

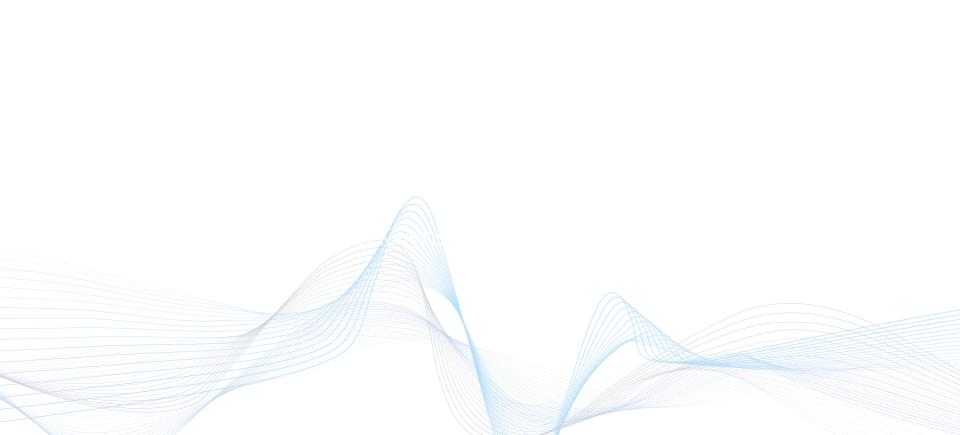
Inteligencia artificial y salud. Cuestiones éticas

David Casacuberta

Dpt de Filosofia. Universidad Autónoma de Barcelona

INTELIGENCIA ARTIFICIAL





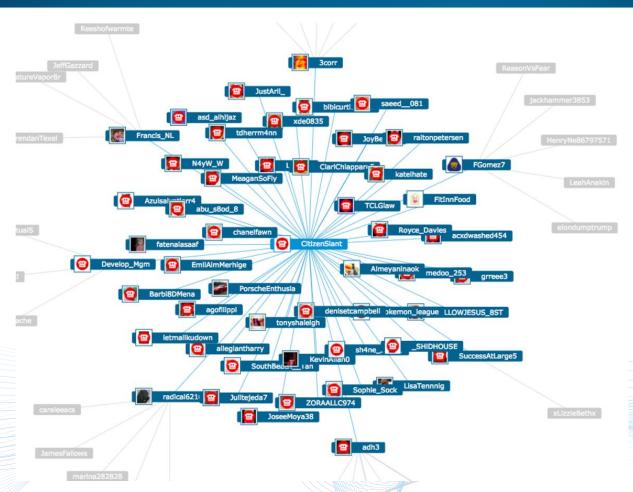






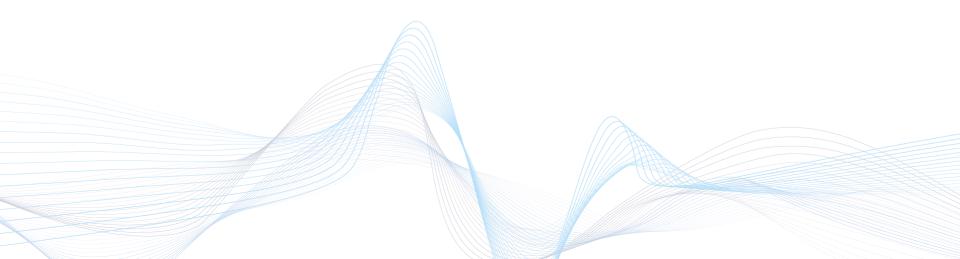






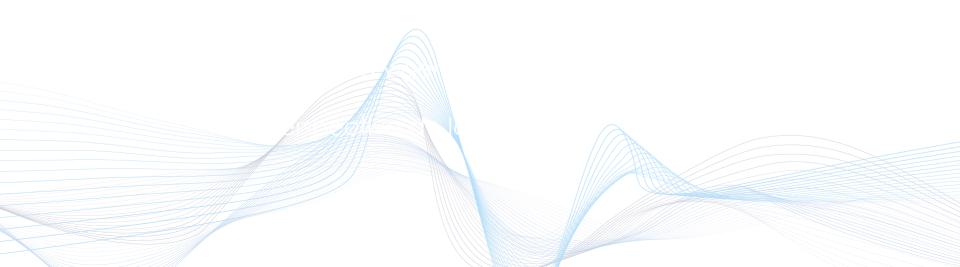


Ética versus legalidad



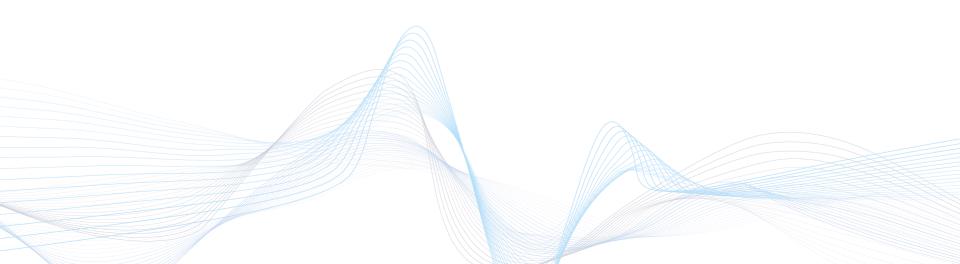


Contenidos de esta clase





No podemos reducir la ética a:

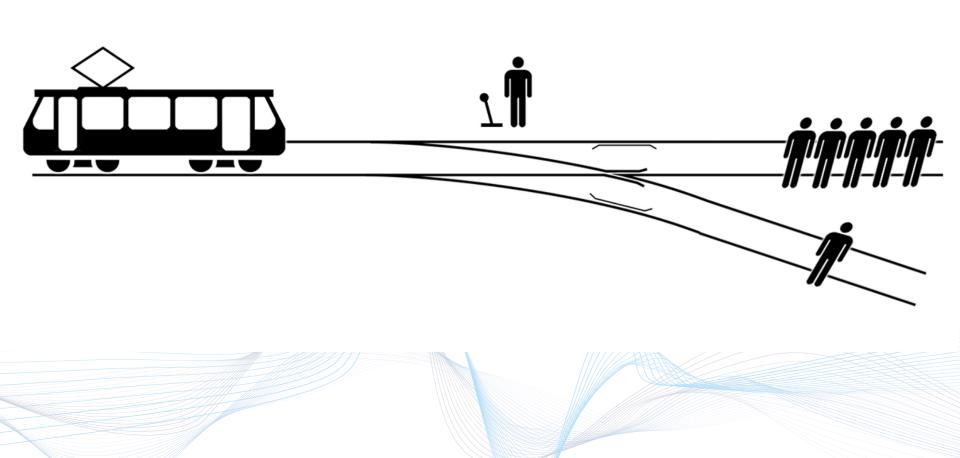




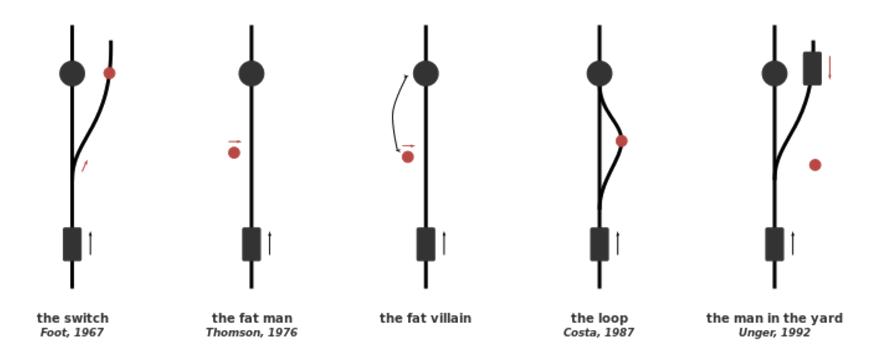
El experimento mental de la vagoneta

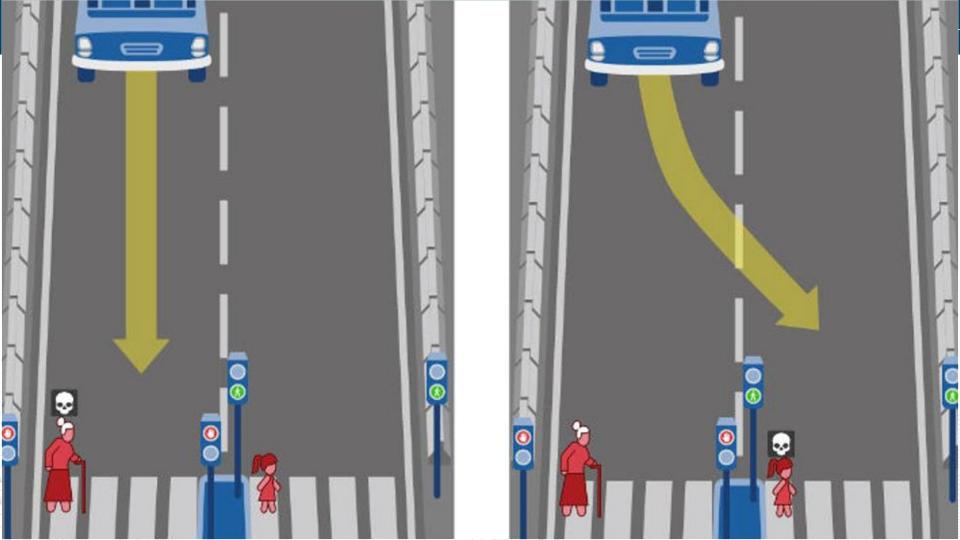






Variantes



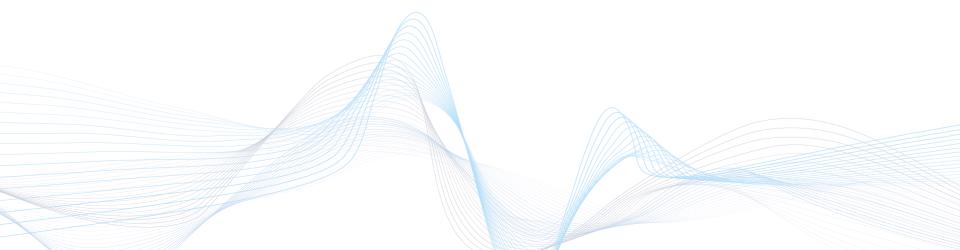




valores



conflicto de valores



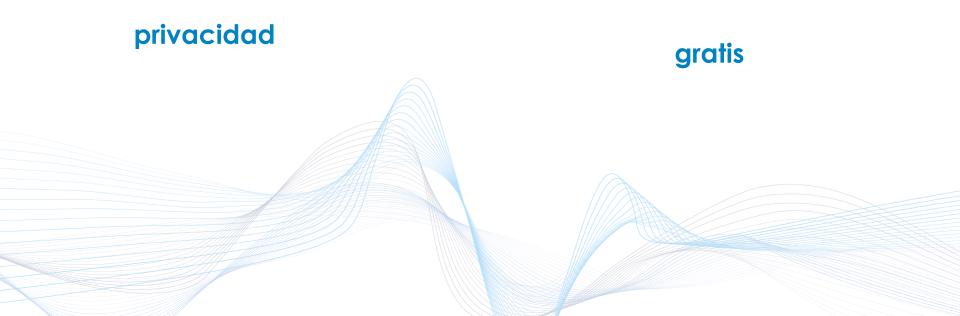


privacidad











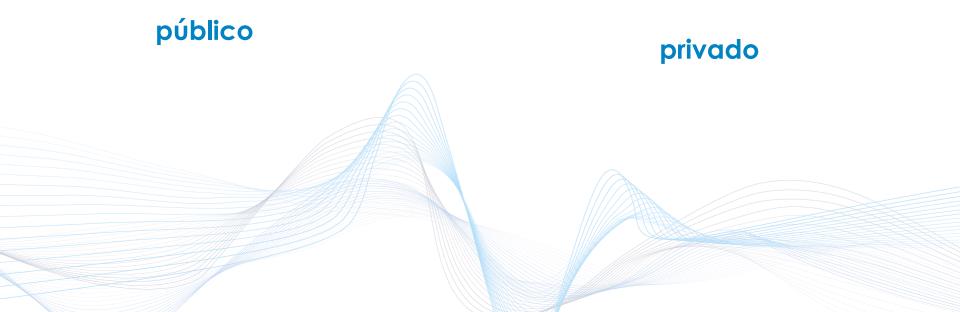






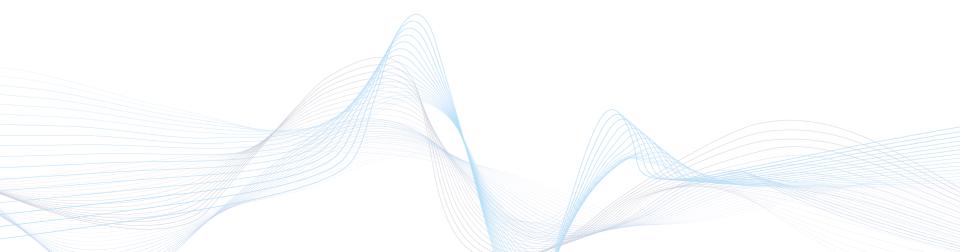
propiedad intelectual

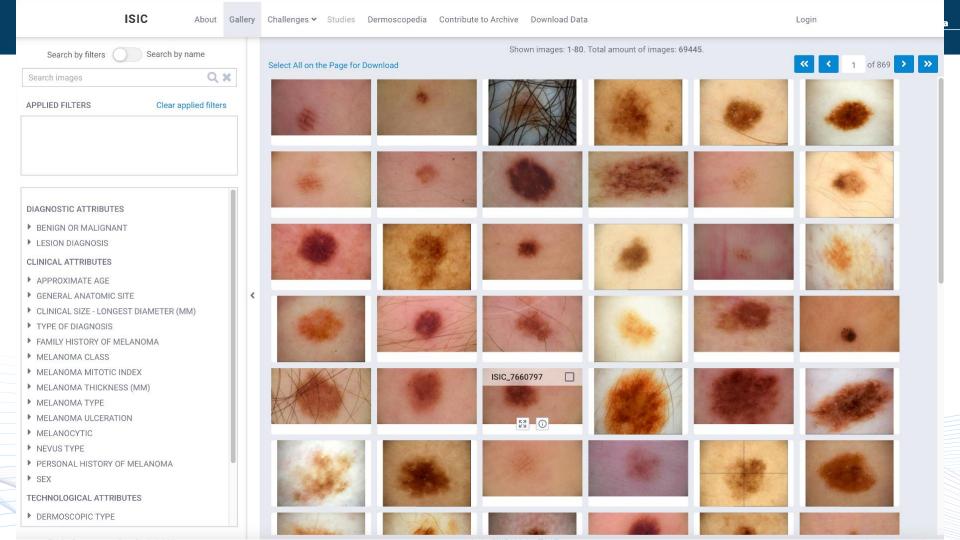






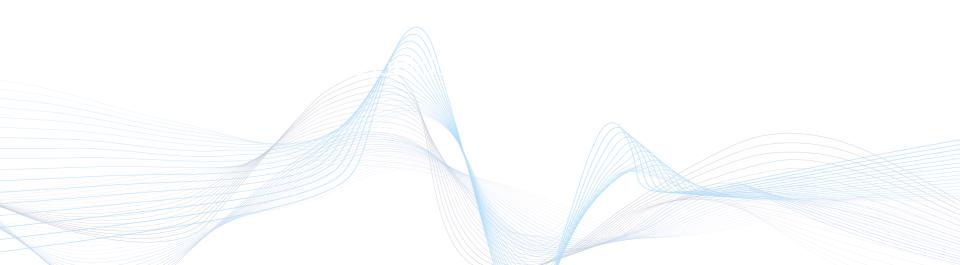
Dos historias para pensar cuestiones éticas y legales



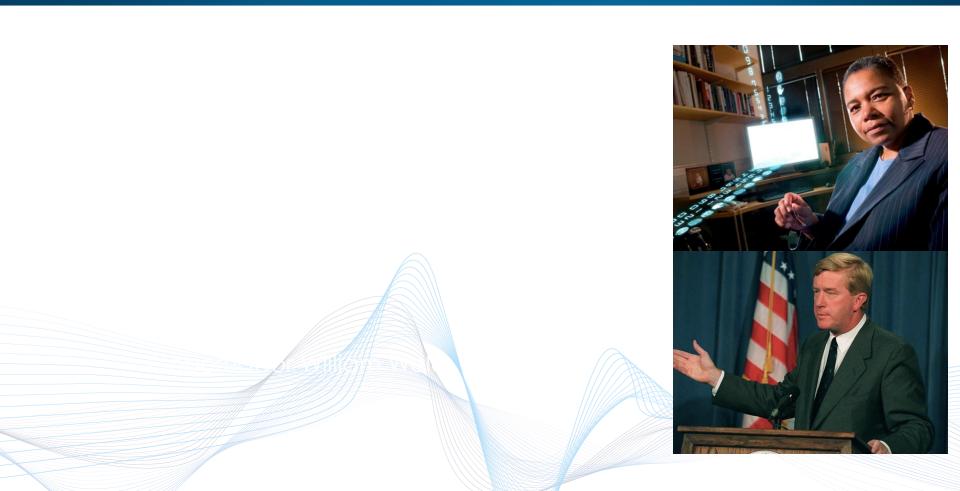




La ética como caja de herramientas



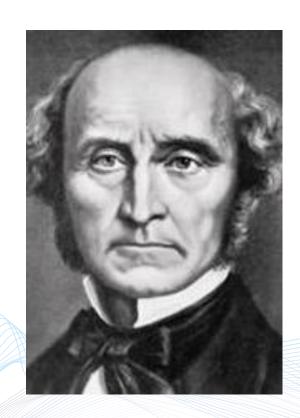






Utilitarismo

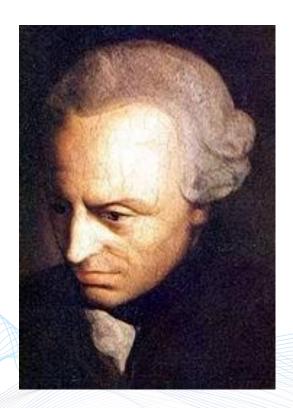
- John Stuart MIII (1806-1873)
- Ética de consecuencias
- La mayor felicidad para el mayor número de personas





Ética de principios

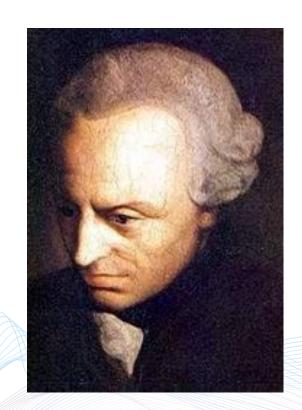
- Immanuel Kant (1724-1804)
- Ética como imperativos categóricos. No negociables.





Ética de principios

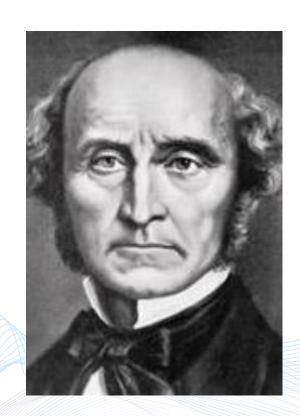
- Actúa como si tu comportamiento pudiera convertirse en ley universal
- Trata a los demás como fines y no solo como medios





Utilitarismo

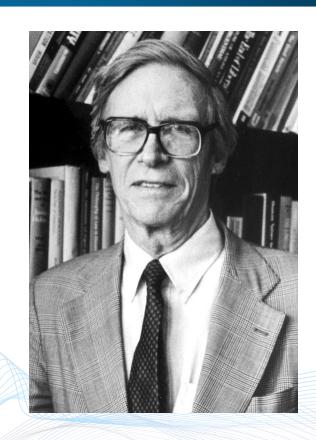
- John Stuart MIII (1806-1873)
- Ética de consecuencias
- La mayor felicidad para el mayor número de personas





Teoría de la justicia

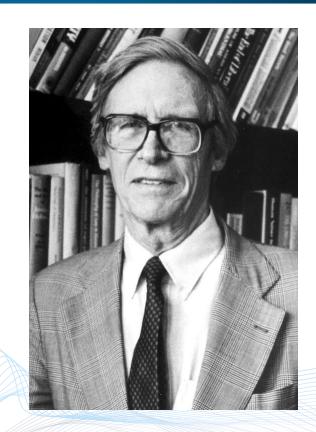
- John Rawls(1921-2002)
- Justicia como equidad
- Justice is the first virtue of social institutions, as truth is of systems of thought. A theory however elegant and economical





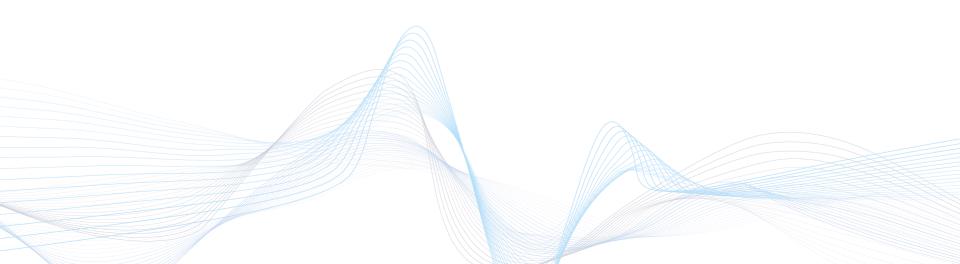
Teoría de la justicia

Justice is the first virtue of social institutions, as truth is of systems of thought. A theory however elegant and economical must be rejected or revised if it is untrue; likewise laws and institutions no matter how efficient and well-arranged must be reformed or abolished if they are unjust





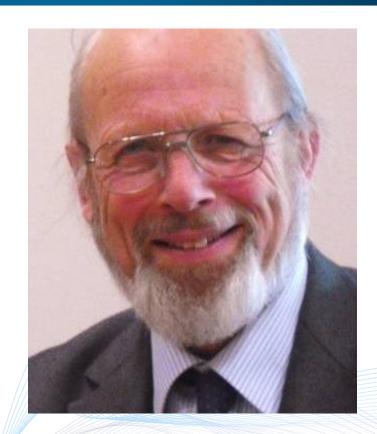
Matriz ética





Ben Mephan

 Desarrolla la herramienta originalmente para facilitar las discusiones sobre ética de los productos GMO

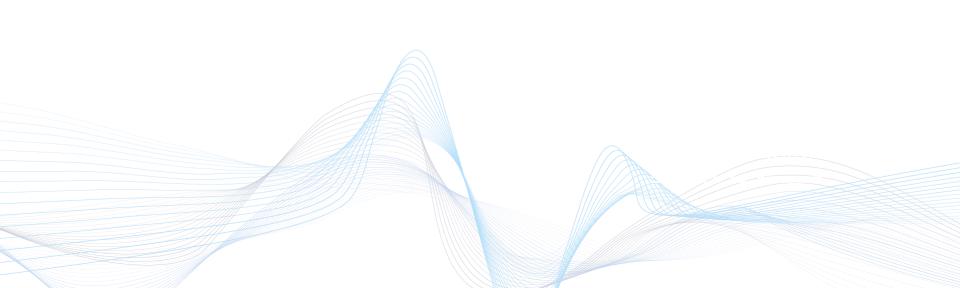


The matrix

Respeto por	Bienestar	Autonomía	Justicia
Granjeros	Salario y condiciones laborales correctas	Libertad de acción en la gestión del negocio	Comercio justo y competitivo
Consumidores	Seguridad alimentaria, calidad de vida	Poder hacer una decisión informada	Acceso a comida a precios adecuados
Vacas de granja	Bienestar del animal	Libertad de movimientos	Valor intrínseco
Entorno	Conservación	Biodiversidad	Sostenibilidad

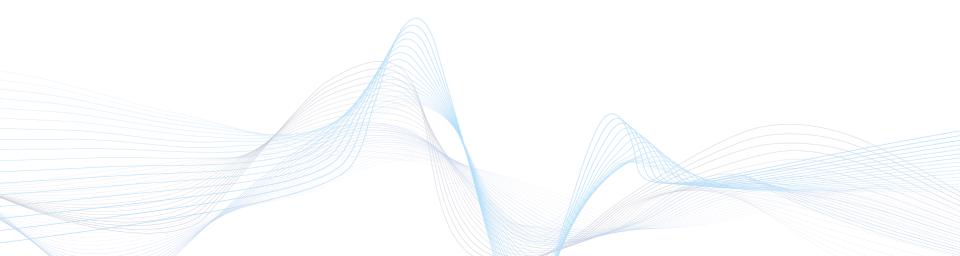


Filas



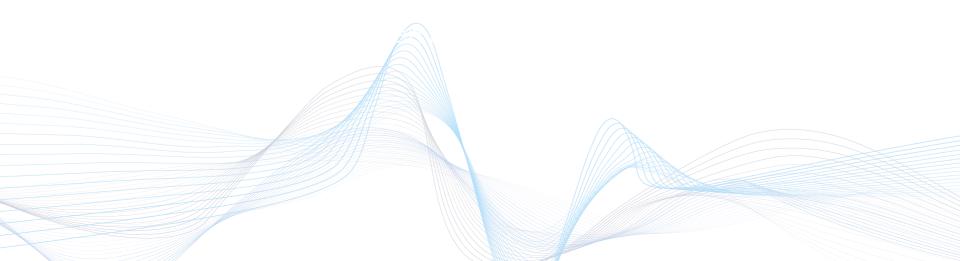


Columnas





Forma de trabajo





Correctional Offender Management Profiling for Alternative Sanctions (COMPAS)











	Bienestar	Autonomía	Justicia
Tribunal	Eficiencia Consistencia	Transparencia Libertad de adoptar o no el software	Eficiencia Falsos negativos Falsos positivos
Acusados de raza negra	Maximizar tratamiento y rehabilitación. Minimizar confinamiento y castigo	Falsos positivos Transparencia	Respeto por derechos civiles datos correctos
Acusados de raza blanca	Maximizar tratamiento y rehabilitación. Minimizar confinamiento y castigo	Transparencia	Respeto por derechos civiles datos correctos
Público	Estabilidad de la comunidad	Transparencia	Falsos positivos Falsos negativos Justicia igualitaria datos correctos
			Métrica propia de justicia



Artificial intelligence Act



Bruselas, 21.4.2021 COM(2021) 206 final 2021/0106(COD)

Propuesta de

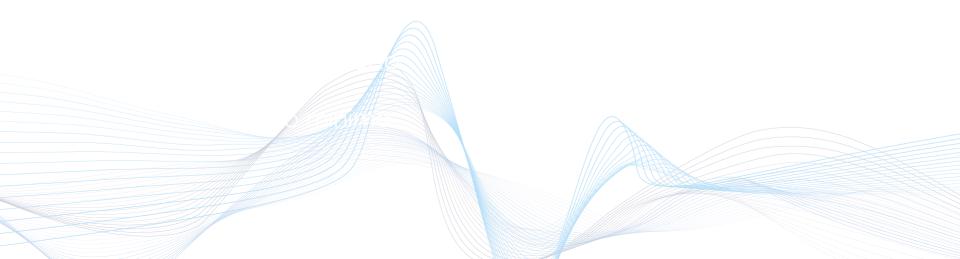
REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO

POR EL QUE SE ESTABLECEN NORMAS ARMONIZADAS EN MATERIA DE INTELIGENCIA ARTIFICIAL (LEY DE INTELIGENCIA ARTIFICIAL) Y SE MODIFICAN DETERMINADOS ACTOS LEGISLATIVOS DE LA UNIÓN

{SEC(2021) 167 final} - {SWD(2021) 84 final} - {SWD(2021) 85 final}

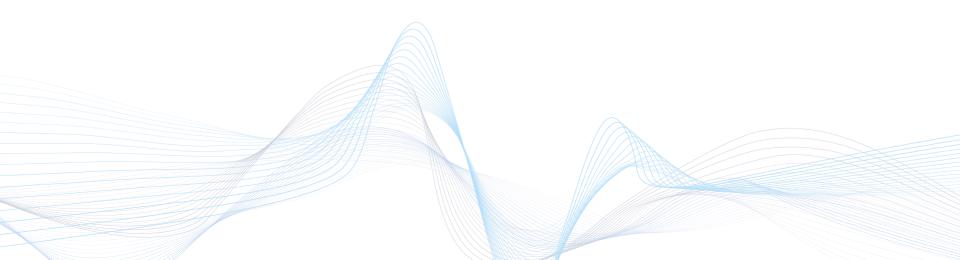


Cuatro tipos de riesgo





Riesgo inaceptable





Riesgo inaceptable. Puntaje social





Riesgo inaceptable. Cambridge Analytica





















SEGMENTACIÓN	SCORE	
AAA - Cliente Excelente	930 – 999 puntos	
AA - Cliente Bueno	850 – 929	
A - Cliente Regular	700 – 849	
B - Cliente Bajo	400 – 699	
C - Cliente Malo	Bajo 400	



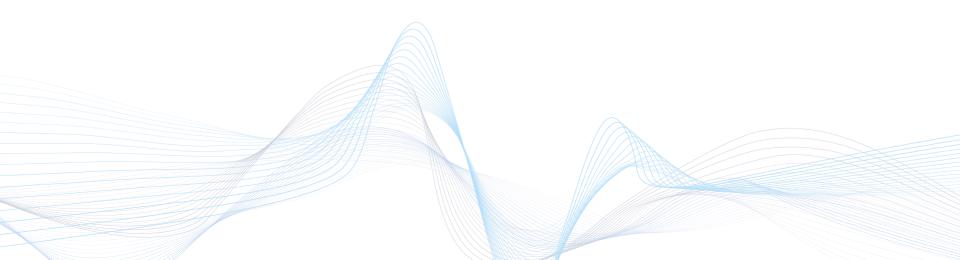






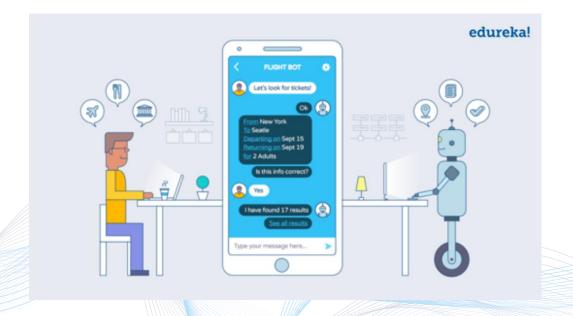








Riesgo limitado





Riesgo mínimo





Explicabilidad

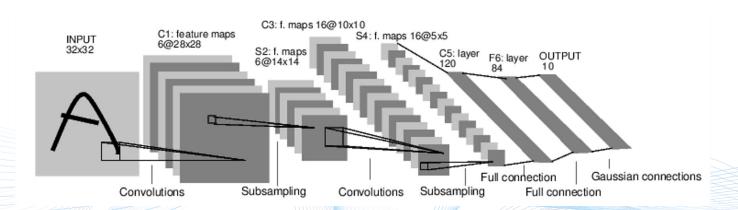


Algoritmos tradicionales



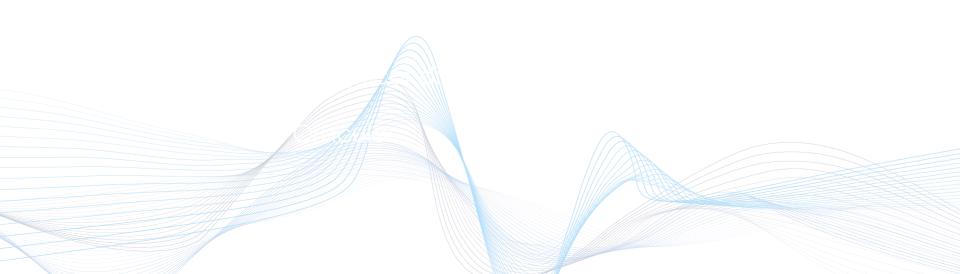


Explicabilidad





Problemas asociados a la falta de explicaciones



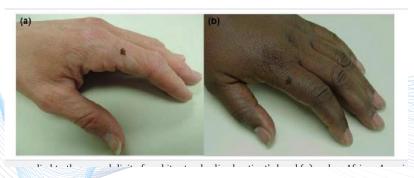


Problemas asociados a la falta de explicaciones











Necesidad de un cambio de paradigma





Gracias por su atención

