomía: derecho del paciente a la autodeterminación;
- Beneficencia: la responsabilidad de "hacer el bien";
- No maleficencia: la responsabilidad de "no hacer daño"; y
- Justicia: tratar a todos por igual y equitativamente.

Si afrontamos estos retos éticos de forma proactiva, podremos maximizar los beneficios de la IA en la atención sanitaria al tiempo que respetamos las normas éticas y garantizamos el bienestar de los pacientes.

El reto de definir la IA ética: A pesar del reconocimiento generalizado de que la IA debe adherirse a principios éticos, existe un debate permanente sobre la definición práctica de "IA ética". Esto incluye incertidumbres en torno a los requisitos éticos específicos, las normas técnicas y las mejores prácticas de aplicación. La Organización Mundial de la Salud (OMS) ha esbozado una serie de principios éticos clave para la IA en la atención sanitaria, entre ellos: proteger la autonomía; promover el bienestar humano; garantizar la transparencia, la explicabilidad y la inteligibilidad; fomentar la responsabilidad y la rendición de cuentas; garantizar la inclusión y la equidad; y promover un desarrollo receptivo.[20] Del mismo modo, una revisión del panorama mundial de las directrices éticas de la IA identificó cinco principios éticos convergentes: transparencia, justicia e imparcialidad, no maleficencia, responsabilidad y privacidad.[21] Teniendo en cuenta estos principios éticos, la siguiente sección profundiza en las cuestiones éticas clave que deben tenerse en cuenta para la aplicación responsable de la IA en la asistencia sanitaria.

4.2.1 Cuestiones éticas

4.2.1.1 Privacidad y seguridad

- El uso del ML y la IA implica el manejo de grandes volúmenes de datos sensibles de pacientes, planteando preocupaciones por las violaciones de la privacidad y la seguridad de los datos.
- Los métodos de anonimato y cifrado de datos deben ser sólidos para proteger la confidencialidad de los pacientes.
- Los investigadores y usuarios finales deben tener en cuenta los riesgos de confidencialidad y privacidad asociados a los datos utilizados.
- Las medidas de privacidad y seguridad deben aplicarse tanto a los datos de formación introducidos en la máquina como a los resultados del ML.

4.2.1.2 Transparencia

- Los modelos de ML pueden ser complejos y difíciles de interpretar. Los investigadores y desarrolladores de IA deben tener en cuenta la transparencia y explicabilidad (o interpretabilidad) de sus modelos.[22]
- Los algoritmos de inteligencia artificial y ML suelen funcionar como "cajas negras", lo que dificulta la comprensión de sus procesos de toma de decisiones.

- Garantizar la transparencia del ML permite mejorar la reproducibilidad y la comprensión de los procesos subyacentes.

- Garantizar la transparencia de los resultados algorítmicos y establecer la responsabilidad por los errores o sesgos algorítmicos son consideraciones éticas fundamentales.[22]

4.2.1.3 Responsabilidades

- En todo el proceso de ML, la responsabilidad es primordial. Los modelos deben utilizarse únicamente para los fines previstos, y las partes interesadas deben ser conscientes de sus responsabilidades.

- Los investigadores, los desarrolladores de IA y los usuarios finales, como los médicos, deben mantener la responsabilidad para evitar consecuencias no deseadas o el uso indebido de los resultados del ML.

4.2.1.4 Sesgos y equidad

- Los modelos de ML entrenados en conjuntos de datos sesgados pueden perpetuar o amplificar los sesgos existentes en la atención sanitaria, lo que conduce a un tratamiento desigual.

- El sesgo también puede introducirse a través de la propia arquitectura del modelo y de las decisiones tomadas durante el proceso de entrenamiento.[23]

- Las prácticas éticas de ML requieren estrategias para identificar y mitigar los sesgos en los datos y algoritmos para garantizar la justicia y la equidad en la prestación de asistencia sanitaria.

4.2.1.5 Consentimiento informado

- Es posible que los pacientes no comprendan del todo las implicancias de los diagnósticos o las recomendaciones de tratamiento basados en la IA.[24]

- Garantizar el consentimiento informado para las intervenciones con IA implica educar a los pacientes sobre el rol y las limitaciones de la IA en la toma de decisiones sanitarias.

4.2.1.6 Impacto en los profesionales sanitarios

- Las tecnologías de ML e IA pueden alterar las funciones y responsabilidades de los profesionales sanitarios, lo que podría provocar el desplazamiento o la reducción de la fuerza de trabajo.

- Las directrices éticas deben abordar las implicancias éticas de los cambios impulsados por la tecnología en las prácticas profesionales y las actividades laborales.

4.2.2 Estrategia de mitigación para el despliegue ético de la IA en la atención sanitaria

Para abordar eficazmente los retos éticos asociados al despliegue de la IA en la atención sanitaria, las partes interesadas de los sectores sanitario y tecnológico deben colaborar para desarrollar directrices y normativas exhaustivas. En el cuadro 6 se esbozan las estrategias y acciones clave de mitigación para la integración ética de la IA, centrándose en la mejora de la toma de decisiones éticas y la participación activa de los farmacéuticos y otras partes interesadas.

Cuadro 6: Estrategias para el despliegue ético de la IA

Estrategia	Acciones
Establecer directrices éticas claras	• Elaborar y difundir directrices éticas que rijan el desarrollo y la implantación de las tecnologías de IA en la atención sanitaria, abordando cuestiones clave como la parcialidad, la transparencia, la privacidad del paciente y las consideraciones del medio ambiente.

Integrar la formación ética en los programas educativos		• Incorporar la formación ética en los planes de estudios de los profesionales sanitarios especializados en salud digital, incluida la IA.[24] • Ofrecer oportunidades de desarrollo profesional continúo centradas en el uso ético de la IA tanto en la práctica clínica como en la ciencia.
Implantar mecanismos sólidos de auditoría y validación		• Desarrollar mecanismos de auditoría y validación periódica de los algoritmos de IA para detectar y corregir sesgos y garantizar la equidad. • Contratar a auditores externos para que realicen evaluaciones independientes de los sistemas de IA.
Implicar a pacientes y comunidades		• Involucrar activamente a pacientes y comunidades en los debates sobre el rol y el impacto de la IA en los procesos de toma de decisiones sanitarias. • Garantizar la transparencia con los pacientes sobre cómo se utilizan sus datos y obtener su consentimiento explícito.
Implementar un marco reforzado de toma de decisiones éticas	Amplia participación de las partes interesadas	• Involucrar a una amplia gama de partes interesadas, incluidos los desarrolladores de IA y ML, los responsables de la toma de decisiones sanitarias, los comités de ética de la investigación, los reguladores y otros comprometidos con la promoción del uso equitativo y responsable de las tecnologías de ML clínicas.
	Integración de normas éticas	• Alinear el diseño ético de las herramientas de IA y ML con las normas establecidas en el entorno sanitario, abarcando los principios de la ética biomédica, la ética de la investigación clínica y la justicia social.
	Orientación ética holística a lo largo del ciclo de vida de la IA y el ML	• Hacer hincapié en las consideraciones éticas en cada etapa de la integración de la IA y el ML en la asistencia sanitaria, reflexionando sobre las implicaciones de estos principios para fomentar la equidad en las aplicaciones de ML.
Capacitar a los farmacéuticos como usuarios finales	Evaluar la transparencia de los modelos	• Formar a los farmacéuticos para que interpreten las recomendaciones de la IA y comprendan los procesos de toma de decisiones subyacentes.[25] • Asegurarse de que los sistemas de IA cuentan con una documentación exhaustiva en la que se detallan su diseño, los datos de entrenamiento y los fundamentos de las decisiones (véase la sección sobre Evaluación del rendimiento de los modelos).
	Identificar y mitigar los prejuicios	• Supervisar periódicamente los resultados de la IA en busca de indicios de parcialidad e informar de cualquier incoherencia a las autoridades o promotores pertinentes. • Utilizar mecanismos de retroalimentación para informar de los sesgos y colaborar con los desarrolladores para mejorar la imparcialidad de los modelos.
	Garantizar la justicia y la equidad	• Abogar por un uso equitativo de la IA en todos los grupos demográficos de pacientes, garantizando que la IA no tenga un impacto desproporcionado en las poblaciones vulnerables.[23] • Verificar las recomendaciones de la IA, sobre todo en casos complejos, para garantizar que mejoran la calidad de la atención al paciente.

	Educación y formación continuas	• Participe en el desarrollo profesional continuo para mantenerse actualizado sobre los avances de la IA y las consideraciones éticas. • Reflexionar sobre las implicaciones éticas de la IA en la práctica diaria e integrar estas consideraciones en la atención al paciente. • Crear una cultura de responsabilidad entre los usuarios finales, que tienen la responsabilidad de prevenir los daños involuntarios causados por los modelos que utilizan.
	Participación de los pacientes y la comunidad	• Comunicarse abiertamente con los pacientes sobre el rol de la IA en su atención, explicándoles los fundamentos de las recomendaciones basadas en la IA. • Fomentar el diálogo comunitario para abordar las preocupaciones de los ciudadanos y aumentar la confianza en las tecnologías de IA.
	Gestión proactiva del riesgo	• Desarrollar estrategias para anticipar y abordar posibles dilemas éticos, como errores de IA, alucinaciones o consecuencias imprevistas. • Colaborar con otros profesionales sanitarios para compartir las mejores prácticas y mejorar la utilización de la IA en la farmacia.

Mediante la aplicación de las estrategias de mitigación del Cuadro 6, los farmacéuticos pueden desempeñar un rol crucial para garantizar el uso ético y responsable de la IA en la asistencia sanitaria. Dotados de las herramientas y conocimientos adecuados, los farmacéuticos pueden ayudar a mitigar los riesgos y maximizar los beneficios de la IA para la atención al paciente, contribuyendo a un enfoque más transparente, justo y centrado en el paciente, fomentando la confianza y la responsabilidad en las aplicaciones de IA.

5 Normativa ISO y cumplimiento

A la hora de integrar la IA en la práctica farmacéutica, es esencial que los farmacéuticos naveguen por varias áreas críticas de forma eficaz para garantizar el cumplimiento de aspectos importantes como la confidencialidad de los datos de los pacientes, la calidad, la reproducibilidad de los algoritmos y la seguridad de las Tecnologías de la Información y las Comunicaciones, TIC. Los sistemas de IA utilizados en la práctica farmacéutica deben cumplir normativas sanitarias y de tecnología de la información específicas que varían según el país o la región.

5.1 Normativa general sobre IA

Las normativas sobre IA varían significativamente entre las distintas regiones del mundo, reflejando el enfoque único de cada zona para equilibrar la innovación con la supervisión. En todas las regiones se observa una clara tendencia a establecer marcos que garanticen que la IA se utiliza de forma ética y responsable sin limitar la innovación. Mientras que Europa hace hincapié en marcos reguladores sólidos, América se caracteriza por un enfoque más descentralizado y sectorial. Asia-Pacífico muestra una estrategia mixta. El cuadro 7 describe con más detalle el panorama normativo de cada región.

Cuadro 7: Regulación regional de la IA

Región	Panorama normativo
Europa	La Unión Europea (UE) ha sido proactiva con su exhaustiva "Ley de IA", que clasifica las aplicaciones de IA en función de los niveles de riesgo y establece requisitos estrictos para los usos de alto riesgo, incluida la vigilancia biométrica y la manipulación subliminal. Esta legislación hace hincapié en la transparencia, la responsabilidad y la equidad en el uso de la IA, con el objetivo de alinearse con los valores europeos de supervisión humana y privacidad. La estrategia del Reino Unido se centra en los principios de seguridad y transparencia, al tiempo que hace hincapié en evitar normativas excesivamente restrictivas para fomentar el desarrollo de la IA y promover la innovación.
América	En EE.UU. aún no existe una legislación unificada sobre IA; en su lugar, el país emplea diversas directrices y marcos para gestionar las aplicaciones de IA, a menudo centrados en medidas específicas de un sector. Canadá ha introducido la Ley de IA y Datos (AIDA), que promueve la seguridad y las prácticas responsables de IA, y Brasil está desarrollando activamente una legislación integral de IA para regular los sistemas de IA de alto riesgo y garantizar la transparencia y la responsabilidad en los despliegues de IA.
Asia-Pacífico	China ha dado pasos significativos en la regulación de la IA con leyes específicas para las recomendaciones algorítmicas y las tecnologías de síntesis profunda, con el objetivo de garantizar la integridad y la equidad de los contenidos, manteniendo al mismo tiempo el control sobre las innovaciones de la IA. Japón se basa más en directrices no vinculantes y normas sectoriales específicas que en una legislación nacional exhaustiva, promoviendo un enfoque flexible para apoyar la innovación en IA. India y otros países asiáticos están formulando actualmente su enfoque, lo que indica un cambio hacia una normativa más definitiva en breve.
África	Varios países se encuentran en distintas fases de desarrollo de estrategias y políticas nacionales para gestionar y aprovechar los beneficios de la tecnología de IA. La legislación y las directrices siguen dependiendo en gran medida de las diversas leyes de protección de datos. Países como Sudáfrica, Mauricio, Egipto y Kenia son pioneros en la redacción y aplicación de directrices sobre IA que aborden tanto las oportunidades como los retos éticos que plantea la IA. Estas iniciativas suelen considerar la adaptación de las normas mundiales, como la Ley de IA de la UE, a los contextos locales, fomentando la innovación al tiempo que se salvaguardan los valores éticos y las normas sociales. La Unión Africana también está desempeñando un papel, lo que sugiere un avance hacia una estrategia continental más unificada. Sin embargo, retos como la limitada infraestructura tecnológica y la necesidad de expertos locales en IA siguen siendo obstáculos importantes. Estas diferencias regionales resaltan la importancia de la colaboración y el diálogo internacional para armonizar la gobernanza de la IA, garantizando que las normas mundiales puedan adaptarse a las necesidades locales y promoviendo al mismo tiempo un desarrollo seguro y beneficioso de la IA en todo el mundo.

5.2 Protección de datos y privacidad en general en la sanidad

Proteger la confidencialidad, integridad y disponibilidad de los datos de los pacientes es primordial. Los sistemas de IA procesan grandes cantidades de datos sensibles, por lo que garantizar su protección frente a accesos no autorizados, infracciones y filtraciones es un requisito ético y jurídico fundamental. Las medidas eficaces de protección de datos ayudan a mantener la confianza de los pacientes, como concepto básico de la prestación de asistencia sanitaria. La falta de protección de los datos puede acarrear pérdidas económicas, pérdida de la confianza de los pacientes y graves sanciones reglamentarias, incluido el enjuiciamiento penal en determinadas jurisdicciones.

En Estados Unidos, la Ley de Portabilidad y Responsabilidad de los Seguros Sanitarios (HIPAA, por sus siglas en inglés) es el principal marco normativo que deben cumplir las farmacias para proteger la información de los pacientes.[27] Las farmacias deben asegurarse de que todos los sistemas que manejan datos de pacientes, incluidas las tecnologías de IA, cumplen las normas de la HIPAA para evitar sanciones como las impuestas a entidades como CVS Pharmacy y Rite Aid por eliminación indebida de información sanitaria personal (PHI, por sus siglas en inglés). Las farmacias de EE.UU. deben adoptar un cifrado robusto para los datos almacenados y protocolos de transmisión seguros para cumplir los requisitos de la HIPAA.

En Europa, el Reglamento General de Protección de Datos (RGPD) ofrece protecciones similares.[28] En virtud del RGPD, los farmacéuticos deben garantizar la confidencialidad de los datos de los pacientes con estrictos controles de acceso y vigilar los accesos no autorizados. Los propietarios de farmacias deben realizar revisiones periódicas de las políticas de conservación de datos y asegurarse de que el personal comprende la importancia de la confidencialidad de los pacientes.

La Unión Africana ha adoptado el Convenio sobre Ciberseguridad y Protección de Datos Personales para facilitar la armonización de las legislaciones de los Estados miembros. Este convenio no es vinculante, pero sirve de orientación. En todo el continente, 36 de los 55 países han promulgado leyes de protección de datos. Sudáfrica ha promulgado la Ley de Protección de Datos Personales (POPIA), que se aplica a cualquier tratamiento de información personal. Impone consecuencias penales por no proteger adecuadamente los datos personales. La ley estipula los derechos y responsabilidades de los titulares de los datos y los usuarios, las medidas de seguridad necesarias y cómo puede enviarse la información fuera del país, lo que tiene implicaciones para cualquier tratamiento de datos en el extranjero. Del mismo modo, Egipto, Kenia, Nigeria y Zimbabue han aplicado sus respectivas leyes de protección de datos.

5.2.1 IA centrada en el ser humano

Los farmacéuticos deben mantener un enfoque centrado en el ser humano, garantizando que la IA apoye, pero no sustituya, la interacción personal y la atención que son fundamentales en la asistencia sanitaria. La equidad en el acceso a la asistencia sanitaria es otra consideración vital; los farmacéuticos deben vigilar que las herramientas de IA no exacerben inadvertidamente las disparidades favoreciendo a ciertas poblaciones en detrimento de otras. La formación continua sobre los avances de la IA y las directrices éticas es esencial, ya que capacita a los farmacéuticos para tomar decisiones informadas y abogar por el uso responsable de la tecnología. Al equilibrar la innovación con la vigilancia ética, los farmacéuticos pueden aprovechar la IA para mejorar la atención al paciente al tiempo que defienden los valores fundamentales de la profesión: confianza, confidencialidad y equidad.

5.2.2 Validación clínica y reproducibilidad de los algoritmos

La regulación de las aplicaciones de IA varía según el país. Las aplicaciones autónomas de IA son aquellas que funcionan sin un ser humano que verifique o compruebe los resultados. Algunos países exigen una validación clínica sólida para la IA autónoma, lo que significa que la aplicación de IA debe probarse y demostrar su seguridad y eficacia para el uso previsto. La regulación de este tipo de aplicaciones suele incluirse en la normativa sobre productos sanitarios. Las aplicaciones que no son autónomas pueden no requerir una regulación tan estricta.

En Europa, el Reglamento de Dispositivos Médicos (MDR) es crucial para los farmacéuticos que integran la tecnología de IA en la atención sanitaria, ya que proporciona un marco integral que garantiza la seguridad, eficacia y calidad de los dispositivos médicos, incluidas las herramientas impulsadas por IA.[29] Según el MDR, las tecnologías de IA clasificadas como dispositivos médicos deben cumplir estrictos requisitos de gestión de riesgos, evaluación clínica y vigilancia posterior a la comercialización. Para los farmacéuticos, la adhesión al MDR es esencial para garantizar que las herramientas de IA utilizadas en la atención al paciente sean fiables, seguras y cumplan las normas europeas. Este reglamento también hace hincapié en la transparencia y la trazabilidad, garantizando que los sistemas de IA no solo sean eficaces, sino también explicables y responsables, protegiendo en última instancia la seguridad del paciente y aumentando la confianza en las soluciones sanitarias basadas en IA.

En EE.UU., la tecnología de IA en la atención sanitaria está regulada principalmente por la Administración de Alimentos y Medicamentos (FDA) bajo sus marcos de Salud Digital y Software como Dispositivo Médico (SaMD).[29,30] La FDA evalúa las herramientas impulsadas por IA en función de su uso previsto, el riesgo potencial para los pacientes y si cumplen los criterios para un dispositivo médico. Para las tecnologías de IA clasificadas como dispositivos médicos, la FDA exige un riguroso proceso de revisión, que incluye la aprobación o autorización previa a la comercialización, para garantizar la seguridad, la eficacia y la calidad. La agencia también hace hincapié en la vigilancia posterior a la comercialización y en la necesidad de aprendizaje y adaptación continuos de los sistemas de IA, garantizando que mantengan los estándares de rendimiento a lo largo del tiempo. Este enfoque normativo pretende equilibrar la innovación con la seguridad del paciente, proporcionando un marco sólido para la integración de la IA en la asistencia sanitaria. Sin embargo, las aplicaciones de IA que se utilizan únicamente como apoyo a la toma de decisiones clínicas no suelen estar obligadas a someterse a pruebas y demostrar su seguridad y eficacia antes de su uso, ya que existe una red de seguridad humana. Aun así, estas aplicaciones de IA deben desplegarse de forma responsable, de modo que se promueva la seguridad y se minimicen los riesgos, entendiendo que la carga de la responsabilidad recae en el usuario clínico final para garantizar la eficacia y la seguridad. Es importante saber distinguir entre las aplicaciones de IA que requieren validación y las que no.

En Asia, la regulación de la tecnología de IA en la atención sanitaria varía de un país a otro, pero varias naciones líderes como Japón, China y Corea del Sur están desarrollando marcos integrales para regular el uso de la IA en contextos médicos. En Japón, la Agencia de Productos Farmacéuticos y Dispositivos Médicos (PMDA) supervisa las tecnologías de IA que se consideran dispositivos médicos, centrándose en la seguridad, la eficacia y la supervisión posterior a la comercialización, de forma similar al enfoque de la FDA. China ha puesto en marcha una estricta normativa a través de la Administración Nacional de Productos Médicos (NMPA), que exige pruebas rigurosas y la aprobación de los dispositivos médicos de IA, con un fuerte énfasis en la seguridad de los datos y la privacidad del paciente. Corea del Sur, bajo el Ministerio de Seguridad Alimentaria y Farmacéutica (MFDS), también regula la IA en la atención sanitaria, exigiendo procesos de aprobación y garantizando la supervisión continua de las herramientas de IA en el mercado. En general, aunque los marcos normativos en Asia siguen evolucionando, existe una clara tendencia a alinearse con las normas internacionales, garantizando que la IA en la asistencia sanitaria sea segura, eficaz y protegida en toda la región.

5.2.3 Normas de interoperabilidad

Garantizar que los sistemas de IA en la práctica farmacéutica se adhieren a estándares de interoperabilidad como Health Level 7 (HL7) o Fast Health Interoperability Resource (FHIR) es crucial para un intercambio de datos fluido entre diferentes sistemas sanitarios.[31] Esta compatibilidad ayuda a mantener un entorno de información sanitaria unificado y eficiente.

5.2.4 Medidas de ciberseguridad

Las farmacias deben aplicar fuertes medidas de ciberseguridad para protegerse de las amenazas. Esto incluye el empleo de cifrado, controles de acceso seguros y auditorías de seguridad periódicas para salvaguardar la información de los pacientes. Se recomiendan estrategias como la autenticación de dos factores y políticas de contraseñas sólidas para mejorar la seguridad. ISO27001 e ISO27799 para proteger los datos sanitarios son certificaciones esenciales a tener en cuenta.

5.2.5 Accesibilidad e inclusión

Las herramientas de IA deben ser accesibles para todos los usuarios, incluidos los menos capacitados, y estar diseñadas para prevenir las desigualdades en la atención sanitaria. Esto requiere un diseño cuidadoso y prácticas de implementación para garantizar que los sistemas de IA sean utilizables por una población de pacientes diversa y no excluyan inadvertidamente a ningún grupo.

La IA puede diseñarse para hacer hincapié en la inclusión garantizando que los algoritmos se entrenen con conjuntos de datos diversos y representativos, lo que ayuda a mitigar los sesgos y mejora el rendimiento del sistema en diferentes grupos demográficos. Por ejemplo, en la atención sanitaria, la IA puede desarrollarse para tener en cuenta distintos antecedentes genéticos, factores socioeconómicos y diferencias de género, garantizando que las herramientas de diagnóstico y las recomendaciones de tratamiento sean eficaces para todas las poblaciones.

Además, las interfaces de IA pueden diseñarse para que sean accesibles a las personas con capacidades diferentes, incorporando funciones como el reconocimiento de voz para las personas con movilidad limitada o problemas visuales. El soporte multilingüe y el contenido culturalmente sensible son otras formas en que la IA puede adaptarse para servir a una gama más amplia de usuarios, garantizando que la tecnología beneficie a todos, independientemente de sus antecedentes o capacidades.

Al abordar estas áreas, los farmacéuticos pueden utilizar eficazmente la IA para complementar su experiencia, permitiendo una atención al paciente más segura, eficaz y personalizada. Cada región puede tener requisitos normativos específicos, y mantenerse informado a través de la formación continua y el seguimiento de las actualizaciones normativas es esencial para mantener el cumplimiento y garantizar el uso eficaz de la IA en la práctica farmacéutica.

5.2.6 Consideraciones para el futuro

A medida que los proveedores de atención farmacéutica implementan cada vez más la tecnología de IA, las tendencias futuras en la regulación de la IA a tener en cuenta incluyen la posibilidad de leyes de privacidad de datos más estrictas, en particular en relación con los datos de los pacientes, y el establecimiento de directrices más claras sobre la responsabilidad y la transparencia de la IA. Los reguladores pueden exigir que los sistemas de IA en la atención sanitaria se sometan a rigurosas pruebas y procesos de validación para garantizar la seguridad y la eficacia, de forma similar al proceso de aprobación de nuevos medicamentos. Además, podría haber un impulso para establecer marcos más sólidos que garanticen que las decisiones impulsadas por la IA sigan siendo explicables e interpretables para los profesionales sanitarios y los pacientes, con el fin de mantener la confianza y defender las normas éticas en la atención al paciente.

6 Obstáculos comunes a la implementación

Aunque la precisión y el rendimiento de un modelo de IA son cruciales, su eficacia clínica final viene determinada por su aplicación práctica en el entorno de despliegue. Incluso los modelos de IA más avanzados pueden quedarse cortos si no se integran eficazmente en la práctica. El despliegue de los modelos de IA presenta diversos obstáculos y retos que deben abordarse para garantizar su implantación y repercusión satisfactorias.

6.1 Sesgo algorítmico

El sesgo en la IA se produce cuando los algoritmos producen una discriminación sistemática e injusta contra determinados individuos o grupos. Esto puede deberse a datos de entrenamiento sesgados, un diseño defectuoso del modelo, la desviación de datos y conceptos, o un despliegue inadecuado. En el contexto de la atención sanitaria y los productos farmacéuticos, las consecuencias de estos sesgos pueden ser especialmente graves, afectando a los resultados de los pacientes y a la eficacia de los tratamientos.[32,33] Afortunadamente, existen estrategias que pueden aplicarse para ayudar a abordar el sesgo, como se describe en el Cuadro 8. Si bien es posible que estas estrategias no eliminen el sesgo, pueden ayudar a mitigarlo.

Cuadro 8: Estrategias clave para abordar el sesgo algorítmico

Estrategia	Detalle
Recogida de datos diversos	Garantizar que los datos de formación sean representativos de poblaciones diversas puede ayudar a mitiga el sesgo. Esto implica recopilar datos de diversos orígenes demográficos, geográficos y socioeconómicos.
Equidad algorítmica	La aplicación de algoritmos que tengan en cuenta la imparcialidad, comprueben y corrijan activamente los sesgos durante el entrenamiento del modelo puede evitar resultados sesgados. Dados los diversos modelos de imparcialidad, cada uno aborda diferentes aspectos del sesgo, como la paridad demográfica Para seguir aprendiendo, el Instituto Alan Turing ofrece un curso[34] sobre evaluación y mitigación del sesgo y la discriminación en la IA, que proporciona técnicas prácticas y una visión completa de la equidad algorítmica.
Control continuo	Es esencial realizar auditorías periódicas de los sistemas de IA para detectar y corregir los prejuicios. Esto implica establecer circuitos de retroalimentación en los que se evalúe continuamente la imparcialidad de los resultados del sistema. Existen múltiples definiciones de imparcialidad. Castelnovo et al.[35] definen la imparcialidad a través de los siguientes conceptos: • Imparcialidad individual: Individuos similares deben recibir resultados similares. Esto incluye conceptos como la Equidad a través de la conciencia (FTA) y la Equidad a través de la ignorancia (FTU), que tratan a los individuos de forma similar en función de sus características sin tener en cuenta atributos sensibles como el género. • Imparcialidad grupal: Se centra en tratar a los grupos por igual e incluye: o Independencia (paridad demográfica): Las decisiones deben ser independientes de los atributos sensibles. o Separación (igualdad de probabilidades): Garantizar tasas de error iguales en todos los grupos. o Suficiencia (calibración): Garantizar que las predicciones son igual de precisas en todos los grupos.
Transparencia y explicabilidad	Hacer que los sistemas de IA sean transparentes y que sus procesos de toma de decisiones sean explicables puede ayudar a identificar y rectificar los sesgos. Las partes interesadas deben comprender cómo y por qué los sistemas de IA toman determinadas decisiones. Aunque los algoritmos de aprendizaje profundo son intrínsecamente inexplicables, pueden utilizarse algunas herramientas para ayudar a predecir cómo toma sus decisiones el modelo.
Añade tus propios barreras	El caso de la IA generativa es distinto. En los resultados del modelo se generan datos que no existen. Dado que estos modelos no están entrenados para ser fácticos, una de sus principales preocupaciones son las alucinaciones, en las que el modelo proporciona una respuesta que es falsa o inexacta. Además, estos modelos no son deterministas: cuando se les da una sola indicación, hay muchas respuestas correctas (e incorrectas). En este caso, la ingeniería pronta desempeña un papel crucial.

6.2 Modelo de deriva

La deriva del modelo, que se produce cuando el rendimiento de un modelo de IA se deteriora con el tiempo debido a cambios en los patrones de datos subyacentes, es un problema crítico.[36] Si no se gestiona adecuadamente, puede hacer que el modelo se vuelva cada vez más impreciso con el paso del tiempo.
El cuadro 9 enumera las estrategias para minimizar el impacto de este inconveniente.

Cuadro 9: Estrategias para gestionar la deriva del modelo

Estrategia	Detalle
Entrenamiento regular	El reentrenamiento periódico de los modelos con nuevos datos garantiza que sigan siendo precisos y pertinentes.
Control del rendimiento	Establecer sistemas de supervisión continua para seguir el rendimiento de los modelos en tiempo real permite detectar a tiempo las desviaciones. Pueden controlarse parámetros como la exactitud, la precisión, la recuperación y la puntuación F1 (F1 score)
Control de versiones	Mantener el control de versiones de los modelos ayuda a seguir los cambios y a revertir a versiones anteriores si es necesario. Esto también ayuda a comprender la evolución del rendimiento del modelo.
Detección de anomalías:	La implantación de sistemas de detección de anomalías puede alertar a las partes interesadas de cambios inesperados en el comportamiento de los modelos, lo que permite intervenir a tiempo.

6.3 Sobreajuste

El sobreajuste se produce cuando un modelo aprende el ruido de los datos de entrenamiento en lugar de la señal, lo que da lugar a una generalización deficiente con datos no observados.[37] Esto hace que los modelos funcionen muy bien durante la fase de entrenamiento, pero muy mal cuando se prueban con datos ajenos al conjunto de datos de entrenamiento. Para evitar el sobreajuste, pueden aplicarse las estrategias del Cuadro 10.

Cuadro 10: Estrategias para reducir el sobreajuste

Estrategia	Detalle
Validación cruzada	Hay técnicas y buenas prácticas que pueden utilizarse, como la validación cruzada, que es una técnica para evaluar la precisión de un modelo entrenándolo y probándolo en diferentes subconjuntos de datos para garantizar que funciona bien en todos ellos. Un ejemplo es la validación cruzada k-fold.
Técnicas de regularización	Se trata de técnicas que penalizan los modelos por ser demasiado complejos, fomentando la simplicidad y la generalización. Algunos ejemplos de métodos de regularización son L1 (Lasso) y L2 (Ridge).
Poda y detención precoz	Podar los árboles de decisión (un tipo de modelo de ML) y detener el proceso de entrenamiento antes de tiempo cuando el rendimiento en los datos de validación empieza a degradarse son métodos eficaces para combatir el sobreajuste.
Métodos de ensamblaje	Los métodos de ensamblaje consisten en utilizar varias técnicas a la vez para mejorar la generalización y reducir la probabilidad de sobreajuste.

6.3.1 Otros obstáculos a tener en cuenta

Además de los retos técnicos, como la minimización de los sesgos, la gestión de las desviaciones y la reducción de los sobreajustes, la integración de la IA en la práctica se enfrenta a importantes obstáculos administrativos y humanos. Entre ellos se encuentran la navegación por flujos de trabajo clínicos complejos y la formación adecuada de los farmacéuticos en las nuevas herramientas y procesos. El éxito de la implantación también depende de la conformidad de los usuarios finales y de su voluntad de adoptar la nueva tecnología. En el cuadro 11 se describen los principales obstáculos que deben superarse para que la integración de la IA tenga éxito.

Cuadro 11: Retos adicionales para la aplicación

Barreras	Detalle
Integración con flujos de trabajo complejos	La integración de la IA en flujos de trabajo farmacéuticos y clínicos complejos implica varias consideraciones. Se trata de un reto importante para la adopción y aplicación de la IA[38].
Interoperabilidad	Es crucial garantizar que los sistemas de IA puedan integrarse sin problemas con la infraestructura informática y los sistemas de software existentes. Esto implica utilizar protocolos y API (Interfaz de programación de aplicaciones) estándar para la comunicación. También es crucial contar con los recursos necesarios (tanto humanos como económicos).
Formación y asistencia al usuario	Ofrecer una formación y una asistencia completa a los usuarios es esencial para facilitar una adopción sin problemas. Esto incluye crear interfaces fáciles de usar y ofrecer asistencia técnica continua.
Personalización del flujo de trabajo	Las soluciones de IA deben adaptarse a las necesidades y procesos específicos de la farmacia, la empresa farmacéutica, la organización o las personas que las utilizan. La personalización garantiza que las herramientas de IA complementen los flujos de trabajo existentes en lugar de perturbarlos.
Escalabilidad	Los sistemas de IA deben diseñarse para adaptarse al crecimiento de la organización. Esto implica planificar el aumento del volumen de datos, las funcionalidades adicionales y la ampliación de las bases de usuarios.
Colaboración y comunicación	Promover la colaboración entre expertos en IA, farmacéuticos y otras partes interesadas es vital. Unos canales de comunicación claros ayudan a garantizar que las soluciones de IA aborden los retos del mundo real con eficacia. Para integrar ambos mundos se necesitan tanto conocimientos técnicos como industriales. La incorporación de la IA a una empresa debe contar con el apoyo de toda una cultura que supere las barreras y las complicaciones.
Costos iniciales elevados	El desarrollo y la implantación de soluciones de IA requieren una inversión sustancial en infraestructura, personal cualificado y tecnología, lo que puede ser un factor disuasorio para muchas organizaciones.[32, 38]
Gestión del cambio	Integrar la IA en los flujos de trabajo existentes exige un cambio cultural en las organizaciones. La resistencia al cambio entre el personal y las partes interesadas puede obstaculizar el proceso de adopción.[32, 38]

6.4 Seleccionar la herramienta adecuada estableciendo las ventajas y las desventajas.

Seleccionar la herramienta adecuada no siempre es fácil y a menudo más de una solución de IA podría ser eficaz. No siempre son las capacidades funcionales de la herramienta las que impulsan la decisión, sino otros factores que entran en juego, como el entorno técnico en el que se implantará, el costo, construir o mantener el modelo, el nivel de transparencia necesario, la normativa o el acceso a los datos, entre otros.

A la hora de seleccionar la herramienta adecuada, hay que tener en cuenta sus capacidades y limitaciones, pero también el diseño de la aplicación. Por ejemplo, crear una herramienta de IA generativa para proporcionar información sanitaria a los pacientes conlleva un mayor riesgo cuando la herramienta está orientada al paciente. Sin embargo, si el modelo proporciona información a un clínico que primero la revisa antes de proporcionársela a un paciente, el riesgo es mucho menor. En el segundo caso, las opciones en cuanto a las herramientas que pueden utilizarse son mucho más amplias, ya que puede que no sea necesario disponer inmediatamente de todas las barreras de seguridad que se necesitarían en el primer caso.

Del mismo modo, la creación de una herramienta de apoyo a la toma de decisiones clínicas para la dosificación de fármacos en un entorno de cuidados críticos implica consideraciones diferentes. Crear un modelo que recomiende qué dosis debe recibir un paciente cada día es muy diferente a crear un modelo que alerte a un clínico de que el estado del paciente está cambiando (o se prevé que cambie pronto) y le pida que evalúe qué ajuste de dosis es necesario. Ambas situaciones requerirían una formación diferente para construir el modelo porque el resultado de cada modelo sería diferente. Para el primer escenario, el resultado del modelo es una dosis recomendada (por ejemplo, 5 mg). En el segundo escenario, el resultado del modelo predice si será necesario o no un ajuste de la dosis, lo que implica menos riesgo porque el médico sigue tomando la decisión, pero se ve reforzado por la IA. Para el primer escenario, un modelo de aprendizaje por refuerzo puede ser más apropiado, mientras que un modelo de aprendizaje supervisado puede ser más apropiado para el segundo.

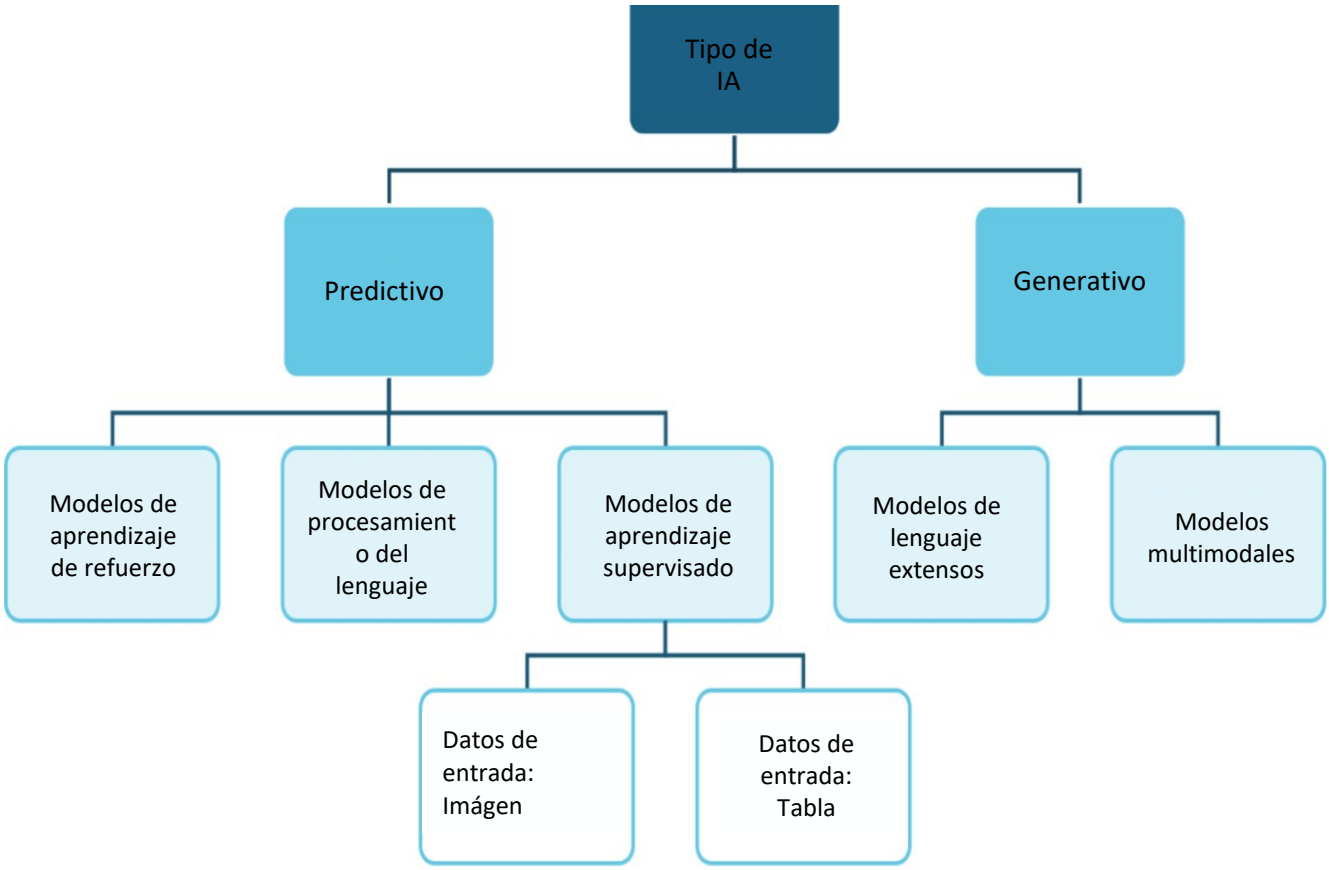
6.4.1 Sopesar las ventajas y desventajas entre modelos

Cuando se utilizan herramientas de IA de aprendizaje profundo, hay muchas ventajas, pero también limitaciones generales a tener en cuenta, como la dependencia de los datos y la interpretabilidad limitada del modelo (véase el cuadro 12). Para profundizar, es esencial sopesar los puntos fuertes y débiles de los distintos modelos de IA en función de su tipo, estilo de aprendizaje y datos de entrada. La Figura 2 describe los modelos de IA más utilizados, mientras que el cuadro 13 ofrece un resumen de alto nivel de las ventajas y desventajas que hay que tener en cuenta para cada tipo de modelo, en función de cómo se implementen y del objetivo final de su despliegue.

Cuadro 12: Ventajas y desventajas generales de los modelos de aprendizaje profundo

Ventajas	Desventajas
Adaptabilidad: Se puede mejorar el rendimiento del modelo a medida que se cuenta con más datos.	Dependencia de los datos: Requiere grandes cantidades de datos para entrenar efectiva y precisamente.
Versatilidad: Aplicable a una amplia gama de problemas, desde el reconocimiento de imágenes al procesamiento del lenguaje natural.	Opacidad: Los modelos complejos como las redes neuronales profundas pueden ser "cajas negras" que dificultan la comprensión de cómo se toman las decisiones.

Figura 2: Modelos de aprendizaje profundo



Cuadro 13: Resumen de alto nivel de las ventajas y desventajas tener en cuenta para los distintos modelos

Modelo	Ejemplos (no exhaustivos)	Ventajas	Desventajas	Consideraciones
Herramientas de procesamiento generativo del lenguaje natural (grandes modelos lingüísticos)	**Chatbot** • Programa informático que procesa texto natural y simula una conversación con un ser humano.[39] **Creación de contenidos** • Genera un texto original a partir de una instrucción **Clasificación de documentos** • Etiqueta un documento en función del tema u otra característica **Resumen** • Genera un resumen de un documento o consulta	• No requiere que los datos de entrada estén estructurados • Versátil • Un mismo modelo puede utilizarse para muchos propósitos. • Muchos productos y recursos no requieren codificación y son fáciles de usar. • Para usos generales, muchos modelos están listos para su uso "preconfigurados" (es decir, no requieren formación adicional ni ajustes finos). • Los modelos pueden conectarse a una fuente de datos externa	• No está entrenado para ser preciso. Propenso a las alucinaciones. • Limitaciones contextuales: comprensión del contexto, entendimiento del sarcasmo y el lenguaje complejo. • Sesgo: puede aprender y propagar inadvertidamente sesgos presentes en los datos de entrenamiento. • Puede tener dificultades para realizar tareas específicas • El modelo no incluye ninguna información (por ejemplo, acontecimientos actuales, nuevos fármacos, actualizaciones de directrices) que se haya producido después de la fecha de corte del conocimiento de los datos de entrenamiento.	• Por lo general, no resulta práctico desarrollarlo internamente. A menudo hay que utilizar un modelo preexistente. • Algunos modelos son patentados, mientras que otros son de uso público. • El rendimiento varía según los modelos. Algunos son más precisos en determinadas competencias que otros (conocimientos médicos, matemáticas, razonamiento, información sobre medicamentos, etc.). • Es importante revisar el rendimiento de un modelo desde varios puntos de vista, cómo se ha entrenado, el límite de sus datos de entrenamiento y las fuentes/riesgos de sesgo. Sin embargo, no todos los modelos comparten estos datos. • Requiere el uso de técnicas como la **Generación de recuperación aumentada** (RAG de sus siglas en inglés) para que el modelo proporcione información en tiempo real o responda a preguntas sobre un documento específico o una base de conocimientos interna.

Modelo	Ejemplos (no exhaustivos)	Ventajas	Desventajas	Consideraciones
Herramientas no generativas de procesamiento del lenguaje natural	**Clasificación:** ● Etiquetar el tema de un documento **Extracción de datos:** ● Extraer un diagnóstico o una lista de medicamentos de una nota clínica **Predictivo:** ● Uso de notas clínicas hospitalarias para estimar el riesgo de que un paciente necesite cuidados al final de su vida	● Los desarrolladores tienen más control sobre los resultados del modelo ● Se puede construir con menos datos ● No requiere tanta potencia de cálculo (puede ejecutarse más fácilmente en dispositivos individuales). ● Eficaz para casos de uso limitados	● Tarda más en construirse; requiere conocimientos de IA ● No son generalizables ni específicos. Los modelos creados para un uso no pueden utilizarse para otros fines sin una nueva formación. ● Suele requerir datos etiquetados para el entrenamiento, cuya elaboración puede requerir muchos recursos.	● Puede construirse internamente ● El uso de valores previamente entrenados puede ayudar a mejorar la precisión ● Dependiendo del modelo subyacente y del uso o no de redes neuronales, estos modelos pueden ser más comprensibles que otros modelos de caja negra que utilizan redes neuronales
Modelos de IA generativa multimodal	**Chatbots:** ● Programa informático que simula una conversación con un ser humano mediante el tratamiento de texto, imágenes, audio y/o vídeo. **Asistentes de IA:** ● Varios modelos se coordinan entre sí y/o con otras herramientas para completar una tarea	● Puede interactuar con los modelos a través de múltiples canales ● Los modelos pueden diseñarse para permitir preguntas de seguimiento y un diálogo de ida y vuelta (por ejemplo, después de pedir a un modelo que mire una imagen, el usuario puede hacer preguntas de seguimiento basadas en la descripción/respuesta del modelo).	● Las mismas limitaciones de la IA generativa se aplican a los modelos multimodales	● No todos los modelos son realmente multimodales. Algunos parecen ser multimodales al combinar varios modelos diferentes (por ejemplo, generador de texto, generador de imágenes, modelo de transcripción, etc.). ● Las consideraciones son similares a las de las herramientas de procesamiento generativo del lenguaje natural anteriores

Modelo	Ejemplos (no exhaustivos)	Ventajas	Desventajas	Consideraciones
Modelo de aprendizaje supervisado - reconocimiento de imágenes	**Diagnóstico:** • Diagnóstico o caracterización de enfermedades a partir de imágenes médicas. [40-42]	• Muchas soluciones de IA han demostrado suficiente precisión, seguridad y fiabilidad según lo determinado por organismos reguladores, como la Administración de Alimentos y Medicamentos (FDA).[43.]	• Los modelos son cajas negras y no se pueden explicar • A menudo sólo puede utilizarse con imágenes del mismo tamaño, tipo y resolución con las que se entrenó. • Puede estar vinculado a un único generador de imágenes específico: si se cambia de aparato, será necesario un nuevo modelo.	• Si utiliza un modelo de aprendizaje profundo para imágenes médicas de un proveedor externo, compruebe si el modelo ha sido aprobado por el organismo regulador de medicamentos o dispositivos correspondiente (por ejemplo, FDA, EE. UU.; Health Sciences Authority, Singapur; South African Health Products Regulatory Authority (SAHPRA), Sudáfrica; INFARMED, Portugal; Medicines and Healthcare Products Regulatory Agency (MHRA), Reino Unido de Gran Bretaña e Irlanda del Norte).[44]
Modelo de aprendizaje supervisado - datos tabulares	**Predicción de riesgos:** • Creación de una puntuación de riesgo para la probabilidad de que un paciente vuelva a ingresar en el hospital tras el alta.[45] • Predecir la escasez de medicamentos.[46]	• Numerosas herramientas de aplicación existentes • Puede alcanzar una gran precisión y rendimiento • Ya existe una cantidad significativa de datos sanitarios en formato tabular • Puede ajustarse en función de los datos locales • Puede utilizarse para asignar datos existentes a una nueva métrica o puntuación de riesgo (por ejemplo, asignar datos de podómetro a una puntuación de estabilidad al caminar).	• Los modelos son cajas negras y no se pueden explicar • Puede que no siempre funcionen mejor que los métodos estadísticos tradicionales (por ejemplo, la regresión logística) • Requiere grandes cantidades de datos para obtener mejores resultados que con métodos estadísticos más sencillos. • Perpetúa errores y sesgos preexistentes en los datos • No generalizable	• No es ideal para los sistemas de "recomendación", ya que no predice realmente la mejor acción, sino las acciones que se han llevado a cabo históricamente (que no siempre han sido correctas). • Su ampliación puede resultar difícil y requerir muchos recursos. Cada nuevo caso de uso suele requerir un nuevo modelo y datos de entrenamiento

Modelo	Ejemplos (no exhaustivos)	Ventajas	Desventajas	Consideraciones
Modelo de aprendizaje por refuerzo	**Sistemas de recomendación:** ● Optimiza una secuencia de acciones/decisiones para alcanzar un objetivo a largo plazo.	● Pueden aprender sin supervisión explícita, lo que los hace adecuados para tareas en las que los datos anotados son escasos o no están disponibles. ● Diseñado para maximizar una recompensa que podría estar lejos en el futuro (por ejemplo, qué serie de acciones deben realizarse en un juego para ganar). ● Destaca en problemas que requieren una secuencia de acciones para alcanzar un objetivo	● Estos modelos pueden tener a menudo consecuencias imprevistas si la función de recompensa no está bien definida ● Caja negra ● El modelo puede tomar decisiones impredecibles para "explorar" otras vías de actuación, lo que podría resultar inseguro en un entorno sanitario ● Es difícil implantar barreras de seguridad en el modelo para evitar que haga recomendaciones inseguras	● Se adapta bien a un entorno sanitario en el que la "recompensa" puede ser un resultado a largo plazo, más que un beneficio a corto plazo. Por ejemplo, durante una hospitalización prolongada se toman múltiples decisiones terapéuticas con el objetivo de aumentar las probabilidades de que el paciente sobreviva hasta el alta (objetivo a largo plazo). ● Estos modelos deben integrarse con sistemas basados en el conocimiento para crear barandillas de seguridad ● Es necesaria una validación rigurosa para garantizar que se eligió la "recompensa" adecuada y que el modelo no está tomando accidentalmente decisiones incorrectas.

7 Lista de comprobación para la aplicación de IA

A modo de resumen de la información contenida en el kit de herramientas de IA, la siguiente lista de comprobación sirve como guía de preguntas para orientar cómo se debe implantar una herramienta o solución de IA. No pretende ser una guía exhaustiva, sino un punto de partida. Cada entorno de trabajo incluirá varios requisitos únicos que deberán tenerse en cuenta.

Lista de comprobación	
Definir el caso de uso	
☐	¿Qué problema pretende resolver la herramienta de IA?
☐	¿Puede resolverse el problema con una solución o herramienta no basada en la inteligencia artificial?
☐	¿Se ha contado con todas las partes interesadas, incluidos los responsables de la toma de decisiones y los usuarios finales?
☐	¿Cómo encajará esta herramienta en el flujo de trabajo existente o cómo se ajustará el flujo de trabajo en función de la herramienta?
Selección del modelo	
☐	¿El modelo de IA se desarrollará internamente o se recurrirá a un proveedor externo?
☐	Si el modelo se desarrolla internamente, ¿existen suficientes datos para su entrenamiento?
☐	Si se utiliza un modelo básico de terceros, ¿requerirá un ajuste fino (entrenamiento adicional) con datos locales o específicos del dominio?
☐	Si se utiliza un producto o software de terceros basado en IA, ¿se necesita evaluar y aprobar su seguridad y eficacia por un organismo regulador (por ejemplo, un software autorizado por la FDA como dispositivo médico)? En caso afirmativo, ¿lo ha sido?
☐	Si se utiliza un modelo existente, ¿cuál es su rendimiento (véase el Cuadro 3)? ¿Cómo se compara con otros modelos o puntos de referencia existentes?
Conformidad	
☐	¿Tendrá el modelo acceso a información sanitaria protegida o la utilizará?
☐	¿El uso del modelo requerirá que los datos se compartan fuera de la organización? Por ejemplo, ¿el modelo requiere el uso de una Interfaz de programación de aplicaciones (API) o los datos se comparten con un servidor en la nube externo a la organización? En caso afirmativo, ¿qué limitaciones plantea esto respecto a los datos que pueden incluirse en la entrada del modelo?
☐	¿Qué normativa de cumplimiento debe seguirse (véase el apartado sobre Cumplimiento y normativa ISO)?
	¿Puede implantarse el modelo a escala local?
Selección de proveedores	
☐	Si se recurre a un proveedor externo, ¿proporciona éste una ficha técnica modelo (como la descrita en el Cuadro 2) o detalles sobre sus datos de entrenamiento y el rendimiento del modelo (véanse las métricas en el Cuadro 3)?
☐	¿Con qué frecuencia audita el proveedor el rendimiento del modelo o vuelve a formarlo?
☐	¿Proporciona el proveedor métricas de rendimiento actualizadas después de implantar cualquier actualización del modelo?
Seguridad	
☐	¿Se han esbozado los posibles fallos del modelo? ¿Cuáles serán las estrategias de mitigación?
☐	¿Cómo se auditará el modelo de forma continua?
☐	¿Con qué frecuencia hay que revalidar el modelo en función de las posibles desviaciones?
☐	Basándose en los datos de entrenamiento del modelo, ¿es éste menos preciso para subpoblaciones específicas? ¿Cómo será esto mitigado?

8 Formación y desarrollo de competencias en IA

8.1 ¿Qué deben saber los farmacéuticos sobre la IA?

Para los farmacéuticos y los equipos de farmacia, tanto en entornos hospitalarios como comunitarios, las siguientes competencias son esenciales para aprovechar todo el potencial de estas tecnologías:

- **Comprensión de las capacidades y limitaciones de la IA generativa:** Los farmacéuticos deben comprender lo que las herramientas de IA generativa pueden y no pueden hacer, incluido su alcance, fiabilidad y los contextos en los que funcionan de forma óptima. Este conocimiento garantiza una confianza adecuada en la IA para la toma de decisiones, evitando el exceso de confianza que podría conducir a errores.

- **Conocimientos de datos:** La capacidad de interpretar y evaluar los datos generados por las herramientas de IA es crucial. Los farmacéuticos deben saber leer, analizar y tomar decisiones informadas basadas en los datos generados por la IA. Este aspecto es esencial para que exista una atención al paciente y una gestión de la medicación precisas.

- **Consideraciones éticas y legales del uso de la IA:** Comprender las implicaciones éticas y los límites legales del uso de la IA en la práctica farmacéutica es primordial. Esto incluye las medidas sobre la privacidad del paciente, la seguridad de los datos y el uso ético de la IA en los procesos de toma de decisiones para garantizar la seguridad del paciente y el cumplimiento de la normativa.

- **Pensamiento crítico y toma de decisiones:** Aunque la IA puede proporcionar recomendaciones, la responsabilidad última de la toma de decisiones recae en el farmacéutico. La capacidad de evaluar críticamente los consejos generados por la IA, teniendo en cuenta los contextos y necesidades únicos de cada paciente, es esencial para una práctica farmacéutica eficaz.

- **Capacidad de comunicación:** Los farmacéuticos deben comunicar eficazmente la información generada por la IA a los pacientes y a los equipos sanitarios. Esto incluye traducir datos complejos de la IA en consejos comprensibles y garantizar que las decisiones respaldadas por la IA sean transparentes y justificables.

- **Aprendizaje continuo y adaptabilidad:** El campo de la IA evoluciona rápidamente; por lo tanto, los farmacéuticos deben comprometerse con la formación continua y la adaptación a las nuevas tecnologías. Este aprendizaje continuo garantiza que la práctica farmacéutica se mantenga a la vanguardia, utilizando las herramientas de IA más actuales para mejorar la atención al paciente.

- **Habilidades de colaboración para equipos interdisciplinarios:** Trabajar con equipos interdisciplinarios, incluidos profesionales de IT, científicos de datos y otros proveedores de atención médica, es crucial para implementar y optimizar las herramientas de IA en la práctica farmacéutica. Una colaboración eficaz garantiza que las implementaciones de IA estén bien coordinadas y satisfagan las diversas necesidades de la prestación de asistencia sanitaria.

- **Atención centrada en el paciente:** Los farmacéuticos deben garantizar que las herramientas de IA se utilicen de forma que den prioridad a las necesidades y los resultados de los pacientes. Esto implica utilizar la IA para personalizar la gestión de la medicación y el apoyo, mejorando la calidad de la atención prestada a los pacientes.

- **Innovación y creatividad:** Por último, como en todos los ecosistemas que trabajan con IA, los farmacéuticos deben cultivar una mentalidad innovadora, buscando formas creativas de aplicar la IA en la práctica farmacéutica. Esto incluye el desarrollo de nuevos flujos de trabajo, estrategias de atención al paciente y prácticas de gestión mejoradas por la IA, impulsando el campo de la farmacia.

Cada una de estas competencias es fundamental para que los farmacéuticos y los equipos de farmacia integren eficazmente las herramientas de IA generativa en su práctica. Juntas, permiten ofrecer una atención al paciente de alta calidad, eficiente y personalizada, garantizando que los profesionales farmacéuticos se mantengan a la vanguardia de la innovación sanitaria.

8.2 Conocimientos y competencias para la IA

La integración de herramientas de IA generativa en la práctica farmacéutica representa un cambio transformador hacia una prestación sanitaria más eficiente y eficaz. En la práctica farmacéutica, los sistemas de IA se utilizan cada vez más para optimizar la gestión de la medicación, mejorar la prescripción de medicamentos a través de las recetas y agilizar el control del inventario de la farmacia. Por ejemplo, los sistemas de verificación de recetas basados en IA pueden detectar automáticamente posibles errores en la dispensación de medicación, reduciendo la carga de las tareas de verificación manual para los farmacéuticos. Aunque esto mejora la eficiencia y la seguridad de los pacientes, también afecta al papel tradicional de los farmacéuticos en la supervisión de los procesos de dispensación de medicamentos y la prestación de atención clínica. Los farmacéuticos tienen que adaptarse a trabajar con sistemas de IA, lo que requiere nuevas competencias para gestionar e interpretar los datos generados por la IA.

Los farmacéuticos que quieran aprender más sobre la IA y sus aplicaciones en la atención sanitaria y la práctica farmacéutica tienen una gran variedad de entornos de aprendizaje a su disposición. Los cursos en línea y los seminarios web ofrecidos por instituciones como Coursera, edX, Udemy y universidades proporcionan conocimientos básicos sobre la IA, sus principios y sus aplicaciones sanitarias, y son fácilmente accesibles.

En Asia, Coursera China ofrece cursos a medida en colaboración con las mejores universidades chinas, centrados en la IA en la atención sanitaria con contenidos disponibles en mandarín. NUS-ISS (Instituto de Ciencia de Sistemas de la Universidad Nacional de Singapur) ofrece programas especializados sobre IA en sanidad, dirigidos a profesionales de Singapur y regiones circundantes. K-MOOC (Korean Massive Open Online Course), apoyado por el Ministerio de Educación de Corea del Sur, ofrece cursos relacionados con la IA centrados en aplicaciones sanitarias y farmacéuticas, disponibles en coreano. Además, la Universidad Tecnológica de Nanyang de Singapur ofrece formación ejecutiva y módulos de aprendizaje en línea sobre aplicaciones de la IA en la sanidad, dirigidos a profesionales sanitarios de toda Asia. Estas plataformas proporcionan a los farmacéuticos conocimientos y habilidades específicas para integrar eficazmente la IA en su práctica.

Las organizaciones profesionales, como la Sociedad Americana de Farmacéuticos del Sistema de Salud (ASHP) y la Federación Internacional Farmacéutica (FIP), ofrecen sesiones de formación especializadas, talleres y conferencias centradas en la IA en la práctica farmacéutica. Las revistas y publicaciones académicas suelen incluir artículos sobre investigación y estudios de casos de IA, lo que mantiene a los farmacéuticos al día de los últimos avances y aplicaciones prácticas.

La colaboración con equipos multidisciplinares, incluidos científicos de datos y profesionales sanitarios a través de seminarios y proyectos prácticos, puede aportar conocimientos prácticos y experiencia del mundo real. Además, la participación en foros y grupos de debate centrados en la IA, tanto en línea como dentro de redes profesionales, fomenta un entorno de aprendizaje continuo e intercambio de conocimientos.

Estos recursos educativos pueden ayudar a dotar a los farmacéuticos con las habilidades y la comprensión necesarias para integrar eficazmente la IA en su práctica, para mejorar la atención al paciente y optimizar la prestación de asistencia sanitaria.

9 Referencias

1. Templin T, Perez MW, Sylvia S, Leek J, Sinnott-Armstrong N. Addressing 6 challenges in generative AI for digital health: A scoping review. PLOS Digital Health. 2024;3(5):e0000503. Disponible en: https://pubmed.ncbi.nlm.nih.gov/38781686.

2. Jiang T, Gradus JL, Rosellini AJ. Aprendizaje automático supervisado: A Brief Primer. Behav Ther. 2020;51(5):675-87. Disponible en: https://pubmed.ncbi.nlm.nih.gov/32800297.

3. Stokes JM, Yang K, Swanson K, Jin W, Cubillos-Ruiz A, Donghia NM, et al. A Deep Learning Approach to Antibiotic Discovery. Cell. 2020;180(4):688-702.e13. Disponible en: https://pubmed.ncbi.nlm.nih.gov/32084340.

4. Arulkumaran K, Deisenroth MP, Brundage M, Bharath AA. Aprendizaje profundo por refuerzo: A Brief Survey. Revista de procesamiento de señales del IEEE. 2017;34(6):26-38. Disponible en: https://ieeexplore.ieee.org/document/8103164

5. Silver D, Hubert T, Schrittwieser J, Antonoglou I, Lai M, Guez A, et al. Mastering chess and shogi by self-play with a general reinforcement learning algorithm. arXiv preprint arXiv:171201815. 2017. Disponible en: https://arxiv.org/abs/1712.01815.

6. Heo J, Kang Y, Lee S, Jeong D-H, Kim K-M. An Accurate Deep Learning-Based System for Automatic Pill Identification: Model Development and Validation. J Med Internet Res. 2023;25:e41043. Disponible en: https://pubmed.ncbi.nlm.nih.gov/36637893.

7. Mitchell M, Wu S, Zaldivar A, Barnes P, Vasserman L, Hutchinson B, et al., editores. Model Cards for Model Reporting2019: ACM. Available at: https://dl.acm.org/doi/10.1145/3287560.3287596.

8. Fang H, Miao H, Shukla K, Nanas D, Xu C, Greer C, et al. Introducing the model card toolkit for easier model transparency reporting. Blog de Google AI. 2020. Disponible en: https://research.google/blog/introducing-the-model- card-toolkit-for-easier-model-transparency-reporting.

9. Jelinek F, Mercer RL, Bahl LR, Baker JK. Perplejidad: una medida de la dificultad de las tareas de reconocimiento del habla. The Journal of the Acoustical Society of America. 1977;62(S1):S63-S. Disponible en: https://pubs.aip.org/asa/jasa/article/62/S1/S63/642598/Perplexity-a-measure-of-the-difficulty-of-speech.

10. Lin C-Y, Och FJ, editores. Orange: un método para evaluar métricas de evaluación automática para la traducción automática. COLING 2004: Actas de la 20ª Conferencia Internacional de Lingüística Computacional; 2004. Disponible en: https://aclanthology.org/C04-1072.

11. Papineni K, Roukos S, Ward T, Zhu W-J, editores. Bleu: un método para la evaluación automática de la traducción automática. Actas de la 40ª reunión anual de la Association for Computational Linguistics; 2002. Disponible en: https://aclanthology.org/P02-1040.

12. Wang A. Glue: A multi-task benchmark and analysis platform for natural language understanding. arXiv preprint arXiv:180407461. 2018. Disponible en: https://arxiv.org/abs/1804.07461.

13. Lin C-Y, editor ROUGE; Un paquete para la evaluación automática de resúmenes. Text summarization branches out; 2004; Barcelona, España: Asociación de Lingüística Computacional. Disponible en: https://aclanthology.org/W04-1013.

14. Hossin M, Sulaiman MN. A review on evaluation metrics for data classification evaluations. Revista internacional de minería de datos & proceso de gestión del conocimiento. 2015;5(2):1. Disponible en: https://www.aircconline.com/ijdkp/V5N2/5215ijdkp01.pdf.

15. Grupo de Expertos de Alto Nivel en Inteligencia Artificial. Directrices éticas para una IA digna de confianza2019. Disponible en: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

16. Harrer S. Attention is not all you need: the complicated case of ethically using large language models in healthcare and medicine. eBioMedicine. 2023;90:104512. Disponible en: https://pubmed.ncbi.nlm.nih.gov/36924620.

17. Morley J, Machado CC, Burr C, Cowls J, Joshi I, Taddeo M, et al. The ethics of AI in health care: a mapping review. Social Science & Medicine. 2020;260:113172. Disponible en: https://pubmed.ncbi.nlm.nih.gov/32702587.

18. Mccradden M, Odusi O, Joshi S, Akrout I, Ndlovu K, Glocker B, et al. ¿Lo que es justo es... justo? Presentando JustEFAB, un marco ético para operacionalizar la ética médica y la justicia social en la integración del aprendizaje automático clínico: JustEFAB. Actas de la 2023 ACM Conference on Fairness, Accountability, and Transparency;

Chicago, IL, USA: Asociación para Informática Machinery; 2023. p. 1505-19. Disponible en at: https://dl.acm.org/doi/10.1145/3593013.3594096.

19. Beauchamp TL, Childress JF. Principios de ética biomédica. 7th ed. Nueva York: Oxford University Press; 2013.

20. Orientaciones de la OMS. Ética y gobernanza de la inteligencia artificial para la salud. Organización Mundial de la Salud. 2021. Disponible en: https://www.who.int/publications/i/item/9789240029200.

21. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Inteligencia de máquinas de la naturaleza. 2019;1(9):389-99. Disponible en: https://www.nature.com/articles/s42256-019-0088-2.

22. Abby Sen A, Joy J, Jennex M. Illuminating Ethical Dilemmas in AI-Driven Regulation of Healthcare Marketing on Social Media. 2025. Disponible en: https://hdl.handle.net/10125/109472.

23. Leben D. Cuestiones éticas de la IA en medicina. Salud digital. 2025:309-19.

24. Hasan HE, Jaber D, Khabour OF, Alzoubi KH. Ethical considerations and concerns in the implementation of AI in pharmacy practice: a cross-sectional study. BMC Medical Ethics. 2024;25(1):55. Disponible en: https://bmcmedethics.biomedcentral.com/articles/10.1186/s12910-024-01062-8.

25. Hefti E, Xie Y, Engelen K. Machine learning model better identifies patients for pharmacist intervention to reduce hospitalization risk in a large outpatient population. Journal of Medical Artificial Intelligence. 2025;8. Disponible en: https://jmai.amegroups.org/article/view/9455/html.

26. Ley EAI. Ley de Inteligencia Artificial de la UE. Recuperado en mayo; 2024. Disponible en: https://artificialintelligenceact.eu.

27. Ley de Portabilidad y Responsabilidad de los Seguros Sanitarios de 1996, 104 (1996). Disponible en: https://aspe.hhs.gov/reports/health-insurance-portability-accountability-act-1996.

28. Reglamento (UE) 2016/679, 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 y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos), (2023). Disponible en: https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng.

29. Baird P, Cobbaert K. El software como producto sanitario. Comparación del enfoque de la UE con el de EE.UU. Available at: https://pages.bsigroup.com/l/35972/2023-12-11/3t7658k.

30. Hermon R. Software as a Medical Device (SaMD): ¿Término útil o inútil? 2021. Disponible en: https://researchnow.flinders.edu.au/en/publications/software-as-a-medical-device-samd-useful-or-useless- término.

31. HL-7. HL7 2023 [Disponible en: https://www.hl7.org/.

32. Meskó B, Topol EJ. The imperative for regulatory oversight of large language models (or generative AI) in healthcare. npj Digital Medicine. 2023;6(1):120. Disponible en: https://www.nature.com/articles/s41746-023- 00873-0.

33. Levi R, Gorenstein D. La IA en medicina debe desplegarse con cuidado para contrarrestar los prejuicios y no afianzarlos. NPR Health Shots. 2023. Disponible en: https://www.npr.org/sections/health-shots/2023/06/06/1180314219/artificial-intelligence-racial-bias-health-care.

34. Kohsiyama A, Lomas E, Polle R, Almeida D, Kazim E, Inness C. Assessing and Mitigating Bias and Discrimination in AI: The Alan Turing Institute; 2024. Disponible en: https://www.turing.ac.uk/courses/assessing-and-mitigating-bias- and-discrimination-ai.

35. Castelnovo A, Crupi R, Greco G, Regoli D, Penco IG, Cosentini AC. Una aclaración de los matices en el panorama de las métricas de equidad. Scientific Reports. 2022;12(1):4209. Disponible en: https://www.nature.com/articles/s41598- 022-07939-1.

36. Holdsworth J, Belcic I, Stryker C. ¿Qué es la deriva del modelo? IBM; 2024. Disponible en: https://www.ibm.com/think/topics/model-drift.

37. AWS. ¿Qué es el sobreajuste? : AWS Amazon; 2024. Disponible en: https://aws.amazon.com/what-is/overfitting/.

38. Ahmed MI, Spooner B, Isherwood J, Lane M, Orrock E, Dennison A. A Systematic Review of the Barriers to the Implementation of Artificial Intelligence in Healthcare. Cureus. 2023;15(10):e46454. Disponible en: https://pmc.ncbi.nlm.nih.gov/articles/PMC10623210.

39. Adamopoulou E, Moussiades L. An Overview of Chatbot Technology. Artificial Intelligence Applications and Innovations; Neos Marmaras, Greece: Springer, Cham; 2020. p. 373-83. Available at: https://link.springer.com/chapter/10.1007/978-3-030-49186-4_31.

40. Kumar Y, Koul A, Singla R, Ijaz MF. Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. J Ambient Intell Humaniz Comput. 2023;14(7):8459-86. Available at: https://link.springer.com/article/10.1007/s12652-021-03612-z.

41. Mello-Thoms C, Mello CAB. Aplicaciones clínicas de la inteligencia artificial en radiología. Br J Radiol. 2023;96(1150):20221031. Disponible en: https://pubmed.ncbi.nlm.nih.gov/37099398.

42. Shen J, Zhang CJP, Jiang B, Chen J, Song J, Liu Z, et al. Inteligencia artificial frente a clínicos en el diagnóstico de enfermedades: Systematic Systematic Review. JMIR Med Inform. 2019;7(3):e10010. Disponible en: https://medinform.jmir.org/2019/3/e10010.

43. Benjamens S, Dhunnoo P, Meskó B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. npj Digital Medicine. 2020;3(1):118. Disponible en: https://www.nature.com/articles/s41746-020-00324-0.

44. Organización Mundial de la Salud. Consideraciones reglamentarias sobre la inteligencia artificial para la salud: Organización Mundial de la Salud; 2023. Disponible en: https://www.who.int/publications/i/item/9789240078871.

45. Davis S, Zhang J, Lee I, Rezaei M, Greiner R, McAlister FA, et al. Effective hospital readmission prediction models using machine-learned features. BMC Health Services Research. 2022;22(1):1415. Disponible en: https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-022-08748-y.

46. Pall R, Gauthier Y, Auer S, Mowaswes W. Predicción de la escasez de medicamentos mediante datos de farmacia y aprendizaje automático. Salud Care Management Science. 2023;26(3):395-411. Disponible en at: https://ideas.repec.org/a/kap/hcarem/v26y2023i3d10.1007_s10729-022-09627-y.html.

2 Anexo - Glosario de términos utilizados en este kit de herramientas

Término	Definición
Algoritmo	Procedimiento secuencial o conjunto de reglas, diseñado para resolver un problema concreto o realizar una tarea específica. Es una secuencia clara y finita de instrucciones que toma una entrada, la procesa y genera una salida.
Interfaz de programación de aplicaciones (API)	Conjunto de directrices y protocolos que permiten a distintas aplicaciones informáticas comunicarse entre sí. Describe los métodos y formatos de datos que utilizan las aplicaciones para solicitar e intercambiar información, lo que permite una interacción fluida. Las API se utilizan con frecuencia para integrar sistemas, servicios o aplicaciones, ayudando en tareas como la recuperación de datos, la invocación de servicios y el intercambio de funcionalidades.
Inteligencia Artificial (IA)	Rama de la informática dedicada a crear sistemas capaces de realizar tareas que suelen requerir inteligencia humana, como el aprendizaje, el razonamiento, la resolución de problemas y el reconocimiento de patrones.
Inteligencia general artificial (AGI)	Tipo de IA que tiene la capacidad humana de comprender, aprender y aplicar conocimientos en una amplia gama de tareas, de forma similar a las capacidades cognitivas humanas. Las AGI pueden adaptarse a nuevos retos, razonar los problemas y transferir conocimientos de un ámbito a otro.
Inteligencia farmacológica artificial (API)	Utilización de técnicas avanzadas de inteligencia artificial para analizar e interpretar datos relacionados con la farmacología. Su objetivo es mejorar la comprensión de las interacciones entre fármacos, el descubrimiento de medicamentos y los efectos de las sustancias en los sistemas biológicos.
Grandes datos (Big data)	Una vasta colección de datos que son demasiado grandes para que los sistemas tradicionales de gestión de datos puedan procesarlos. La IA utiliza los big data para mejorar la eficiencia de los modelos en el aprendizaje y el análisis.
Ciencia de datos	Campo interdisciplinar centrado en la extracción de ideas y conocimientos a partir de datos mediante diversas técnicas, como el análisis estadístico, la minería de datos, el ML y la visualización de datos.
Aprendizaje profundo	Implica el uso de grandes redes neuronales artificiales multicapa que procesan los datos utilizando representaciones continuas (números reales). Ofrece una mejor generalización a partir de conjuntos de datos pequeños y se adapta más eficazmente a grandes conjuntos de datos y recursos informáticos.
Despliegue	Proceso de integración de un modelo de IA en un entorno de producción para realizar predicciones y análisis basados en datos.
IA Generativa (farmacia)	La IA generativa en farmacia se refiere a la aplicación de técnicas avanzadas de IA que pueden crear nuevos datos, soluciones o conocimientos basados en la información existente. Estas tecnologías aprovechan los modelos de ML, en particular los modelos generativos, para mejorar diversos aspectos de la práctica y la investigación farmacéuticas.
Grandes modelo lingüístico (LLM)	Los LLM son un tipo específico de modelo de IA generativa centrado en la producción de texto similar al humano. Mientras que la IA generativa se refiere a una amplia gama de técnicas y modelos de IA diseñados para crear nuevos contenidos -ya sea texto, imágenes, audio o vídeo-, los LLM se especializan en la generación de texto.
Aprendizaje automático (ML)	Rama de la IA que se centra en cómo los agentes informáticos pueden mejorar su percepción, conocimiento, razonamiento o acciones a través de la experiencia o los datos. El ML integra conceptos de campos como la informática, la estadística, la psicología, la neurociencia, la economía y la teoría de control.
Generación de lenguaje natural (NLG)	Rama de la inteligencia artificial dedicada al desarrollo de sistemas que crean automáticamente textos similares a los humanos a partir de datos estructurados. El NLG permite a los ordenadores generar contenidos escritos, como informes, resúmenes o respuestas, en lenguaje natural. Este proceso transforma datos, hechos o percepciones en frases o párrafos coherentes y significativos, mejorando la forma en que los humanos entienden la información e interactúan con ella. El lenguaje natural se aplica ampliamente en ámbitos como los chatbots, los informes automatizados, la creación de contenidos y la comunicación personalizada.

Procesamiento del lenguaje natural (PLN)	Campo de la inteligencia artificial cuyo objetivo es capacitar a los ordenadores para comprender, interpretar y generar lenguaje humano. Implica el desarrollo de algoritmos y modelos que permitan a las máquinas procesar y analizar datos de texto o voz de forma significativa y útil.
Red neuronal	Modelos informáticos inspirados en la estructura del cerebro humano. Estas neuronas artificiales interconectadas, organizadas en capas, aprenden de los datos para hacer predicciones en el ML, en el que se basa el aprendizaje profundo.
Prompt	La información que un usuario introduce en un modelo de IA para recibir una respuesta específica.
Generación de recuperación aumentada (RAG)	Técnica avanzada de procesamiento del lenguaje natural que combina la recuperación de información con la generación de texto.
Automatización robótica de procesos (RPA)	Tecnología rival que utiliza robots informáticos para automatizar tareas repetitivas basadas en reglas y entradas fijas, a menudo realizando tareas con más eficacia que los humanos.

International
Pharmaceutical
Federation

Fédération
Internationale
Pharmaceutique

Andries Bickerweg 5
2517 JP The Hague
The Netherlands
-
T +31 (0)70 302 19 70
F +31 (0)70 302 19 99
fip@fip.org
-
www.fip.org