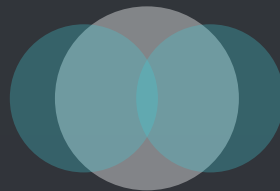


The image features a split background. On the left is a heatmap with a color gradient from blue to red, overlaid on a street grid. On the right is a standard street map of Los Angeles, California, with several red rectangular boxes highlighting specific areas. The text is centered over the heatmap portion.

# Hacia un uso (y diseño) ético de la IA en la seguridad ciudadana

Fernando Miró Llinares  
Cofundador de PlusEthics  
Catedrático de Derecho penal y Criminología  
Director del Centro de Investigación CRÍMINA  
Universidad Miguel Hernández de Elche

CRÍMINA



P L U S  
ethics

CRÍMINAI

# IA, delincuencia y seguridad pública

☰ **eldiario.es**

**TECNOLOGÍA** **Contra la violencia machista, el odio y las denuncias falsas: los algoritmos**

Contra la violencia machista, el odio y las denuncias falsas: los algoritmos que usa la Policía

☰ **EL PAIS** 🔔 🔍

**El policía que aprendió a programar para predecir crímenes**

El policía italiano Elia Lombardo desarrolla un programa que se utiliza en Nápoles y Venecia y que permite reducir un 24% robos, hurtos y estafas

☰ **EL HERALDO** DE SALTILLO 🔍

Inicio > EL MUNDO

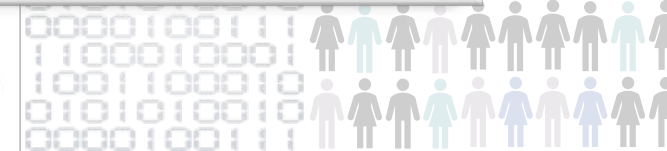
**EL MUNDO** **MEXICO**

Cámaras con inteligencia artificial de la CDMX ya podrán evitar delitos

Por **El Heraldo de Saltillo** - 5 junio, 2019 👁️ 76 💬 0

🔍 **DV** Tecnología | Algoritmos inteligentes para combatir el crimen

**Algoritmos inteligentes para combatir el crimen mejor que en 'Breaking Bad'**



# IA, delincuencia y seguridad pública



**sinc**  
La ciencia es noticia

PORTADA CIENCIAS NATURALES **TECNOLOGÍAS** BIOMEDICINA Y SALUD MATEMÁTICAS, FÍSICA Y QUÍMICA HUMANO

NOTICIAS REPORTAJES ENTREVISTAS MULTIMEDIA AGENDA ESPECIALES

**MATEMÁTICAS, FÍSICA Y QUÍMICA: Matemáticas**

Mathy O'Neil desmonta la supuesta neutralidad del 'big data'

## Algoritmos para la manipulación política y la discriminación



Abogacia Española  
CONSEJO GENERAL

Blogs → Nuevas Tecnologías → Ese algoritmo, el discriminador

NUEVAS TECNOLOGÍAS 14 MAYO, 2018

## Ese algoritmo, el discriminador



## REVISTA *CLOUD COMPUTING*

### Discriminación, prejuicio, injusticia, el impacto del sesgo en los algoritmos



MIT  
Technology  
Review

Secciones +

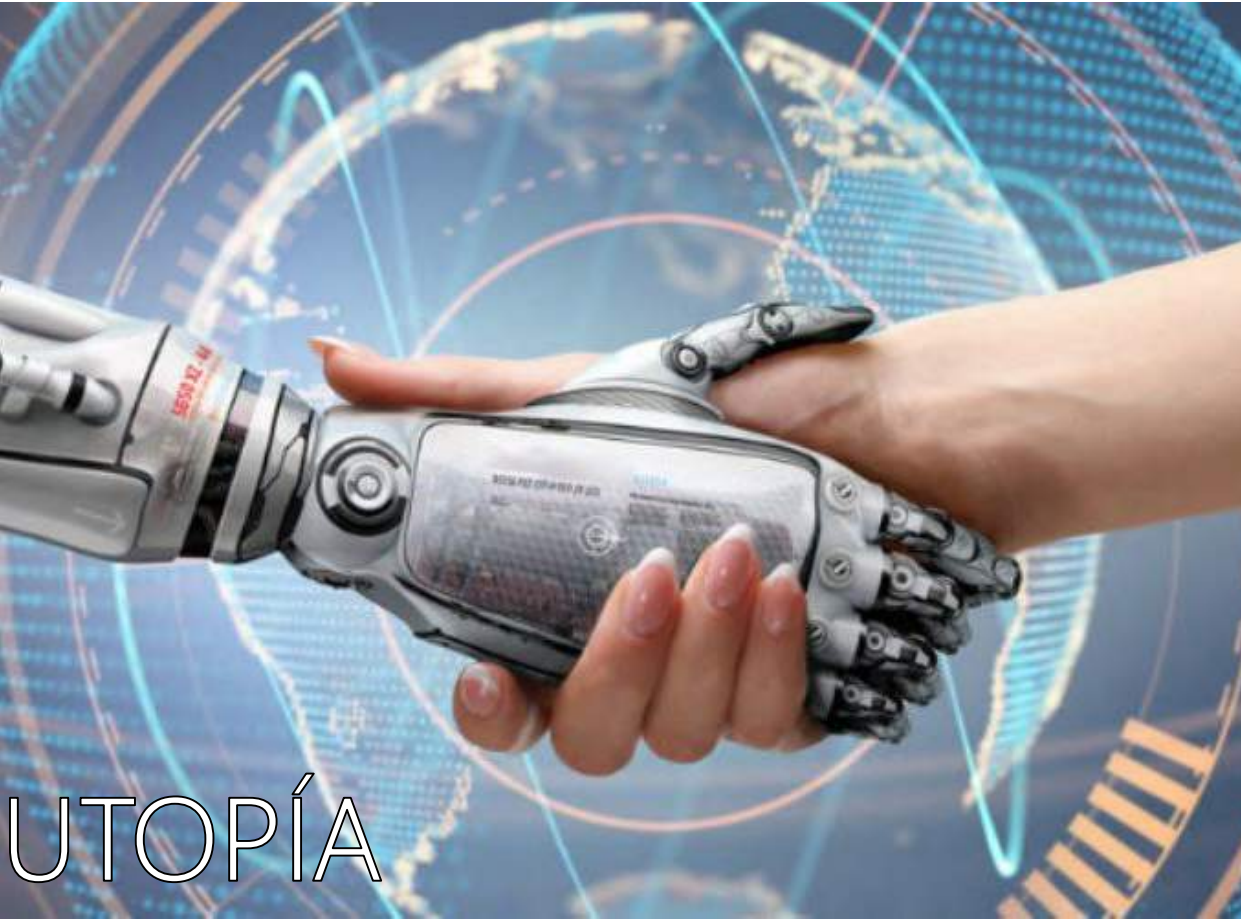
Computación

## La inteligencia artificial ya es capaz de evitar la prisión preventiva innecesaria

Un programa de aprendizaje automático analiza el historial del acusado para determinar si debe esperar a su juicio desde casa o presenta riesgo de fuga. Y resulta que funciona mejor que los jueces



# 0. Actitudes ante la IA: entre extremos anda el juego

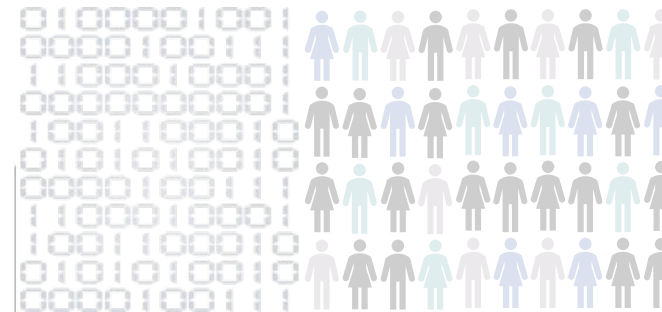
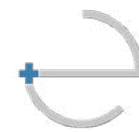


UTOPIÍA



DISTOPÍA

CRÍMINAI



# Predictive Policing: Utopia or Dystopia? On attitudes towards the use of Big Data algorithms for law enforcement<sup>1</sup>

Dr. Fernando Miró-Llinares

Full Professor of Criminal Law and Criminology, Universidad Miguel Hernández de Elche

Date of submission: October 2019

Accepted in: December 2019

Published in: March 2020

## Abstract

The use of predictive AI tools to improve decision-making in relation to crime prevention and investigation is a reality. They are being implemented almost before we fully understand how they work, while we make legal decisions that may determine the progress of the technology, and long before we can predict their full impact. This paper addresses the attitudes towards this technological revolution applied to criminal justice, focusing in particular on its use by police. The first section summarizes and describes the techniques and technologies that make up predictive policing. Subsequently, the main part of the study analyzes how this technology has been received. This ranges from optimism from those who defend its immediate implementation as a way to improve police objectivity and efficiency, to pessimism from those who see its use as fostering a dystopia of State control and surveillance. These two extremes reflect the shift from optimism to technological pessimism in the twentieth century. The article concludes with a defense of a realistic, critical and informed view of the use of these predictive algorithms. This vision accepts that there are no neutral technologies yet does not fall into fatalism and technophobia. It is one that places the human being and legitimate policing at the centre of the algorithmic equation while redefining its objectives based on the scientific evidence applied to each individual technology.

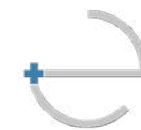
1. This research has been conducted as part of the project "Criminology, empirical evidence and criminal policy", Reference: DER2017-86204-R, financed by the State Research Agency (AEI)/Ministry of Science, Innovation and Universities and the European Union through the European Fund for Regional Development -FEDER- "A way to make Europe"; and within the framework of the grants for dynamization activities "Networks of Excellence" project: Development of a criminological and empirical model of criminal policy - EmpiriC. Financed by MCIUAEI (Ref. DER2017-90552-REDT); and thanks to a internationalization grant from UMH in the Max Planck Institute for Foreign and International Criminal Law of Freiburg. I would like to thank Steven Kemp for the translation of the article.

Actitud  
filosófica ante  
el Uso de la IA  
en el ámbito  
policial y de  
justicia:

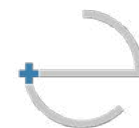
Realista

Informada  
empíricamente

Crítica (ética)



0. Entre extremos anda el juego: actitudes ante el uso policial/penal de IA
1. Marco normativo de la IA (ética para una revolución en construcción)
  - 1.1. IA en construcción (realismo frente a la mitología)
  - 1.2. Normas mutables, principios éticos rígidos (y flexibles)
2. Consideraciones éticas específicas de la IA para la seguridad ciudadana
  - 2.1. Problemas éticos del uso de la IA en el ámbito policial
  - 2.2. Principios éticos para el uso de IA en el ámbito policial y de justicia



# MARCO NORMATIVO DE LA **INTELIGENCIA ARTIFICIAL**

Ética para una revolución en construcción





## 1.1. IA en construcción (realismo frente a la mitología)

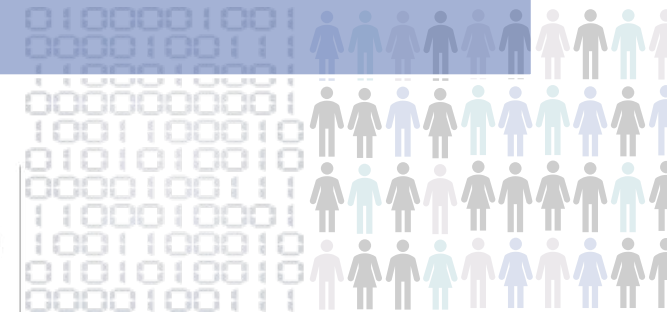
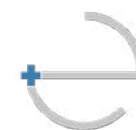
### EL PROBLEMA DEL NOMBRE Inteligencia Artificial

Eufemismo

- Sistemas para el tratamiento y análisis automático de la información

Desiderátum

- Voluntad de emular los procesos cognitivos humanos



# Mitos de la IA vs realidad de la IA (débil)

MITO	REALIDAD (IA DÉBIL)
El mito de la predicción del futuro	Estimación a partir de los datos del pasado
El mito de la decisión totalmente autónoma	Man in the Loop (control del contexto)
El mito de la alquimia matemática	Garbage in, garbage out; cream in, cake out
El mito de la distopía de control estatal	Tanto IA como datos en manos privadas
El mito de la perfección humana	La caja gris tiene más sesgos (y ruido) que la caja plateada

## 1.2. Normas mutables, principios éticos rígidos (y flexibles)



La **fiabilidad** de la inteligencia artificial (IA) se apoya en tres componentes que deben satisfacerse a lo largo de todo el ciclo de vida del sistema:

- a) la IA debe ser **lícita**, es decir, cumplir todas las leyes y reglamentos aplicables;
- b) ha de ser **ética**, de modo que se garantice el respeto de los principios y valores éticos; y
- c) debe ser **robusta**, tanto desde el punto de vista técnico como social, puesto que los sistemas de IA, incluso si las intenciones son buenas, pueden provocar daños accidentales.



# Legalidad- el reto de la privacidad

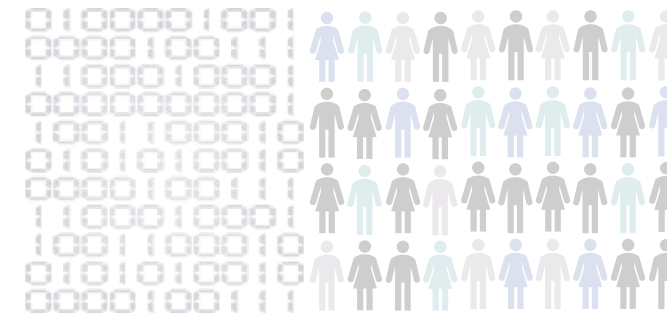
## NOVEDADES – OBLIGACIONES QUE INCORPORA EL INFORME

### 1. Información mínima a aportar por el responsable:

- Datos empleados y su importancia relativa para la toma de decisión.
- Calidad de los datos de entrenamiento.
- Tipo de patrón utilizado.

### 2. Trazabilidad de la información:

Preciso incluir un modelo de gobernanza de la información que permita a los usuarios identificar al responsable del tratamiento y ejercitar sus derechos.

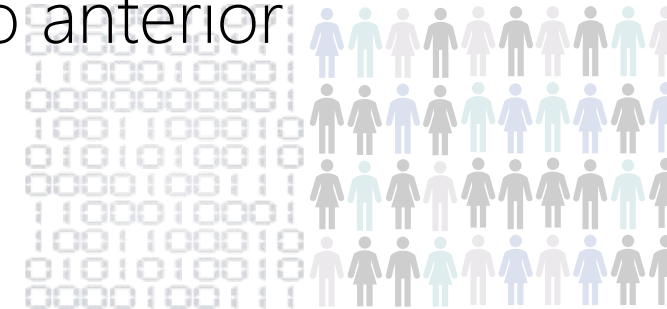


# Legalidad- Sentencia caso SyRI

El 5 de febrero de 2020 el Tribunal de Distrito de la Haya falla en contra del Gobierno de Países Bajos por la utilización del algoritmo SyRI (System Risk Indication) para detectar el riesgo de los ciudadanos de cometer fraude contra la SS:

El algoritmo es ilegal porque:

1. Viola el derecho a la privacidad del art. 8 CEDH.
2. Desproporcionalidad en los medios
3. Falta de transparencia: ni algoritmo, ni el modelo de riesgo
4. Indefensión del interesado como consecuencia de lo anterior



# Desde la perspectiva ética

## Principios éticos:

I) Respeto de la autonomía humana

Prohibición de subordinación

Obligación de supervisión

II) Prevención del daño

Protección Integridad y Dignidad

Atención a Vulnerabilidad y asimetría

III) Equidad

Prohibición de discriminación y estigmatización

Proporcionalidad entre medios y fines

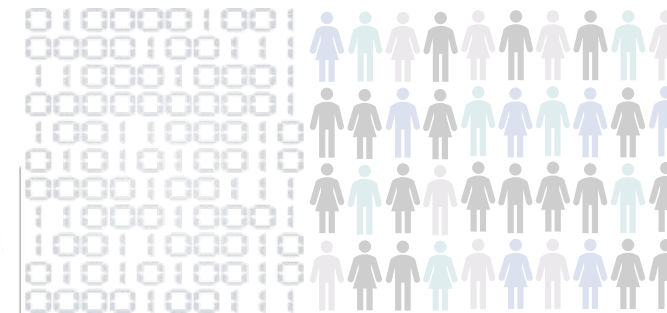
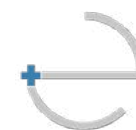
IV) Explicabilidad

Obligación de trazabilidad y auditabilidad

Participación democrática



# CONSIDERACIONES ÉTICAS ESPECÍFICAS DE LA **IA PARA LA SEGURIDAD CIUDADANA**



# 2.1. Problemáticas específicas del uso de la IA en el ámbito policial



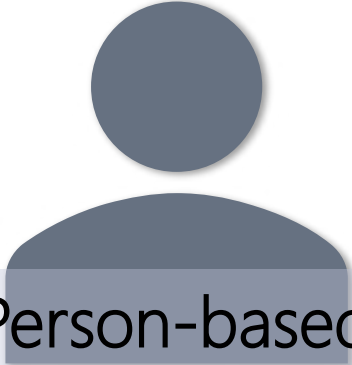
Predictive Policing



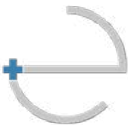
Predictive Sentencing



Place-based

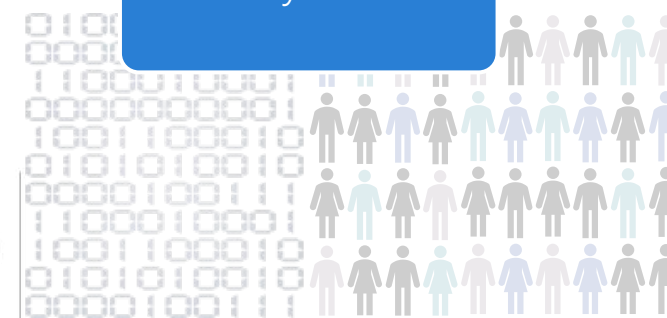
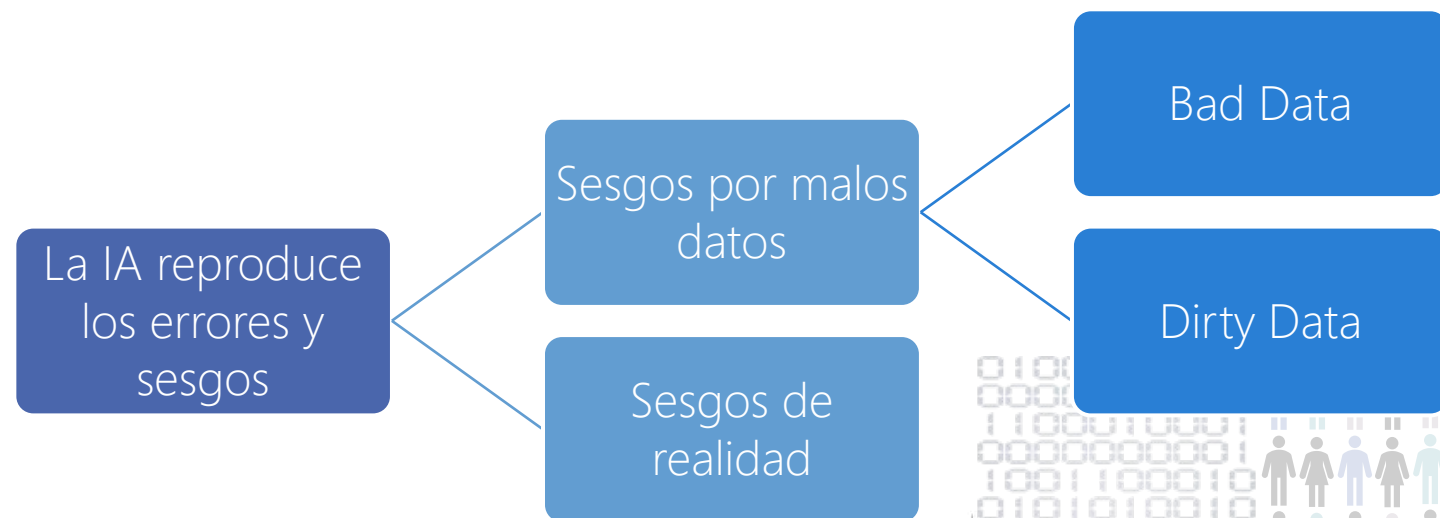


Person-based





| The biased reality, the bad and the dirty data





### CAUTION: BAD DATA



BAD DATA QUALITY  
MAY RESULT IN  
FRUSTRATION AND  
LEAD TO DROP  
KICKING YOUR  
COMPUTER

# VALIDEZ vs BAD DATA

---

¿QUÉ PASA CUANDO TENGO  
UN DATO MALO, UNA  
CORRELACIÓN ESPÚREA, UNA  
MUESTRA NO  
REPRESENTATIVA?



# IGUALDAD vs DIRTY DATA

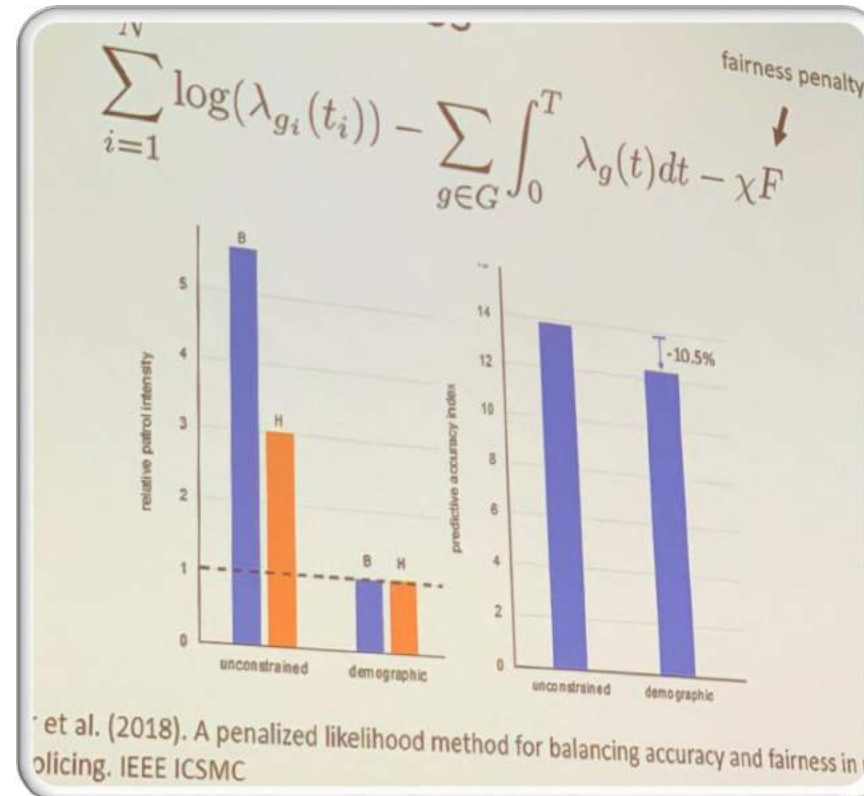
---

- ES MALO NO PORQUE CIENTÍFICAMENTE ESTÉ MAL CONSTRUIDO, SINO PORQUE EL DATO DE LA REALIDAD ESTÁ SESGADO.



# IGUALDAD vs BIASED REALITY

- SE SABE QUE HAY CARACTERÍSTICAS QUE INFLUYEN DECISIVAMENTE EN LA DELINCUENCIA ¿QUITAMOS LAS VARIABLES EDAD Y SEXO?



# 2.2. Principios éticos para el uso de IA en el ámbito policial y de justicia

## LOS 3 (X2) PRINCIPIOS | MIRÓ-LLINARES (2020)



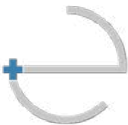
Validez (+ ajuste)



Equidad (+ flexibilidad)



Transparencia (+ trazabilidad)



# Reflexión ética desde el diseño



1. ¿CON QUÉ DATOS?



2. ¿PARA QUÉ USARLA?



3. ¿CÓMO USARLA?

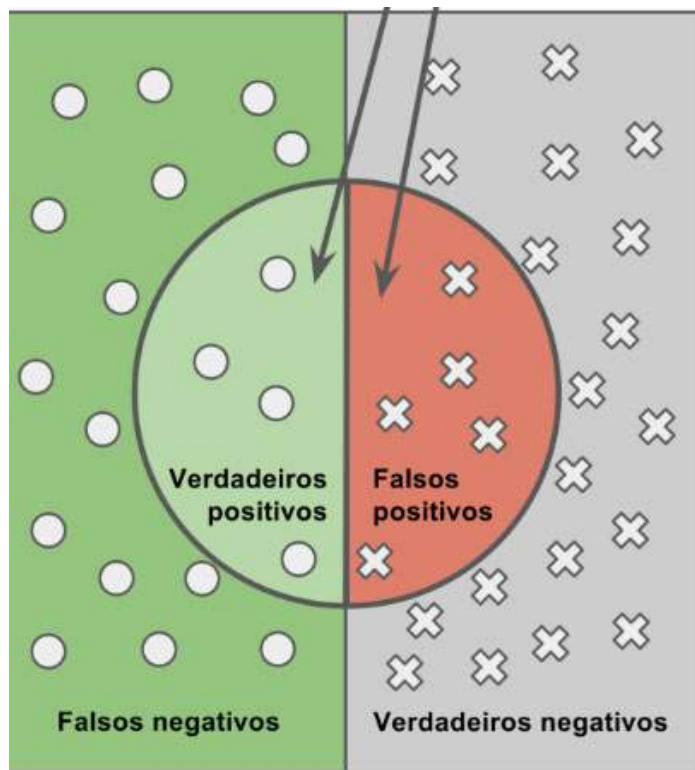


# 1. ¿CON QUÉ DATOS?



- Datos bien etiquetados
- Muestras representativas
- Correlaciones fuertes
- Que tengan en cuenta la multicausalidad
- Que incluyan evidencias relativas a los procesos de medición, validez, dependencias entre los elementos de los datos
- Método científico





## 2. PARA QUÉ USARLA

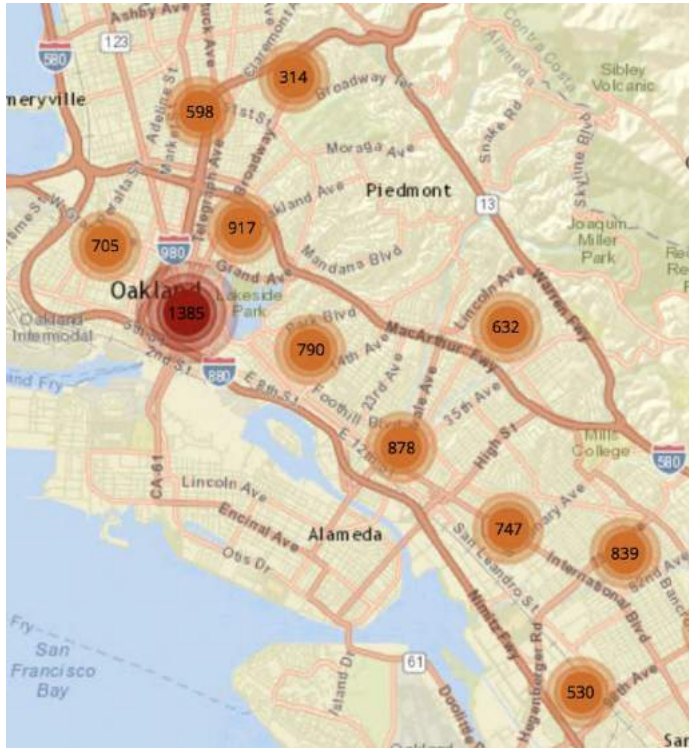
### | ANTE QUÉ ESCENARIOS Y CON QUÉ CONSECUENCIAS

- ¿Ante qué escenarios es coherente usar una IA?
  - Debemos tener en cuenta cómo sufre la capacidad predictiva de la IA en relación con la multicausalidad de la delincuencia
- ¿Cuáles son las consecuencias?
  - Atribución de un riesgo de violencia o reincidencia erróneo
  - Toma de decisiones influenciada por este error



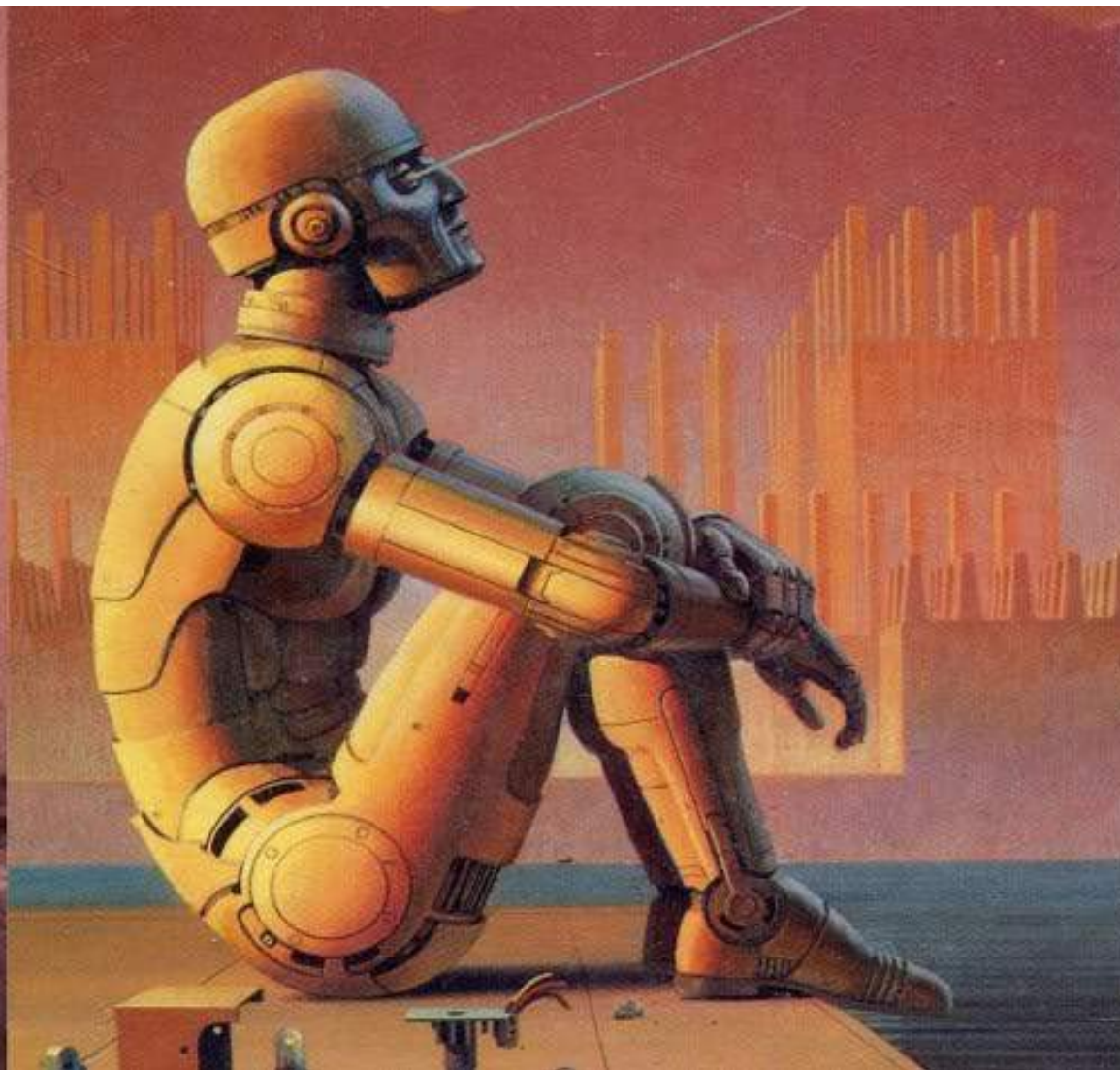
### 3. CÓMO USARLA

## | ¿ES SUFICIENTE CON IDENTIFICAR EL RIESGO?



- Existe una confusión a la hora de distinguir entre riesgo y remedio
- Los algoritmos predictivos permiten identificar lugares y personas con riesgo
- Saber en qué lugar o qué persona tiene un riesgo asociado elevado no informa de cómo afrontar el problema
- Este desconocimiento puede provocar la implementación de soluciones equivocadas
- Es necesario que los encargados de desarrollar estas herramientas sean equipos interdisciplinarios





CRIMINALAI



**GRACIAS**

F.MIRO@CRIMINA.ES

@FERNANDOQPH

CRÍMINA



P L U S  
ethics