



Curso DM

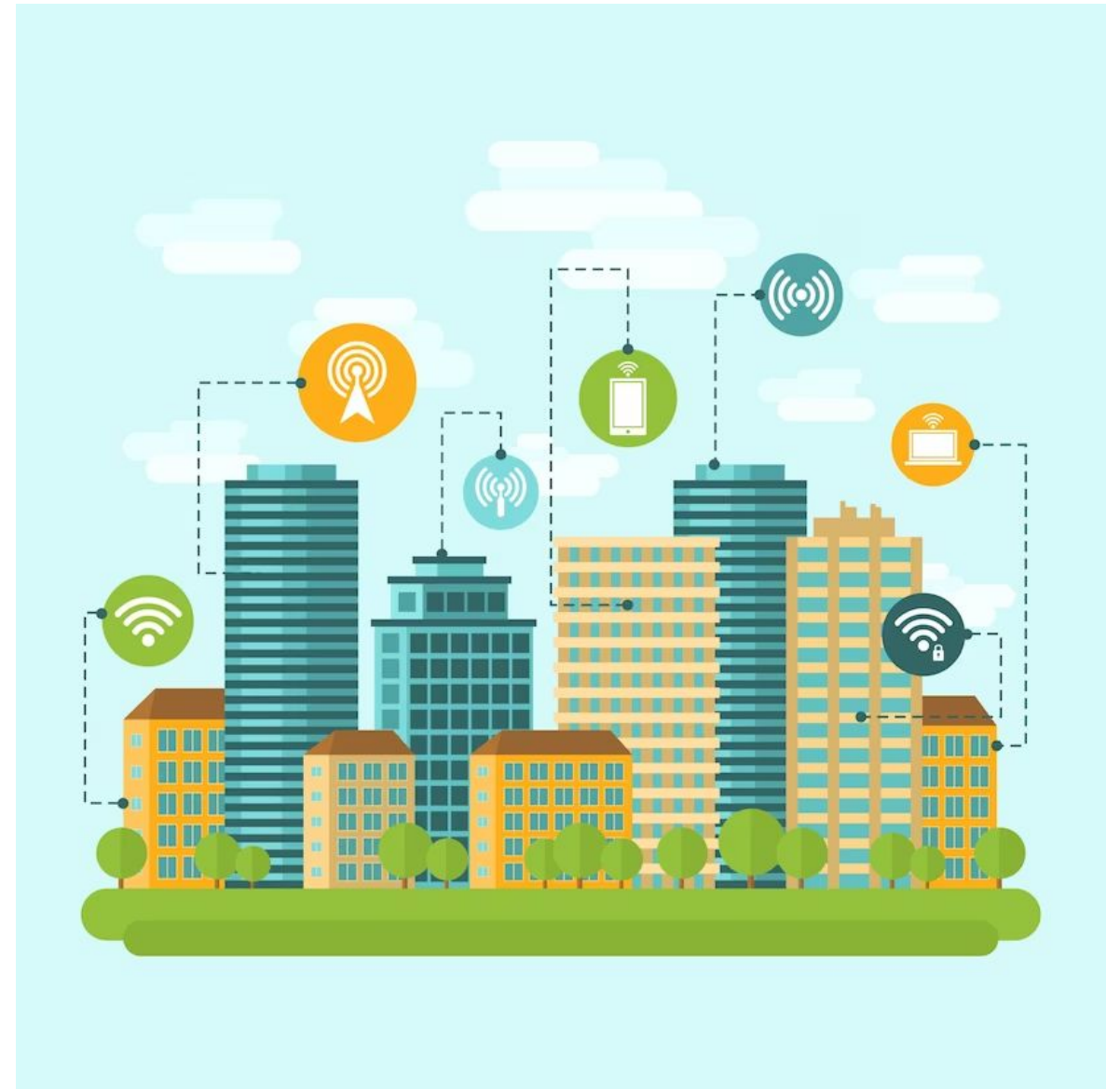
“Ética en Datos”

Primavera 2023

Jazmine Maldonado y Cinthia Sánchez

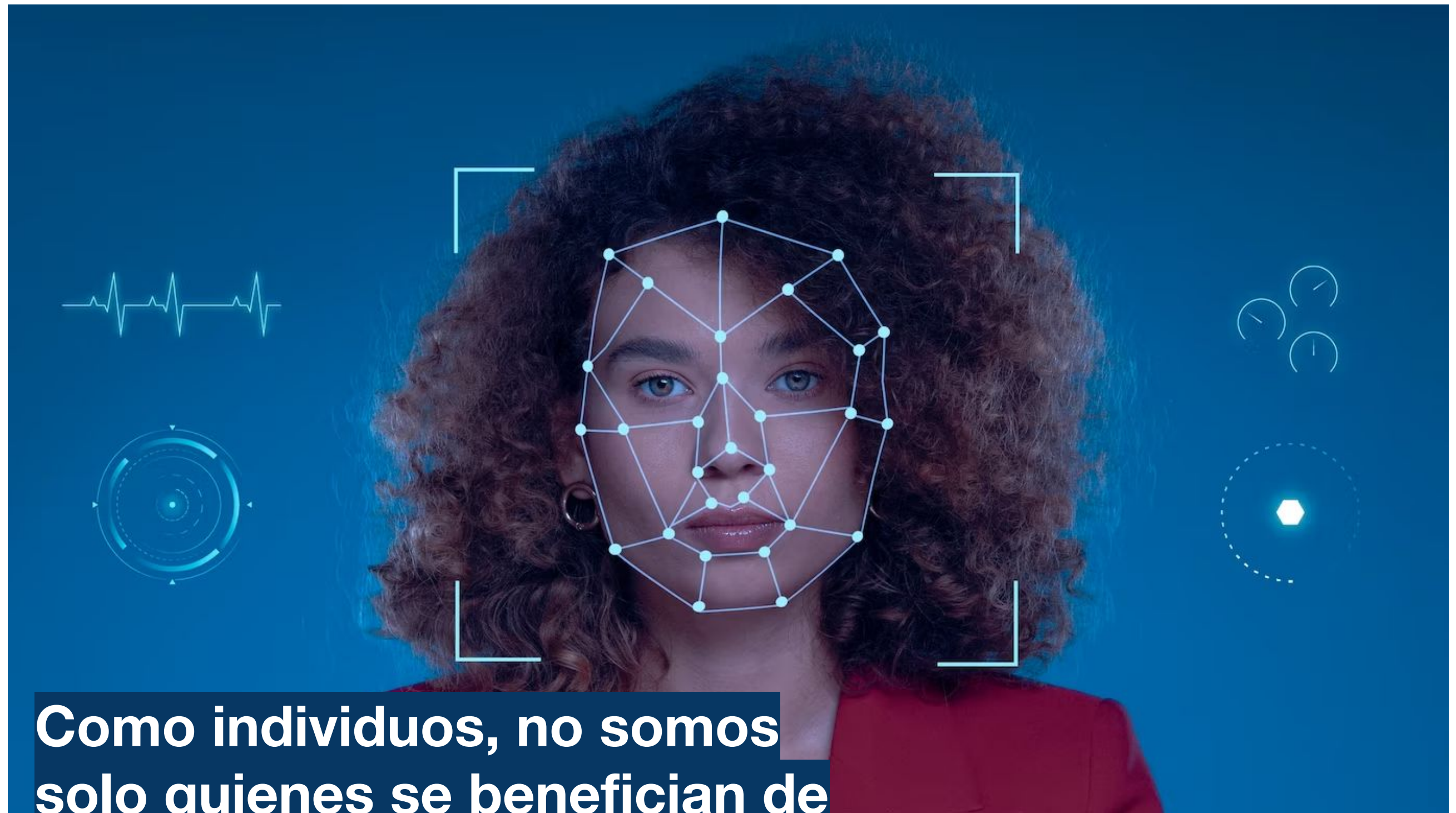
Actualmente se genera una cantidad sin precedentes de información día a día.

Es relativamente fácil obtener acceso.



Los descubrimientos
que podemos hacer a
partir la gran cantidad
de datos que existe hoy
pueden ser fabulosos





**Como individuos, no somos
solo quienes se benefician de
los análisis generados por
estos datos....**

nosotros somos los datos



Los datos se ocupan para tomar decisiones sobre nosotros.

Algunas de ellas con importantes consecuencias.

Caso: Datos anónimos

New York Times obtuvo un conjunto de datos comerciales.

Contenía información de ubicación recopilada de apps móviles.

Datos superficialmente anónimos.



El New York Times
pudo identificar a una
profesora de
matemáticas de 46
años llamada Lisa
Magrin.

Era la única persona
que viajaba 14 millas
de distancia
diariamente.



El Times siguió el rastro de Lisa a la consulta de un dermatólogo y a la casa de su exnovio.

<https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html>



Caso: Herramientas policiales predictivas

Modelos basados en la localidad para predecir cuándo y dónde podría ocurrir un crimen.

Modelos basados en la persona para predecir participación futura en un crimen.



Los sesgos y discriminación en el mundo físico se amplifican en modelos predictivos

<https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/>



Caso: embeddings

Modelos de lenguaje son entrenados con millones de documentos en la web.

Nos permiten transformar palabras en vectores numéricos de dimensión fija



Nos permiten comparar palabras y detectar similitudes semánticas.

También nos permiten realizar analogías entre palabras.

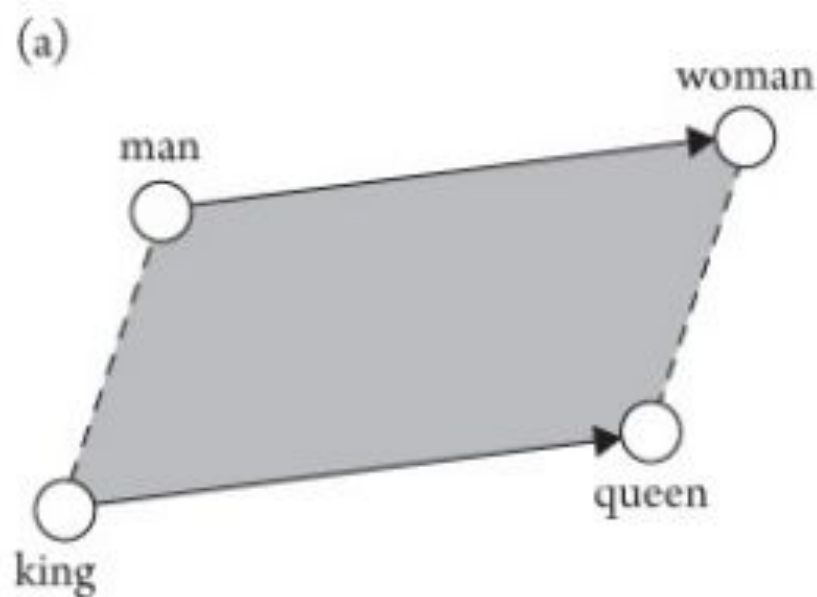


Fig. 7. Word analogies as parallelograms.



Tokio -> Japan

Paris -> France

man -> computer
programmer

women -> homemaker

[https://proceedings.neurips
.cc/paper_files/paper/2016/
file/a486cd07e4ac3d27057
1622f4f316ec5-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2016/file/a486cd07e4ac3d270571622f4f316ec5-Paper.pdf)



Ética

La ética, o filosofía moral, es la rama de la filosofía que estudia la conducta humana, lo correcto y lo incorrecto, lo bueno y lo malo, la moral, el buen vivir, la virtud, la felicidad y el deber.

- Se remonta a los griegos (Platón 427 A.C.)
- Permite crear un orden de convivencia humano y social.

Ética en Datos

Rama emergente de la ética aplicada

Objetivo: Promover el uso responsable y sostenible de los datos en beneficio de las personas y la sociedad y garantizar que el conocimiento adquirido a partir de los datos no se utilice en contra de intereses legítimos de un individuo o grupo.

(Source: Luciano Floridi, Mariarosaria Taddeo (2016) 'What is data ethics?'; Pernille Tranberg, Gry Hasselbalch, Birgitte Kofod Olsen & Catrine Søndergaard Byrne (2018) 'Data Ethics. Principles and Guidelines for Companies, Authorities & Organisations')

Datos Personales

Los datos personales son la información que permite identificar de una manera directa o indirecta a una persona.

Categorías de datos personales

Datos personales generales

Cualquier tipo de información o identificadores que sirvan para identificar a una persona física

Datos personales sensibles

Aquellos que afectan la intimidad del titular o cuyo uso indebido puede generar su discriminación.

Categorías de datos personales

**Ojo: Son ejemplos,
pueden ser más.**

Datos personales generales

Edad
Teléfono
Domicilio
Correo electrónico
Documento nacional
de identidad
Número de seguridad
social
Ingresos
Fecha de nacimiento

Datos personales sensibles

Origen étnico o racial
Opiniones políticas
Convicciones Religiosas
Afiliación sindical
Genéticos
Biométricos
Relativos a la salud
Relativos a la vida sexual o
la orientación sexual

Derechos

A

R

C

O

Acceso

Tienes derecho a ser informado acerca de la inclusión de tus datos en un banco de datos, y a toda información relativa a ella, como su procedencia y destinatario, el propósito del almacenamiento y la individualización de las personas u organismos a los cuales tus datos son o serán transmitidos regularmente.

Rectificación

En caso de que tus datos personales sean erróneos, inexactos, equívocos o incompletos, y así se acredite, tienes derecho a que se modifiquen.

Cancelación

Sin perjuicio de las excepciones legales, podrás, además, exigir que se eliminen tus datos personales, en caso de que su almacenamiento carezca de fundamento legal o cuando estuvieren caducos.

Oposición

Podrás oponerte al tratamiento de tus datos, o pedir que éste cese cuando se haya iniciado sin tu consentimiento, salvo excepciones legales. En particular, puedes oponerte a la utilización de tus datos personales con fines de publicidad, investigación de mercado o encuestas de opinión.

Principios de Protección de Datos Personales

Legalidad, licitud y legitimidad

Consentimiento

Transparencia

Finalidad

Lealtad

Principios de Protección de Datos Personales

Proporcionalidad

Minimización

Calidad

Responsabilidad

Seguridad

Anonimización de Datos

Proceso de eliminar información de identificación personal de los conjuntos de datos, de modo que las personas a quienes describen los datos permanezcan en el anonimato.

Su intención es la protección de la privacidad.

Caso Netflix

En 2006 Netflix lanzó el Netflix Prize Competition

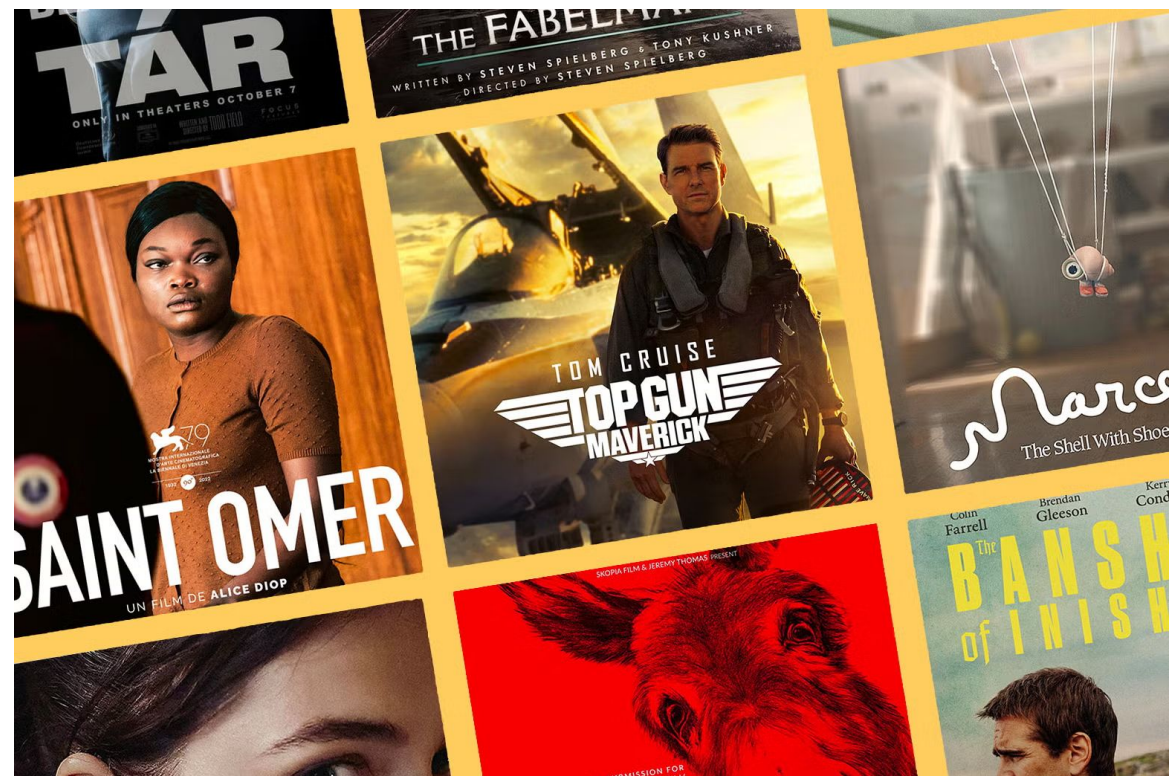
Competencia data science para desarrollar el mejor algoritmo de filtro colaborativo para potenciar su motor de recomendación.



Caso Netflix

Publicó datos de ratings de aproximadamente 18000 películas.

Netflix era consciente de la importancia de la protección de los datos, por lo que incluyó sólo ratings (1-5) y la fecha.



Caso Netflix

2 semanas después Arvind Narayanan (estudiante PhD de la universidad de Texas Austin) y su supervisor, Vitaly Shmatikov, publicaron un paper en el que decían que podrían identificar los nombres reales de muchos de los registros.



Caso Netflix

En el paper explicaban que si el atacante sabía las fechas aproximadas en que una persona había calificado 6 películas, podía identificar con un 99% de probabilidad a la persona.

Para ello cruzaba datos con información obtenida desde la plataforma IMDB.



Caso Netflix

Bueno... pero si publicó su calificación en IMDB,
¿Por qué importa que se revelen sus valoraciones
en Netflix?

Caso Netflix

La persona decidió publicar algunas valoraciones en IMBD, pero en el dataset de Netflix pueden haber muchas más.

Las valoraciones privadas de Netflix podrían revelar información como intereses políticos u orientación sexual.

De hecho, una mujer gay madre de dos niños que no había dado a conocer su orientación sexual demandó a Netflix.

Privacidad Diferencial

Objetivo: Otorgar a las personas incluidas en una base de datos el mismo grado de privacidad que si sus datos se excluyeran por completo. Es decir, garantiza la preservación de la privacidad de los sujetos

Privacidad Diferencial

Se logra a través de una variedad de técnicas estadísticas.

En esencia, agrega una cantidad calculada de ruido (datos aleatorios) a la base de datos de manera controlada, para que sigan siendo útiles para realizar análisis o entrenar modelos.

Esto oscurece la relación entre el individuo y los puntos de datos.

Sesgo (Bias)

Error sistemático o desviación del valor real o deseado de un atributo de datos, medida o modelo.

Puede surgir en diferentes etapas:

- Recopilación de datos
- Preprocesamiento
- Análisis
- Interpretación

Sesgo (Bias)

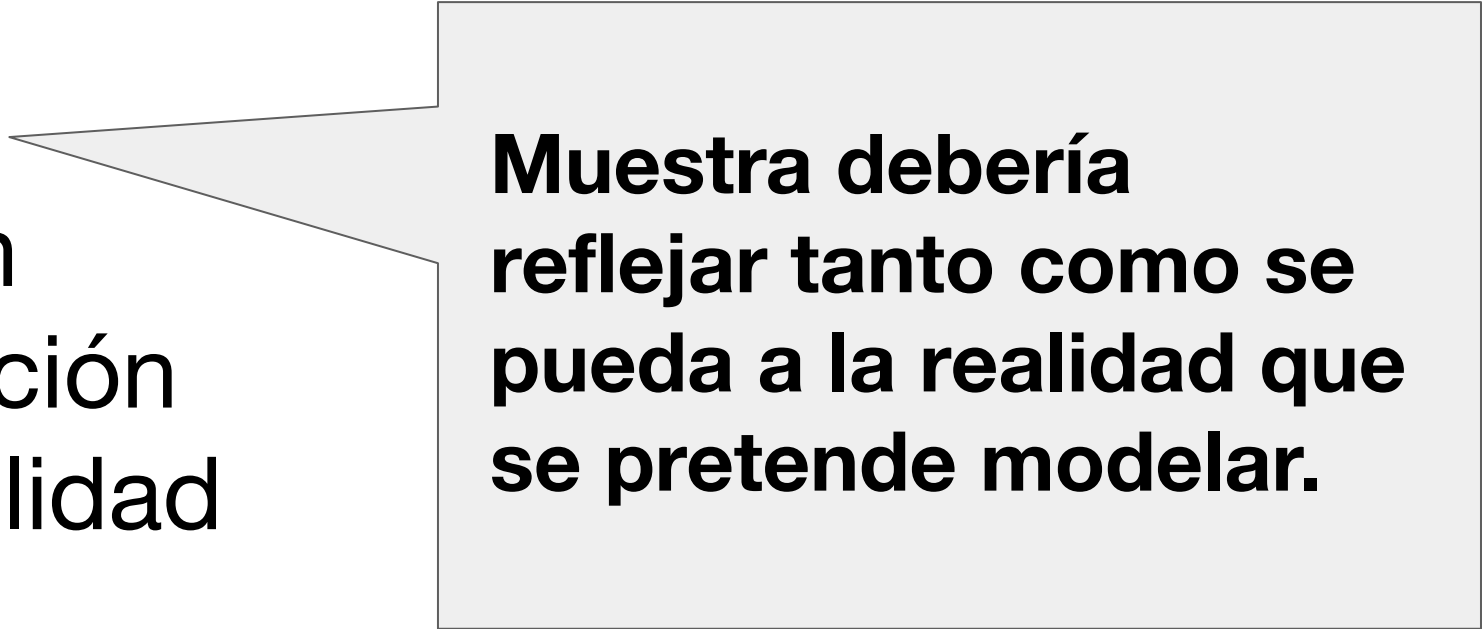
Efectos:

- Poca validez y confiabilidad de los resultados
- Bajo rendimiento y precisión de los modelos
- Pérdida de significado y valor de los resultados
- Implicaciones éticas y sociales

Sesgo (Bias)

Tipos de sesgos durante la selección y preprocesamiento de datos:

- Sesgo de muestra
- Sesgo de selección
- Sesgo de confirmación
- Sesgo de disponibilidad



Muestra debería reflejar tanto como se pueda a la realidad que se pretende modelar.

Sesgo (Bias)

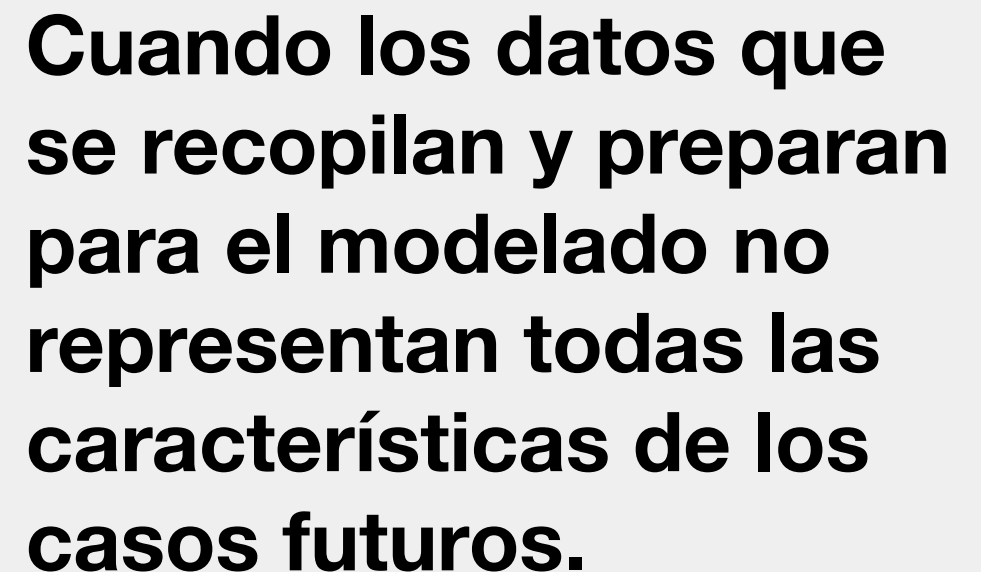


Ej. Detección patrones de movilidad urbana en santiago

Sesgo (Bias)

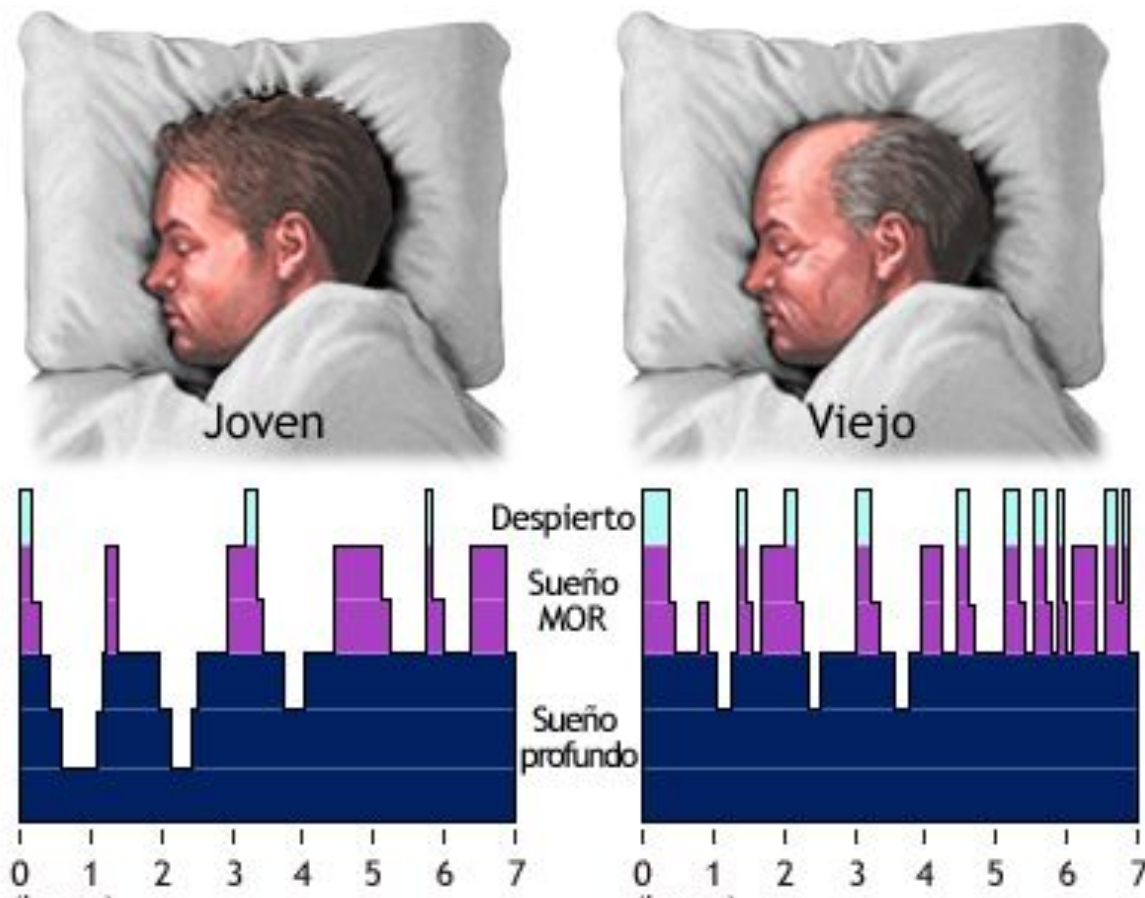
Tipos de sesgos durante la selección y preprocesamiento de datos:

- Sesgo de muestra
- Sesgo de selección
- Sesgo de confirmación
- Sesgo de disponibilidad



Cuando los datos que se recopilan y preparan para el modelado no representan todas las características de los casos futuros.

Sesgo (Bias)

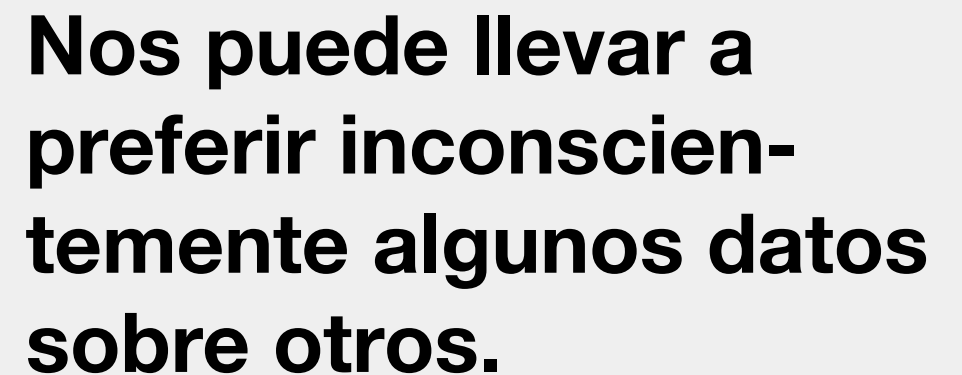


Ej. Caracterización de los patrones de sueño

Sesgo (Bias)

Tipos de sesgos durante la selección y preprocesamiento de datos:

- Sesgo de muestra
- Sesgo de selección
- Sesgo de confirmación
- Sesgo de disponibilidad

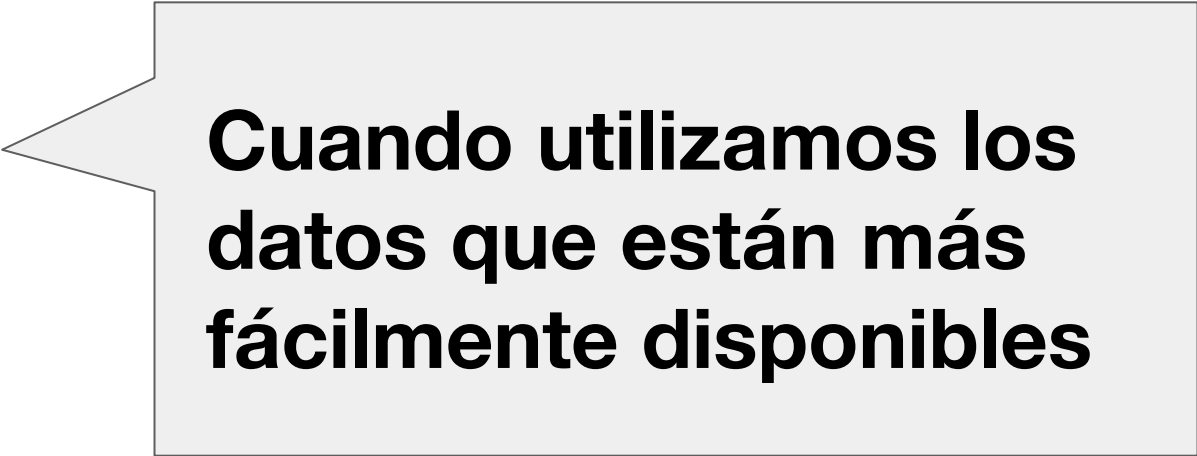


Nos puede llevar a preferir inconscientemente algunos datos sobre otros.

Sesgo (Bias)

Tipos de sesgos durante la selección y preprocesamiento de datos:

- Sesgo de muestra
- Sesgo de selección
- Sesgo de confirmación
- Sesgo de disponibilidad



Cuando utilizamos los datos que están más fácilmente disponibles

Sesgo (Bias)



Ej. Estudio de opinión pública en base a redes sociales

Sesgo (Bias)



Ej. Sistema COMPAS para predecir la probabilidad de volver a delinquir

Sesgo (Bias)



Ej. Algoritmos de contratación y reclutamiento

The Ethical Algorithm: The Science of Socially Aware Algorithm Design. Michael Kearns y Aaron Roth

Datasheets for Datasets. Timnit Gebru et. al.

Iniciativa FairLac <https://fairlac.iadb.org/>

AEDIA - Asociación de Ética en Datos e Inteligencia Artificial conformada recientemente por estudiante de Beauchef.



dcc

CIENCIAS DE LA COMPUTACIÓN
UNIVERSIDAD DE CHILE

www.dcc.uchile.cl

f @ in  / DCCUCHILE