



# Microcredencial de IA ética

GUÍA DE APOYO

UC1 | ¿Qué es el sesgo algorítmico?

# ¿Cómo utilizar esta guía de apoyo?

Este documento es interactivo. A lo largo del documento encontrará enlaces a información adicional.



Botón que le lleva al principio del documento. Este icono aparece en la esquina superior derecha de las páginas.



Siempre que vea esta flecha, significa que tiene un **texto interactivo en color** sobre el que hacer clic, que tiene asociado un enlace externo.

**DESCARGA DE RESPONSABILIDAD:** Tenga en cuenta que no podemos garantizar la disponibilidad permanente de contenidos externos, como vídeos, ya que pueden estar sujetos a cambios o ser retirados por sus autores o las plataformas que los albergan.

# Índice

Haga clic en el menú

**01. Introducción**

**02. Contenidos del curso y resultados esperados**

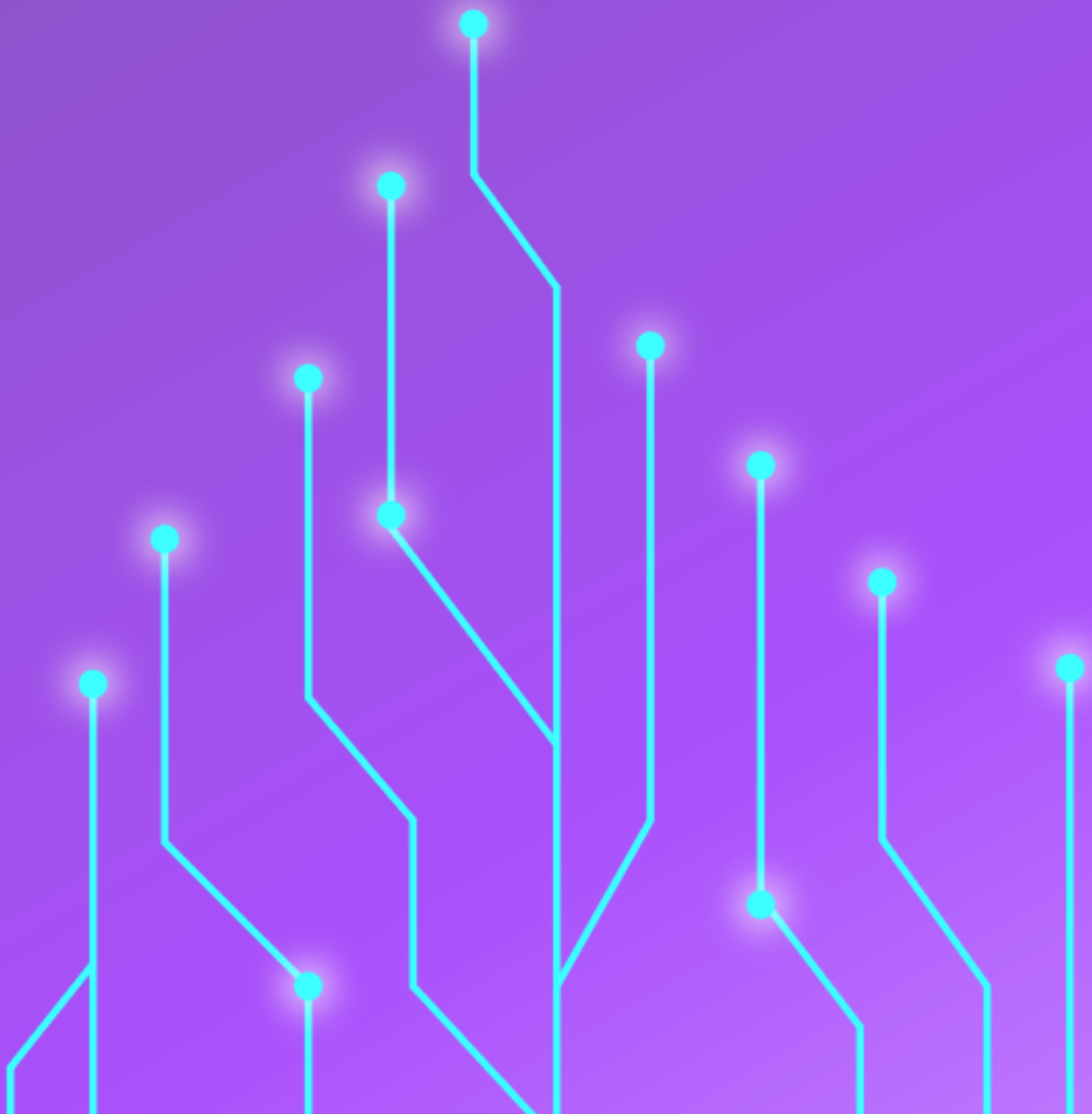
**03. ¿Qué es el sesgo algorítmico?**

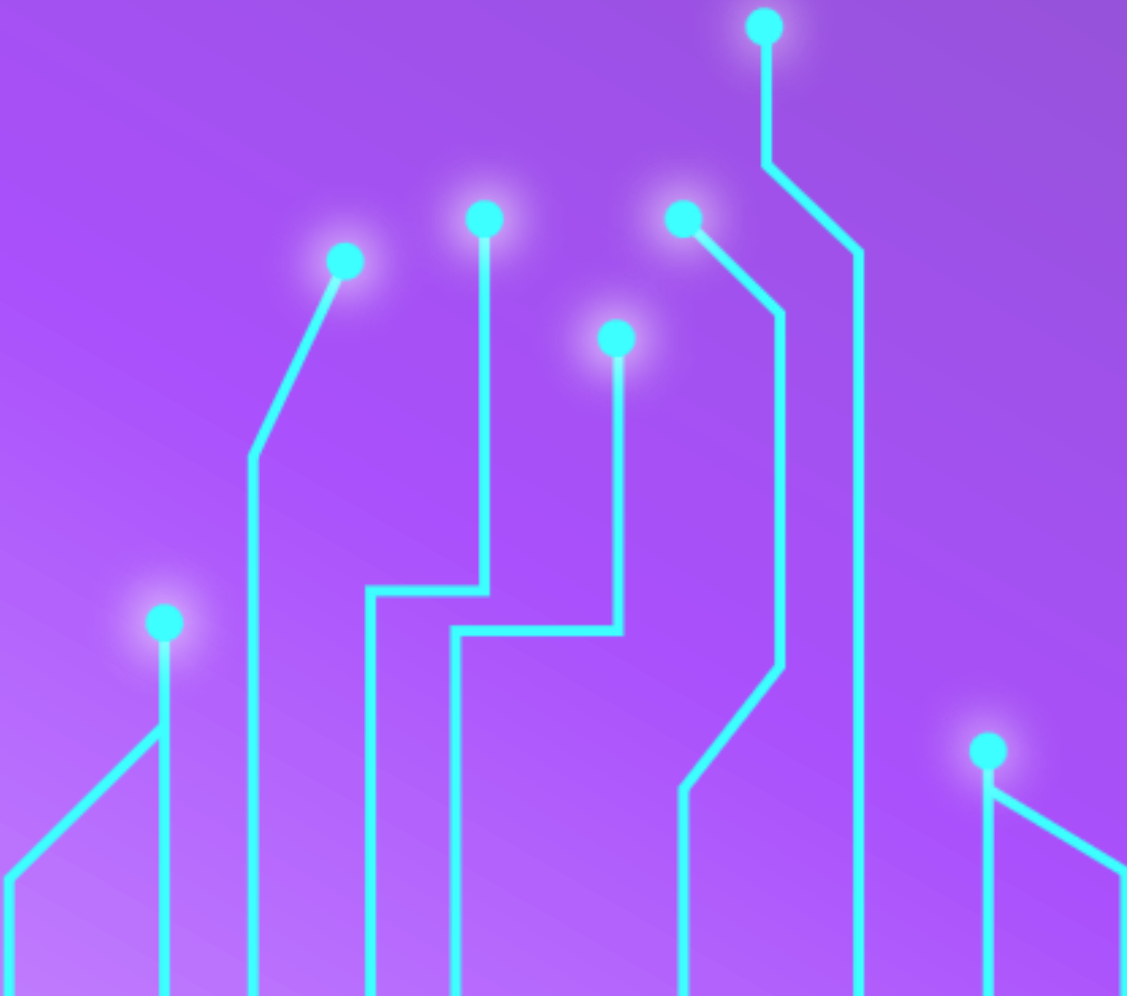
**04. Definición del sesgo algorítmico**

**05. Comprender los sesgos en los sistemas de IA**

# 01. Introducción

CU1 | ¿Qué es el sesgo algorítmico?





## 01. Introducción

En el panorama en rápida evolución de la inteligencia artificial (IA), es primordial garantizar el desarrollo y el uso éticos de las tecnologías de IA. Este folleto sirve como guía completa para la microcredencial de IA ética, centrándose en seis unidades de competencia diseñadas para equiparle con los conocimientos y habilidades necesarios para navegar por las complejidades éticas de la IA.

Al embarcarse en este viaje, se explorarán seis unidades de competencia distintas, cada una de las cuales aborda aspectos cruciales del desarrollo y el despliegue éticos de la IA. Desde la comprensión del sesgo algorítmico a la promoción de la transparencia y la defensa de los derechos humanos, estas unidades de competencia están diseñadas para proporcionar las herramientas necesarias para navegar por los desafíos éticos inherentes a las tecnologías de IA.

A lo largo de esta guía de apoyo, se profundizará en las siguientes Unidades de Competencia (UC a partir de ahora):

- UC1 - ¿Qué es el sesgo algorítmico?
- UC2 - No maleficencia
- UC3 - Rendición de cuentas
- UC4 - Transparencia
- UC5 - Derechos humanos y equidad
- UC6 - Ética de la IA, un enfoque práctico



Cada unidad le proporcionará una comprensión más profunda de los principios y prácticas éticas clave en la IA, junto con ideas prácticas y ejemplos del mundo real para reforzar su aprendizaje.

Tanto si eres un estudiante adulto, un profesional o un entusiasta de la IA, esta guía de apoyo te ofrece un valioso recurso para ampliar tus conocimientos y experiencia en IA ética. Te invitamos a embarcarte en este viaje con nosotros mientras exploramos las dimensiones éticas de la IA y trabajamos para crear un futuro más responsable y equitativo.

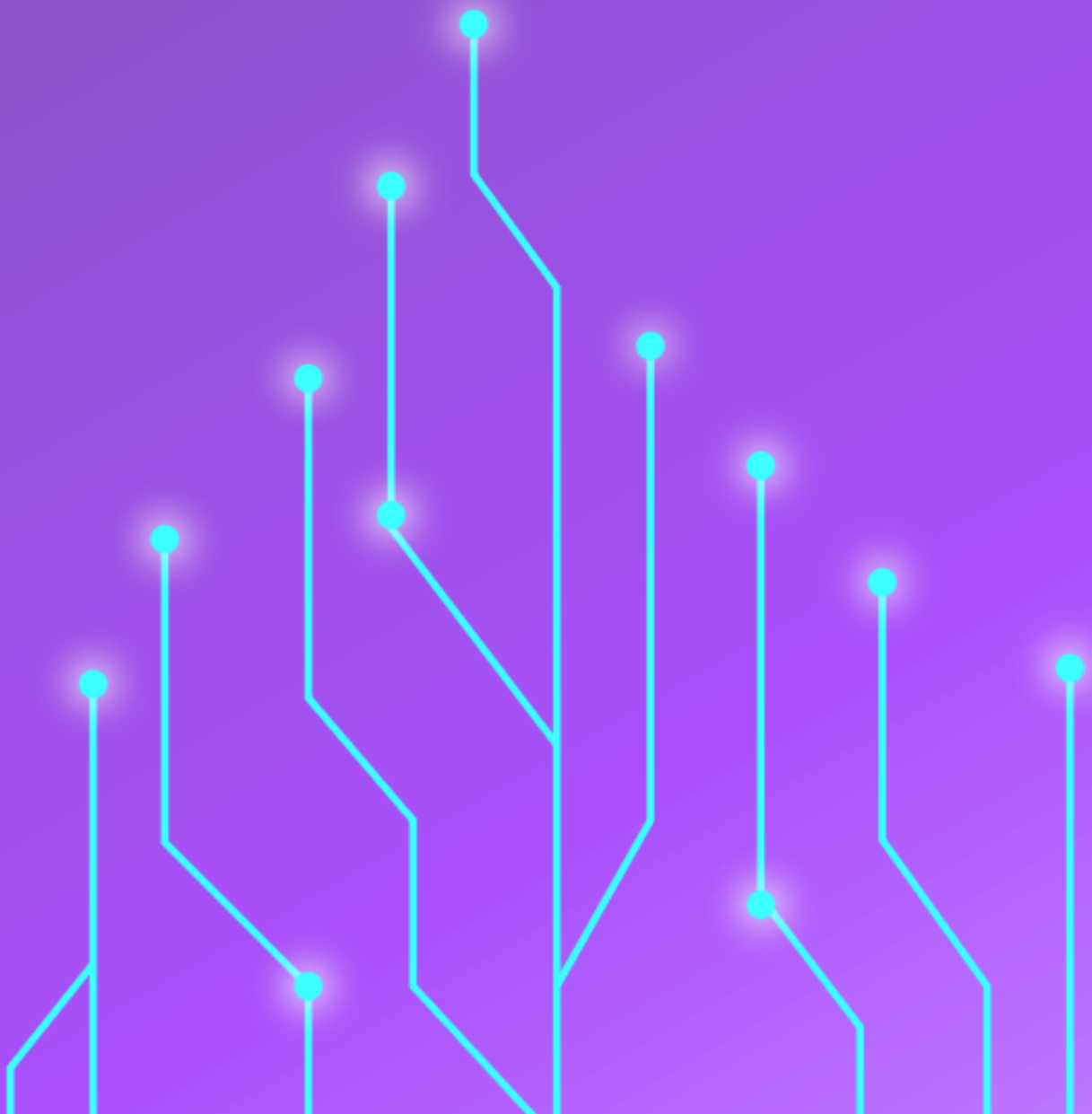
Gracias por leer esta guía de apoyo para el desarrollo y la práctica ética de la IA.

¡Empecemos juntos este viaje transformador!

*El equipo del proyecto CHARLIE*

# 02. Contenidos del curso y resultados esperados

UC1 | ¿Qué es el sesgo algorítmico?







## 02. Contenidos del curso y resultados esperados

La "Microcredencial de IA Ética" de nivel EQF4 está diseñada para alcanzar los siguientes **resultados**:

1. Establecer una **comprensión básica del sesgo algorítmico**, explorando sus orígenes e implicaciones para los individuos y la sociedad.
  - Profundice en la definición, las fuentes y las manifestaciones del sesgo algorítmico.
  - Analizar las ramificaciones sociales e individuales de los algoritmos sesgados.
2. Fomentar la concienciación y la aplicación del **principio ético de no maleficencia en el desarrollo de la IA**.
  - Evaluar los riesgos y daños asociados a los algoritmos sesgados.
  - Desarrollar estrategias para mitigar los daños y promover el desarrollo ético de la IA.
3. Apreciar la importancia de **la responsabilidad en los sistemas de IA**, examinando los marcos jurídicos y éticos pertinentes.
  - Investigar el papel de las distintas partes interesadas en la responsabilidad de la IA.
  - Conozca las mejores prácticas para fomentar la responsabilidad en el desarrollo de la IA.



4. Conocer el concepto de **transparencia en los sistemas de IA** y su papel fundamental en la toma de decisiones algorítmica.
  - Explorar metodologías y herramientas para mejorar la transparencia en la IA.
  - Comprender los retos y las limitaciones a la hora de hacer más comprensibles los algoritmos complejos.
5. Explorar **la intersección de la IA, los derechos humanos y la justicia**, y sus implicaciones para el desarrollo ético de la IA.
  - Evaluar el impacto de los algoritmos sesgados en los derechos humanos, incluidos los de no discriminación, privacidad y libertad de expresión.
  - Desarrollar estrategias para garantizar la justicia y la equidad en el desarrollo y despliegue de la IA.
6. Aplicar principios éticos en el desarrollo y despliegue de la IA mediante **enfoques prácticos y escenarios del mundo real**.
  - Examinar diversos marcos y directrices éticos y su aplicación a los sistemas de IA.
  - Comprender la importancia de la participación de las partes interesadas, la colaboración interdisciplinar y los procesos éticos de desarrollo de la IA.



Al finalizar este curso, los participantes poseerán una comprensión holística de los sesgos algorítmicos, sus impactos en sectores específicos y las herramientas y estrategias para abordarlos. Este conocimiento equipa a los profesionales y académicos/estudiantes en campos basados en algoritmos para contribuir a resultados más equitativos y justos en un mundo basado en datos.

El curso de microcréditos está estructurado en torno a **6 unidades de competencia (UC)**, cada una de ellas diseñada para dotar a los participantes de los conocimientos y habilidades necesarios para navegar por los retos y oportunidades en el ámbito del sesgo algorítmico.

**UC1 - ¿Qué es el sesgo algorítmico?** En esta unidad, los estudiantes explorarán el concepto de sesgo algorítmico y sus diversas manifestaciones. Abarca su definición, causas e implicaciones sociales. Los alumnos analizarán los orígenes del sesgo, las fuentes dentro de los algoritmos y las posibles repercusiones en las personas y la sociedad.

**UC2 - No maleficencia:** Esta unidad profundiza en el principio ético de no maleficencia, dando prioridad a la evitación del daño en el desarrollo y despliegue de la IA. Los participantes explorarán los riesgos y daños inherentes a los algoritmos sesgados, al tiempo que descubren estrategias para mitigar estos riesgos y fomentar prácticas éticas de IA.



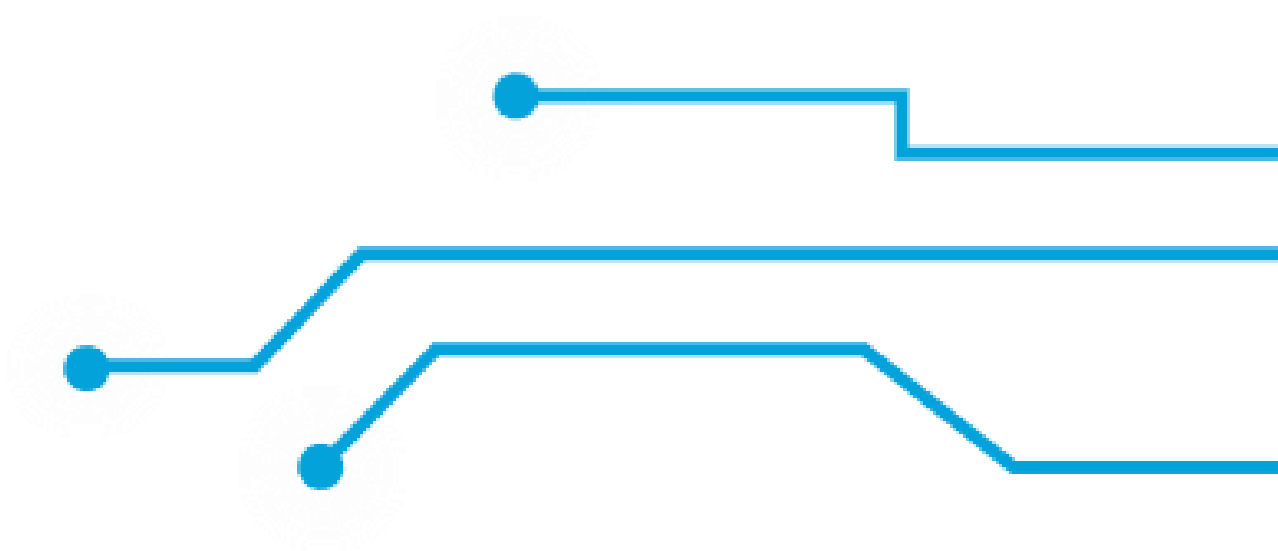
**UC3 - Responsabilidad:** En esta unidad, los estudiantes se adentran en el crítico ámbito de la rendición de cuentas en el desarrollo y la utilización de la IA. Exponiendo la necesidad de delinear líneas claras de responsabilidad, los participantes exploran los marcos legales y éticos que rigen la rendición de cuentas. Además, el plan de estudios examina las funciones de las distintas partes interesadas y profundiza en las mejores prácticas para garantizar la rendición de cuentas en todas las iniciativas de desarrollo de la IA.

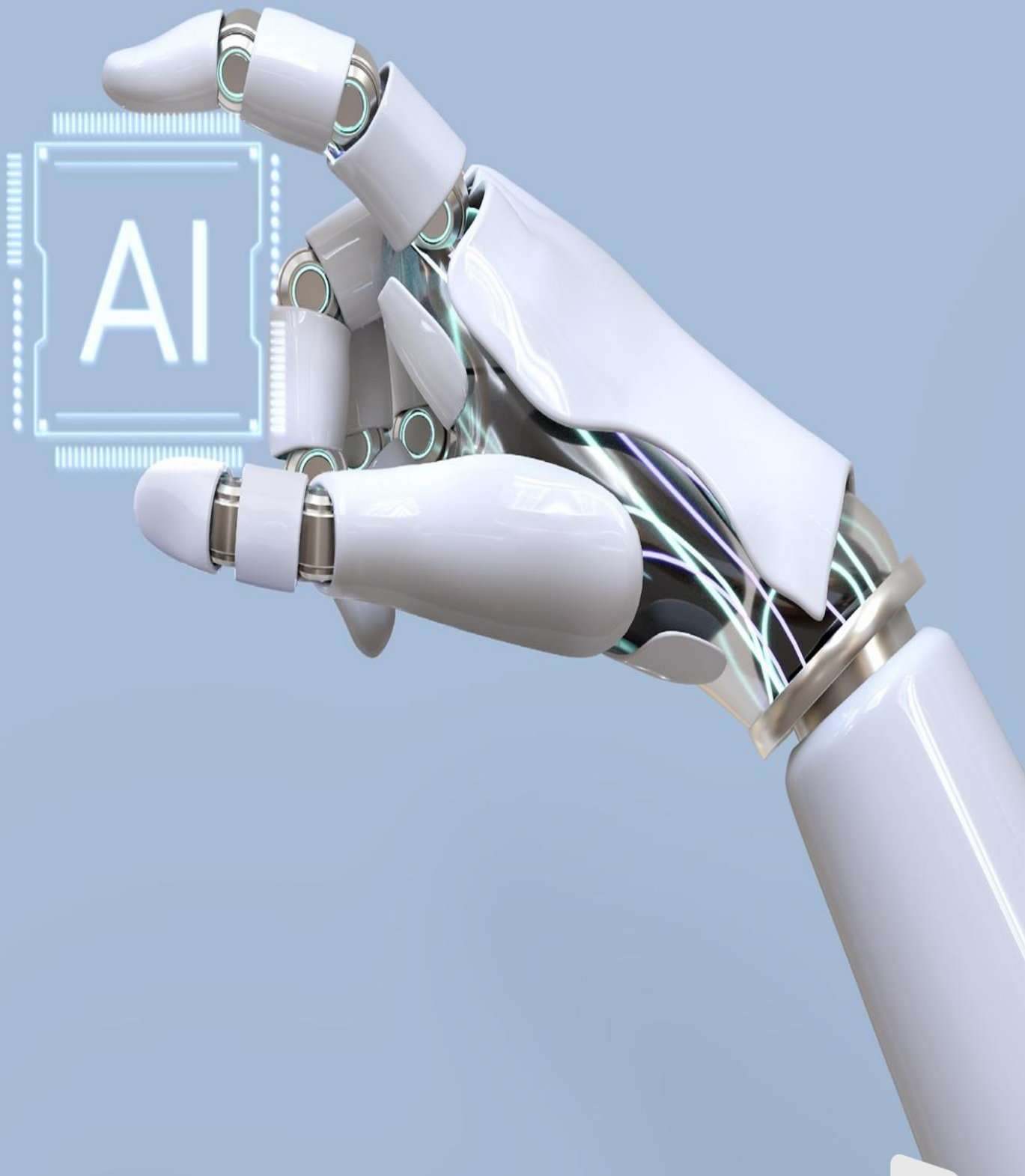
**UC4 - Transparencia:** Esta unidad pone de relieve la importancia de la transparencia en los sistemas de IA, haciendo hincapié en los valores de apertura, comunicación y explicabilidad en la toma de decisiones algorítmicas. Los participantes utilizarán técnicas y recursos destinados a aumentar la transparencia en la IA, al tiempo que se enfrentan a los retos y limitaciones inherentes a la comprensión de algoritmos complejos.

**UC5 - Derechos humanos y equidad:** En la unidad Derechos humanos y equidad, los estudiantes explorarán la intersección entre la IA, los derechos humanos y la equidad. Examinarán cómo los algoritmos sesgados pueden afectar a los derechos humanos, incluido el derecho a la no discriminación, la privacidad y la libertad de expresión. También aprenderán estrategias para garantizar la justicia y la equidad en el desarrollo y el despliegue de la IA.

**UC6 - Ética de la IA, un enfoque práctico:** Esta unidad hace hincapié en la aplicación pragmática de los principios éticos en el desarrollo y la implantación de la IA. Los participantes se adentran en diversos marcos y directrices éticos, y adquieren conocimientos sobre su aplicación en el mundo real en escenarios de IA. Además, la unidad subraya la importancia de la participación de las partes interesadas, la colaboración interdisciplinar y la integración de procesos éticos de desarrollo de la IA para fomentar la innovación responsable de la IA.

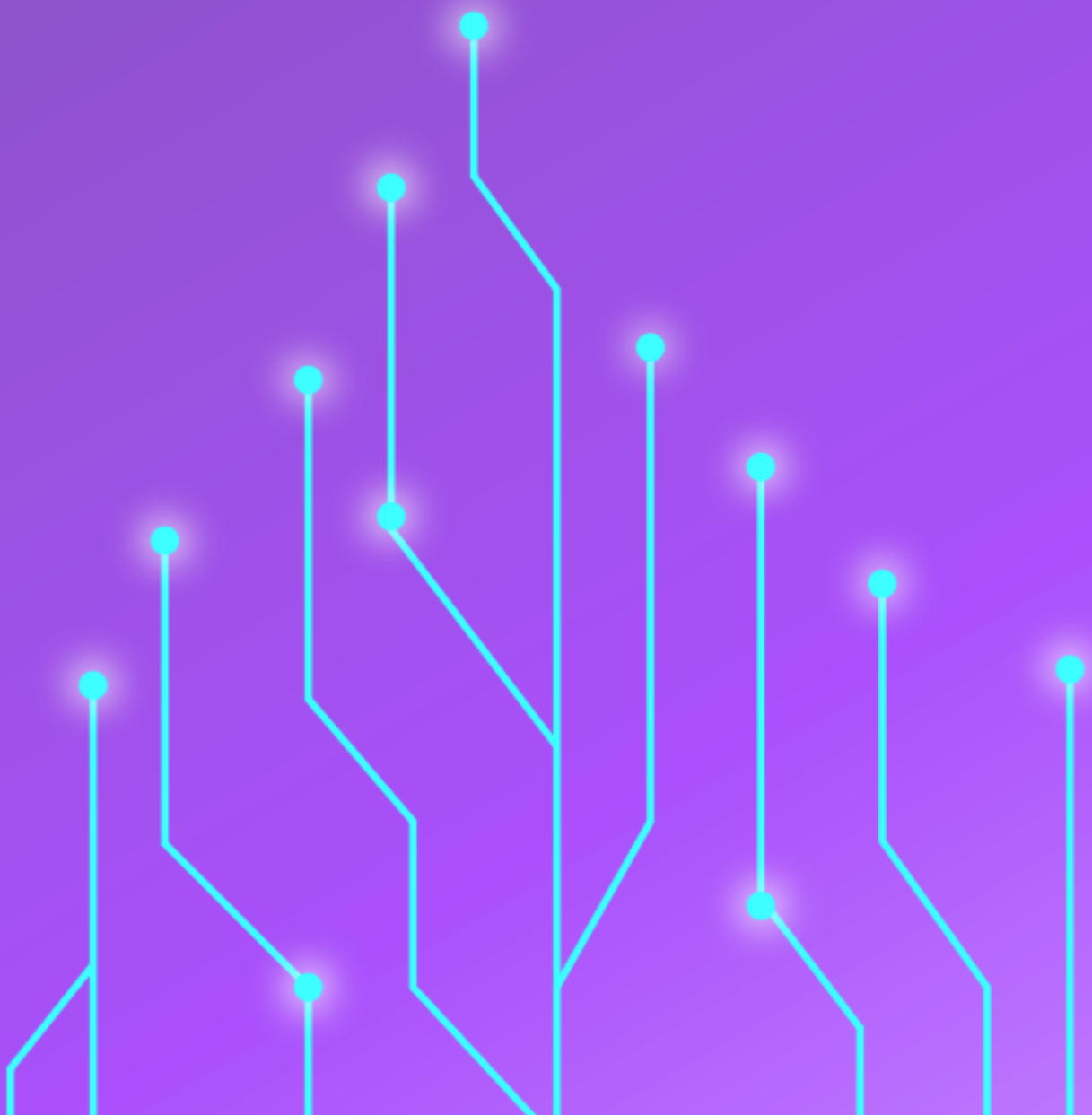
La sección siguiente detalla el contenido de cada unidad de competencia.





# 03. ¿Qué es el sesgo algorítmico?

UC1 | ¿Qué es el sesgo algorítmico?







### 03. ¿Qué es el sesgo algorítmico?

Los algoritmos se utilizan para tomar decisiones importantes. Sin embargo, a veces pueden ser parciales e injustos con ciertos grupos de personas. Es lo que se conoce como sesgo algorítmico.

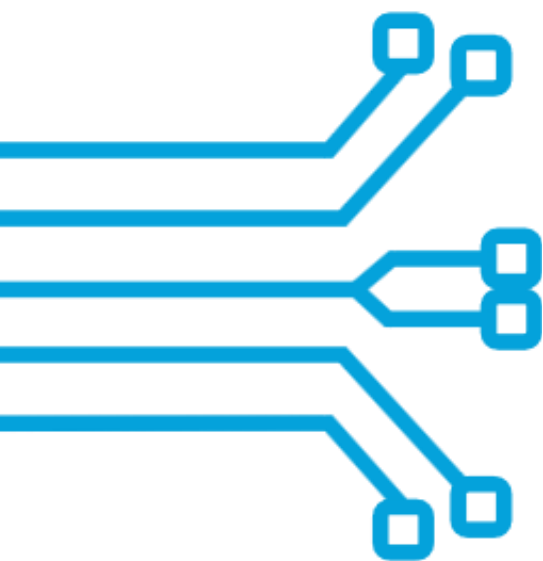
En esta unidad de competencia, se aprenderá sobre el sesgo algorítmico, sus diferentes formas y cómo identificarlo. También se explorarán las razones que subyacen al sesgo en los algoritmos, incluido el impacto del sesgo humano en la toma de decisiones. Además, se examinarán las posibles consecuencias de los algoritmos sesgados para las personas y la sociedad, que pueden conducir a la discriminación y el trato injusto. Al final de esta unidad, los alumnos comprenderán mejor el sesgo algorítmico y cómo abordarlo en su trabajo futuro.

Los resultados de conocimiento para esta unidad incluyen:

- **Definición de sesgo algorítmico:** se aprenderá sobre el sesgo algorítmico y sus causas, incluida la recopilación de datos sesgados, los datos de entrenamiento sesgados y la toma de decisiones humanas. Este conocimiento ayudará a entender cómo el sesgo puede afectar a las aplicaciones de IA, como los sistemas de reconocimiento facial que identifican erróneamente a ciertos grupos.

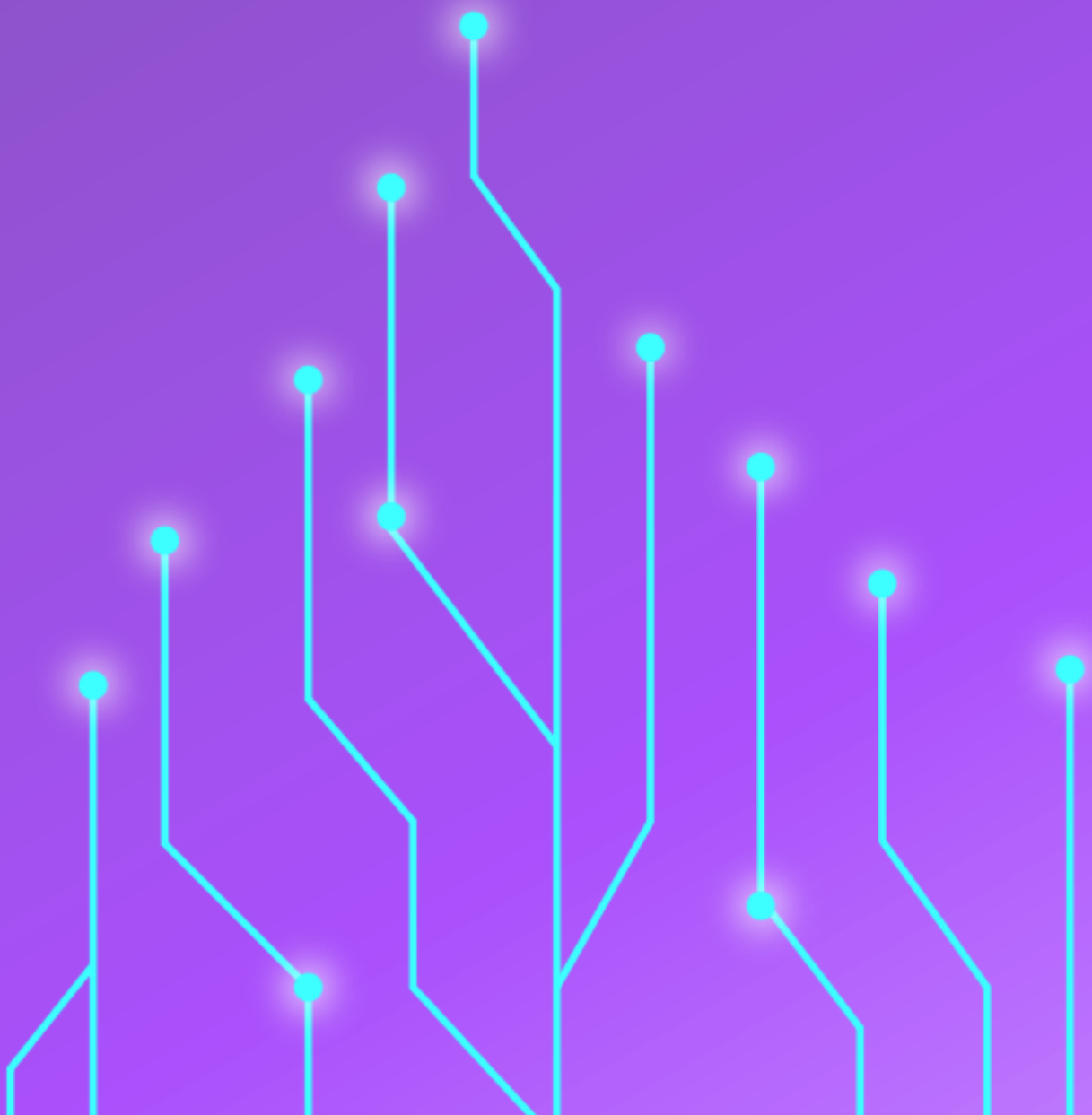


- **Identificación de los tipos de sesgos algorítmicos:** Los alumnos aprenderán sobre los sesgos algorítmicos, incluidos los sesgos basados en datos, modelos y humanos. Entenderán cómo estos sesgos pueden causar injusticias en los sistemas de IA. Por ejemplo, el sesgo basado en los datos puede ser el resultado de datos de entrenamiento no representativos, lo que lleva a predicciones sesgadas en áreas como la calificación crediticia o la selección de solicitantes de empleo.
- **Implicaciones en el mundo real del sesgo algorítmico:** En este curso, los estudiantes aprenderán sobre las consecuencias del sesgo algorítmico en diferentes sectores como la sanidad, las finanzas y la justicia penal. Comprenderán la necesidad de minimizar el sesgo algorítmico en los sistemas de IA para promover la justicia y la equidad. Se debatirán ejemplos de sistemas de IA sesgados que conducen a resultados negativos en la sanidad y la justicia penal.



# 04. Definición del sesgo algorítmico

UC1 | ¿Qué es el sesgo algorítmico?





## 04. Definición del sesgo algorítmico

El sesgo algorítmico es un aspecto crítico de la IA que ha acaparado la atención en los últimos años. Entenderlo es esencial para cualquiera que participe en el desarrollo, la implantación o la regulación de la IA. Definamos qué es el sesgo algorítmico y por qué es crucial estudiarlo.

### > ¿Qué es el sesgo algorítmico?

El sesgo algorítmico se refiere a errores sistemáticos o injustos en los resultados de los sistemas de IA debido a diversos factores, como datos sesgados, algoritmos defectuosos o toma de decisiones humanas. Estos sesgos pueden dar lugar a un trato discriminatorio o injusto de personas o grupos, perpetuando las desigualdades sociales existentes y reforzando los estereotipos.

### ¿Por qué estudiar el sesgo algorítmico?

Para explorar las diferentes formas, causas y consecuencias del sesgo algorítmico, es importante entender primero su definición y significado. Con este conocimiento, podemos equiparnos con las herramientas para identificar, mitigar y prevenir el sesgo algorítmico en los sistemas de IA.



- 1. Implicaciones éticas:** El sesgo algorítmico puede dar lugar a un trato injusto de las personas por motivos de raza, sexo, edad u otras características protegidas, violando los principios de imparcialidad y equidad.
- 2. Impacto social:** Los sistemas de IA sesgados pueden exacerbar las desigualdades sociales y la discriminación, afectando al acceso a oportunidades, recursos y servicios de las comunidades marginadas.
- 3. Preocupaciones legales y reglamentarias:** A medida que las tecnologías de IA se vuelven más omnipresentes, aumenta el escrutinio de los legisladores y los organismos reguladores para abordar el sesgo algorítmico con el fin de garantizar el cumplimiento de las leyes contra la discriminación y proteger los derechos de las personas.
- 4. Reputación y confianza:** Las organizaciones que despliegan sistemas de IA sesgados se arriesgan a sufrir daños en su reputación y a perder la confianza del público, lo que puede tener consecuencias importantes para su imagen de marca y su credibilidad en el mercado.

## ➤ Factores que contribuyen a unos resultados sesgados

Varios factores interrelacionados contribuyen a la aparición de sistemas de IA sesgados, lo que socava su fiabilidad, imparcialidad y eficacia. En esta sección explicaremos algunos de los factores más comunes que contribuyen a los resultados sesgados en los sistemas de IA.

- **Datos sesgados:** Los datos sesgados utilizados para entrenar sistemas de IA dan lugar a sesgos algorítmicos, que pueden conducir a resultados discriminatorios. Para mitigar estos sesgos, deben tenerse muy en cuenta la recopilación y el preprocesamiento de los datos, incluidos el muestreo representativo, los algoritmos de detección y mitigación de sesgos y el aumento de datos diversos.
- **Algoritmos defectuosos:** Los sistemas de IA pueden tener resultados sesgados debido a algoritmos defectuosos, elecciones de diseño, arquitecturas de modelos, procedimientos de optimización o variables de entrada. El aprendizaje automático consciente de la imparcialidad, la transparencia algorítmica y las técnicas de interpretabilidad pueden ayudar a mitigar estos sesgos.
- **Sesgos humanos:** Los sesgos en los sistemas de IA pueden ser el resultado de influencias inconscientes de desarrolladores, científicos de datos y responsables de la toma de decisiones. Para evitar estos sesgos, los equipos de desarrollo de IA deben centrarse en la diversidad, las directrices éticas y los mecanismos de rendición de cuentas.



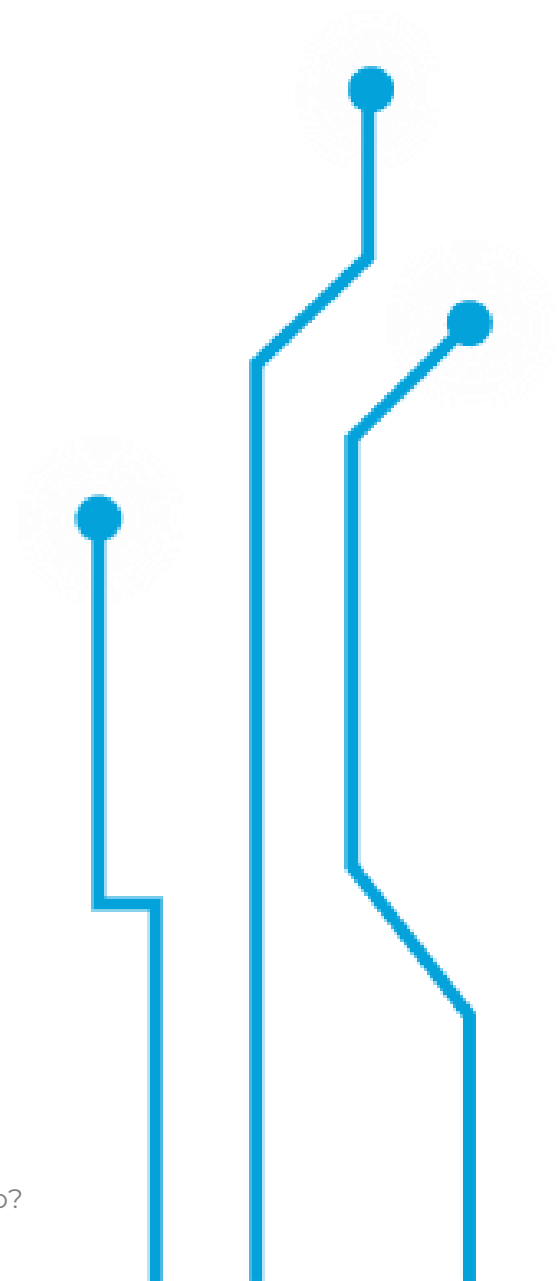


## > Ejemplos de sistemas sesgados

Los sistemas de IA pueden ser parciales y dar lugar a resultados injustos. A continuación se presentan algunos ejemplos reales de sistemas de IA que suelen estar sesgados, lo que pone de relieve las posibles consecuencias de los sesgos algorítmicos. Los analizaremos más a fondo en secciones posteriores de este curso.

- **Algoritmos de reconocimiento facial:** La tecnología de reconocimiento facial puede tener sesgos que perpetúen las disparidades raciales o de género, dando lugar a detenciones erróneas o a la vigilancia de grupos específicos. Abordar estos sesgos es crucial para garantizar la justicia y la equidad en los sistemas de IA y restaurar la confianza pública.
- **Algoritmos policiales predictivos:** Los algoritmos de predicción policial pueden perpetuar los sesgos presentes en los datos históricos sobre delincuencia, lo que puede llevar a una vigilancia excesiva de determinadas comunidades o grupos demográficos. Los algoritmos sesgados pueden exacerbar las disparidades existentes en las prácticas de aplicación de la ley y suscitar preocupación sobre la imparcialidad, la responsabilidad y la posibilidad de resultados discriminatorios en los sistemas de justicia penal.

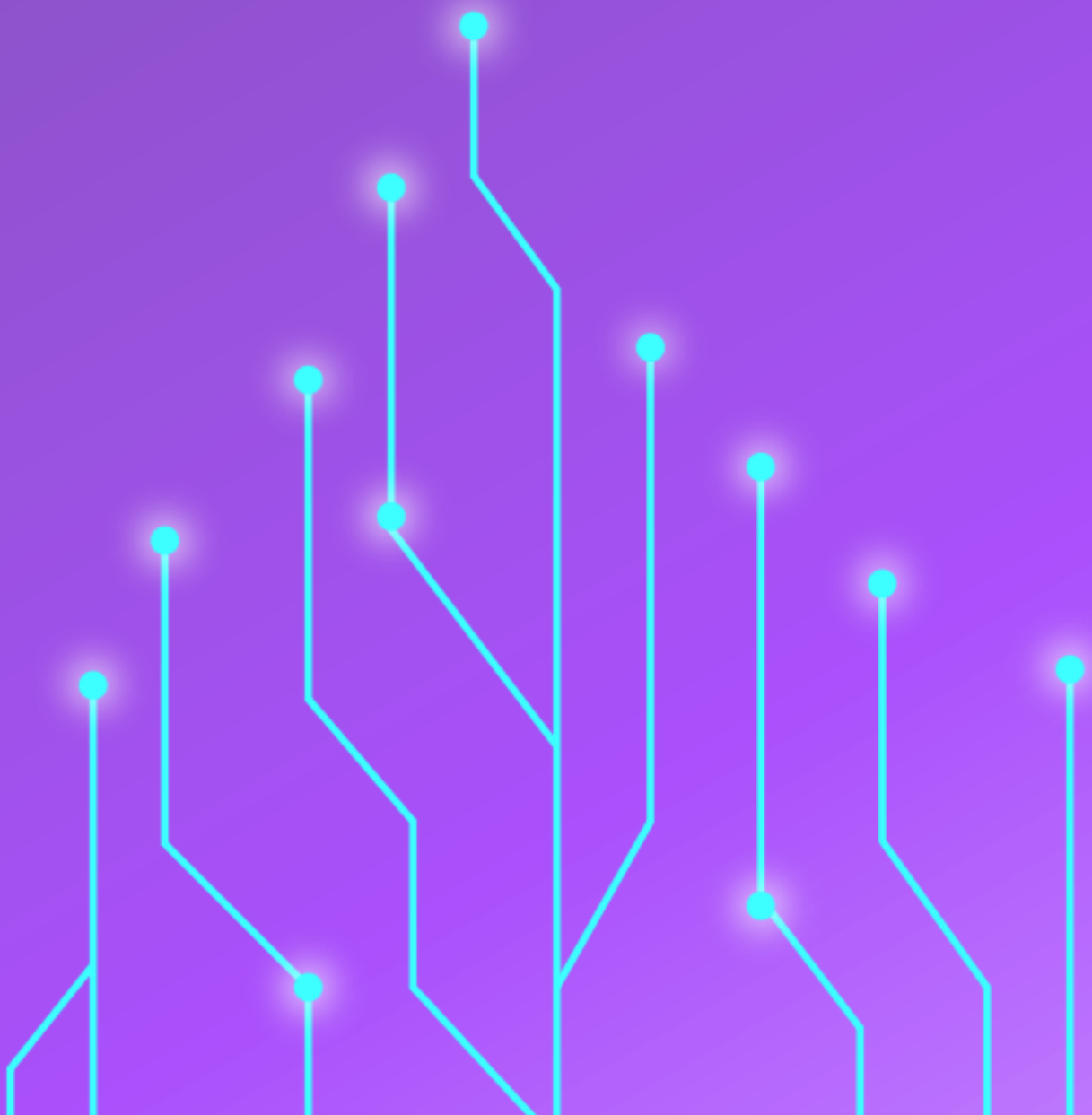
- **Sistemas de contratación automatizados:** Los sistemas automatizados de contratación pueden perpetuar los prejuicios, dando lugar a prácticas discriminatorias y limitando la diversidad en la mano de obra. Los algoritmos sesgados pueden aprender patrones de sesgo a partir de datos históricos, lo que resulta en un trato preferencial de ciertos grupos demográficos.





# 05. Comprender los sesgos en los sistemas de IA

UC1 | ¿Qué es el sesgo algorítmico?







## 05. Comprender los sesgos en los sistemas de IA

En esta sección se explorarán tres tipos de sesgo: el basado en datos, el basado en modelos y el basado en el ser humano.

Estos sesgos pueden afectar a la precisión y fiabilidad de los sistemas de IA, y comprenderlos es el primer paso para prevenirlos.

### > Sesgo basado en los datos

¿Qué es el sesgo basado en los datos?

El sesgo basado en los datos se refiere a los sesgos que surgen de las características o la distribución de los datos de entrenamiento utilizados para desarrollar modelos de aprendizaje automático.

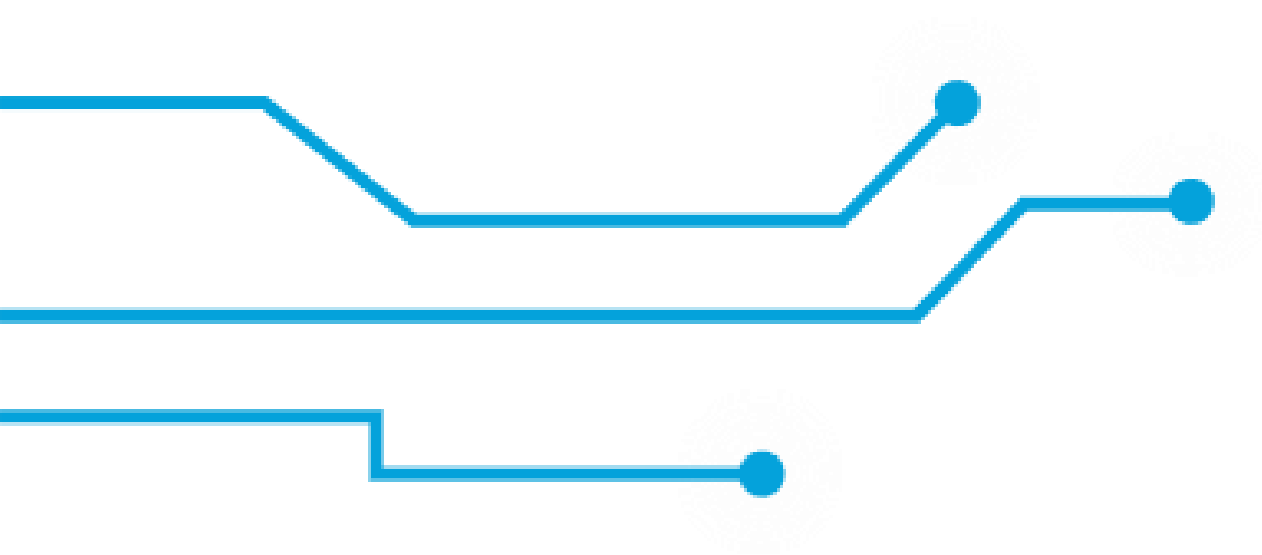
Los datos de formación sesgados pueden reflejar desigualdades históricas, prejuicios sociales o discriminación sistémica, lo que conduce a representaciones sesgadas de determinados grupos demográficos o a la infrarrepresentación de otros.

Comprender el sesgo impulsado por los datos es esencial para reconocer cómo los datos de entrenamiento sesgados pueden perpetuar y exacerbar los estereotipos, las desigualdades y las prácticas discriminatorias existentes en los sistemas de IA.



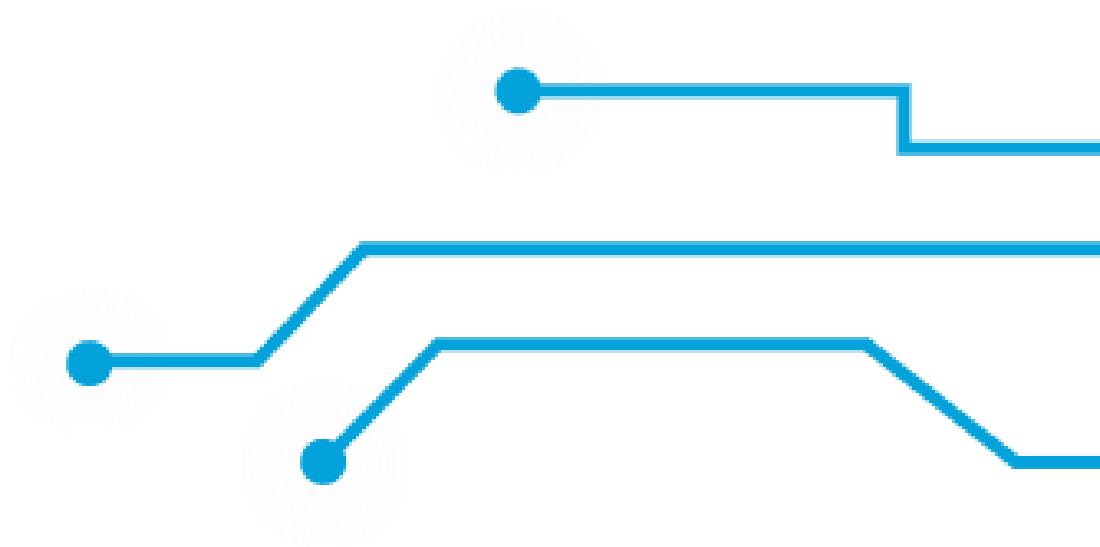
## Causas del sesgo de los datos

- 1. Muestreo incompleto o sesgado:** Los conjuntos de datos de entrenamiento pueden carecer de diversidad o no representar adecuadamente a determinados grupos demográficos, lo que da lugar a representaciones sesgadas y predicciones sesgadas del modelo.
- 2. Prejuicios históricos:** Los datos de entrenamiento pueden reflejar desigualdades históricas o sesgos sistémicos presentes en la sociedad, perpetuando resultados discriminatorios en los sistemas de IA.
- 3. Sesgos de etiquetado:** Las prácticas de etiquetado sesgadas o subjetivas pueden introducir sesgos en los datos de entrenamiento, influyendo en las predicciones del modelo y reforzando los estereotipos existentes.



## Ejemplos de sesgo basado en datos

- 1. Reconocimiento facial sesgado:** Los algoritmos de reconocimiento facial entrenados en conjuntos de datos desequilibrados pueden mostrar sesgos raciales o de género, lo que lleva a una identificación errónea y a la discriminación de ciertos grupos demográficos.
- 2. Prejuicios sexistas en los modelos lingüísticos:** Los modelos lingüísticos entrenados a partir de corpus de texto sesgados pueden generar un lenguaje estereotipado o discriminatorio en función del género, reflejando y perpetuando los prejuicios sociales.
- 3. Prejuicios raciales en la policía predictiva:** Los algoritmos de predicción policial basados en datos delictivos sesgados pueden dirigirse de forma desproporcionada a comunidades minoritarias, exacerbando las disparidades raciales en la aplicación de la ley.







## Impacto del sesgo de los datos

- 1. Refuerzo de los estereotipos:** Los datos de entrenamiento sesgados pueden reforzar los estereotipos y prejuicios existentes, perpetuando la discriminación y la desigualdad en los sistemas de IA.
- 2. Amplificación de la desigualdad:** El sesgo basado en los datos puede exacerbar las desigualdades y disparidades existentes, dando lugar a un trato injusto y a resultados discriminatorios para los grupos marginados.
- 3. Erosión de la confianza:** Los sistemas de IA sesgados socavan la confianza en la tecnología, exacerbando la preocupación por la imparcialidad, la rendición de cuentas y la transparencia.

El sesgo basado en los datos es un reto importante para los sistemas de IA justos y equitativos. Al comprender sus causas y consecuencias, las partes interesadas pueden tomar medidas proactivas para mitigar el sesgo en los datos de entrenamiento y promover la inclusión en la IA.

## > Sesgo de modelo

### ¿Qué es el sesgo basado en modelos?

El sesgo impulsado por el modelo se refiere a los sesgos que surgen del diseño, la estructura o la optimización de los modelos de aprendizaje automático, que conducen a resultados discriminatorios o predicciones sesgadas.

### Causas del sesgo basado en modelos

- 1. Sesgos en la selección de características:** Las características del modelo seleccionadas durante el proceso de modelado pueden codificar inadvertidamente sesgos presentes en los datos de entrenamiento, lo que conduce a predicciones sesgadas o resultados discriminatorios.
- 2. Complejidad algorítmica:** Los algoritmos complejos de aprendizaje automático pueden captar y reforzar sesgos sutiles presentes en los datos de entrenamiento, amplificando su impacto en las predicciones del modelo.
- 3. Objetivos de optimización:** Los objetivos de optimización definidos durante el proceso de entrenamiento del modelo pueden priorizar inadvertidamente ciertos resultados sobre otros, dando lugar a predicciones sesgadas o injustas.





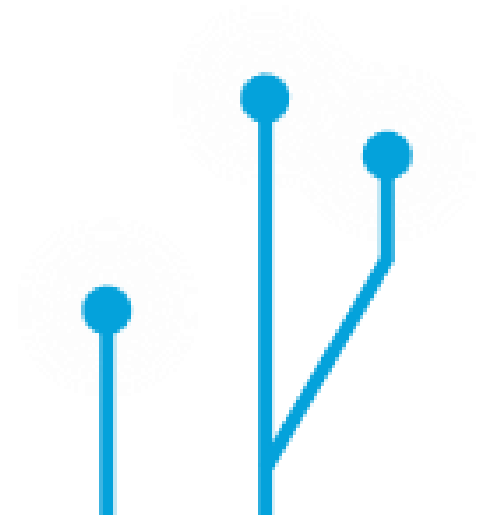
## Ejemplos de sesgo basado en modelos

- 1. Sesgo de género en los algoritmos de contratación:** Los algoritmos automatizados de contratación pueden favorecer inadvertidamente a los candidatos masculinos en detrimento de las candidatas femeninas debido a una selección sesgada de características o a objetivos de optimización, perpetuando las disparidades de género en la mano de obra.
- 2. Sesgo racial en los algoritmos de imposición de penas:** Los algoritmos de predicción de penas utilizados en los sistemas de justicia penal pueden recomendar de forma desproporcionada penas más duras para los acusados pertenecientes a minorías, amplificando las disparidades raciales en las tasas de encarcelamiento.
- 3. Sesgo socioeconómico en los modelos de aprobación de préstamos:** Los modelos de aprendizaje automático utilizados para la aprobación de préstamos pueden denegar sistemáticamente préstamos a personas de comunidades marginadas, exacerbando las desigualdades socioeconómicas en el acceso a los servicios financieros.

## Impacto del sesgo basado en modelos

- 1. Perpetuación de la discriminación:** Los prejuicios basados en el modelo pueden perpetuar y reforzar la discriminación y las desigualdades existentes en la sociedad, dando lugar a un trato injusto y a resultados sesgados para los grupos marginados.
- 2. Falta de responsabilidad:** Los modelos de IA sesgados pueden carecer de transparencia y responsabilidad, lo que dificulta identificar y abordar las prácticas discriminatorias en los sistemas de IA.
- 3. Implicaciones éticas:** El sesgo basado en modelos plantea problemas éticos relacionados con la equidad, la justicia y los derechos humanos, lo que pone de relieve la necesidad de directrices y normativas éticas que rijan el desarrollo y la implantación de la IA.

El sesgo basado en modelos plantea importantes retos para el desarrollo y despliegue de sistemas de IA justos y responsables. Al comprender los mecanismos y las implicaciones del sesgo impulsado por modelos, las partes interesadas pueden aplicar estrategias para mitigar el sesgo y promover la justicia y la equidad en las tecnologías de IA.





## > Prejuicios humanos

### ¿Qué son los prejuicios humanos?

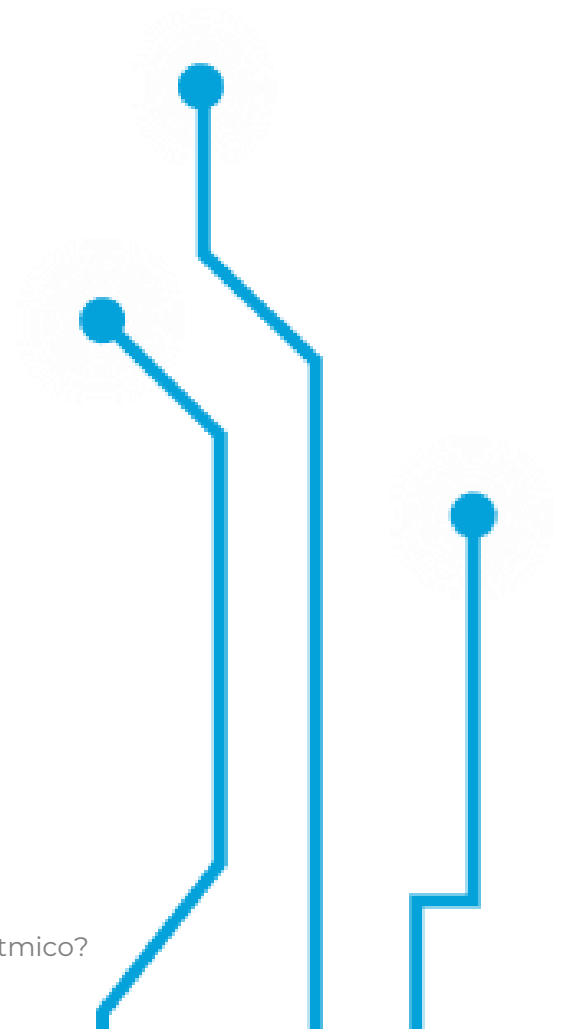
El sesgo de origen humano en la IA se refiere a los sesgos que surgen de las decisiones, acciones o juicios de las personas implicadas en el desarrollo y despliegue. Puede derivarse de sesgos cognitivos, influencias culturales y prejuicios sociales, que conducen a resultados sesgados o prácticas discriminatorias.

### Causas de los prejuicios humanos

- 1. Sesgos en la recogida de datos:** Los sesgos en la recogida de datos, como los de muestreo o selección, pueden dar lugar a datos de entrenamiento sesgados y a predicciones sesgadas de los modelos.
- 2. Sesgos en el diseño algorítmico:** Los algoritmos de IA pueden estar sesgados debido a las decisiones de los diseñadores y desarrolladores humanos, lo que perpetúa los resultados sesgados en los sistemas de IA.
- 3. Sesgos de interpretación y despliegue:** Los intérpretes humanos y los responsables de la toma de decisiones pueden mostrar sesgos a la hora de desplegar sistemas de IA, lo que puede dar lugar a prácticas discriminatorias y a un trato injusto.

## Ejemplos de prejuicios humanos

- 1. Sesgos en los sistemas de reconocimiento facial:** Los sesgos humanos en la recopilación de datos de entrenamiento y el diseño algorítmico pueden dar lugar a sesgos raciales o de género en los sistemas de reconocimiento facial, dando lugar a una identificación errónea o a una infrarrepresentación de determinados grupos demográficos.
- 2. Equidad en los algoritmos de contratación:** Los sesgos en los procesos humanos de toma de decisiones, como la selección de currículos o las evaluaciones de entrevistas, pueden perpetuar las disparidades de género o raciales en los resultados de la contratación, incluso cuando se utilizan algoritmos de contratación basados en IA.





## Impacto de los prejuicios humanos

- 1. Exacerbación de las desigualdades existentes:** Los sesgos humanos en la IA pueden exacerbar las desigualdades y disparidades existentes en la sociedad. La recopilación de datos, el diseño algorítmico y la interpretación sesgados pueden dar lugar a un trato injusto para los grupos marginados, perpetuando la discriminación y obstaculizando el progreso social.
- 2. Erosión de la confianza pública:** Los sistemas de IA contaminados por prejuicios humanos pueden erosionar la confianza pública en la tecnología. Pueden surgir preocupaciones sobre la imparcialidad, la transparencia y la rendición de cuentas, obstaculizando la adopción y aceptación de la IA en diversos sectores.
- 3. Menor eficacia de los sistemas de IA:** Los prejuicios humanos pueden socavar la eficacia de los sistemas de IA. Los datos de formación sesgados o las interpretaciones sesgadas de los humanos pueden dar lugar a predicciones inexactas, recomendaciones erróneas y resultados subóptimos, lo que obstaculiza los beneficios potenciales de la IA.

El sesgo humano es un reto importante para crear sistemas de IA justos y responsables. Al comprender y mitigar los sesgos algorítmicos, las partes interesadas pueden crear sistemas de IA más fiables y transparentes.



# Charlie



Cofinanciado por  
la Unión Europea

Financiado por la Unión Europea. Las opiniones y puntos de vista expresados solo comprometen a su(s) autor(es) y no reflejan necesariamente los de la Unión Europea o los de la Agencia Ejecutiva Europea de Educación y Cultura (EACEA). Ni la Unión Europea ni la EACEA pueden ser considerados responsables de ellos.



Universitat  
de les Illes Balears



helixconnect



2022-1-ES01-KA220-HED-000085257