# Inteligencia Artificial y Ética

#### Miguel A. Gutiérrez Naranjo

Departamento de Ciencias de la Computación e Inteligencia Artificial Universidad de Sevilla

15 de Junio, 2018



#### Según la Real Academia Española

- ...
- Conjunto de normas morales que rigen la conducta de la persona en cualquier ámbito de la vida.
- Parte de la filosofía que trata del bien y del fundamento de sus valores.



#### Problema abordado durante milenios ... sin solución

- Sócrates, Platón, Aristóteles, Agustín de Hipona, Tomás de Aquino, ...
- David Hume (1711-1776): Sentimiento de simpatía hacia aquello que es útil a la sociedad.
- Friedrich Nietzsche (1844-1900): Si la razón no puede alcanzar la verdad del conocimiento tampoco podrá saber nada respecto del bien.

• . . .



#### Problema abordado durante milenios ... sin solución

- Sócrates, Platón, Aristóteles, Agustín de Hipona, Tomás de Aquino, . . .
- David Hume (1711-1776): Sentimiento de simpatía hacia aquello que es útil a la sociedad.
- Friedrich Nietzsche (1844-1900): Si la razón no puede alcanzar la verdad del conocimiento tampoco podrá saber nada respecto del bien.

• ...



#### Problema abordado durante milenios ... sin solución

- Sócrates, Platón, Aristóteles, Agustín de Hipona, Tomás de Aquino, . . .
- David Hume (1711-1776): Sentimiento de simpatía hacia aquello que es útil a la sociedad.
- Friedrich Nietzsche (1844-1900): Si la razón no puede alcanzar la verdad del conocimiento tampoco podrá saber nada respecto del bien.

• . . .



#### Problema abordado durante milenios ... sin solución

- Sócrates, Platón, Aristóteles, Agustín de Hipona, Tomás de Aquino, . . .
- David Hume (1711-1776): Sentimiento de simpatía hacia aquello que es útil a la sociedad.
- Friedrich Nietzsche (1844-1900): Si la razón no puede alcanzar la verdad del conocimiento tampoco podrá saber nada respecto del bien.
- •



- Deontología: Significa lo conveniente o lo que es necesario
- Puede verse como una rama práctica o aplicada de la ética . . .



• ...o quizá como necesaria.

- El código deontológico es un mecanismo de autorregulación en la que una comunidad profesional acuerda una serie de normas que definen los límites que no se deben sobrepasar en el ejercicio de la profesión.
- ¿Es una limitación en mi libertad en el ejercicio de mi profesión . . . ?
- ... o es un guía a la que acogerme en decisiones difíciles?



- El código deontológico es un mecanismo de autorregulación en la que una comunidad profesional acuerda una serie de normas que definen los límites que no se deben sobrepasar en el ejercicio de la profesión.
- ¿Es una *limitación* en mi libertad en el ejercicio de mi profesión ...?
- ...o es un guía a la que acogerme en decisiones difíciles?



- El código deontológico es un mecanismo de autorregulación en la que una comunidad profesional acuerda una serie de normas que definen los límites que no se deben sobrepasar en el ejercicio de la profesión.
- ¿Es una *limitación* en mi libertad en el ejercicio de mi profesión ...?
- ... o es un guía a la que acogerme en decisiones difíciles?



- Abogados, psicólogos, periodistas, médicos, . . .
- ¿Tiene sentido algún mecanismo de autorregulación para los profesionales que nos dedicamos a la Inteligencia Artificial?



# Autorregulación

- ¿Es necesaria una autorregulación ...?
- ¿...o al menos una reflexión para saber hacia dónde vamos?
- ¿Alguien ve algún motivo para dejar de correr y mirar hacia dónde estamos corriendo?



# Inteligencia Artificial y Ética

- Pero. . .
- ... ¿realmente se presentan conflictos éticos en la vida real en los que hay que tomar decisiones concretas ...



 ... o son sólo elucubraciones metafísicas ajenas a mi vida?

6 de abril, 2018



- Fuente: https://tinyurl.com/y7zvdp5g
- ¿...y si la tecnología que desarrollas (Generative Adversarial Networks, Variational Autoencoders, Capsule Networks, ...) tu empresa decide usarla para fines distintos a los fines que a ti te gustaría?



6 de abril, 2018



- Fuente: https://tinyurl.com/y7zvdp5g
- ¿...y si la tecnología que desarrollas (Generative Adversarial Networks, Variational Autoencoders, Capsule Networks, ...) tu empresa decide usarla para fines distintos a los fines que a ti te gustaría?



16 de Mayo, 2018

portaltic / sector

#### Trabajadores de Google dimiten en protesta por la colaboración de la compañía con el proyecto militar con drones Maven

Publicado 16/05/2018 17:45:41 CET

MADRID, 16 May. (EDIZIONES/Portaltic) -

Más de **4.000 empleados de Google** han pedido a la compañía tecnológica deje de colaborar con el proyecto de monitorización a través de drones militares **Maven**, desarrollado por el Departamento de Defensa de Estados Unidos, y alrededor de **una docena de ellos ha dimitido como protesta**.

- Una docena de trabajadores deciden dejar su trabajo por razones éticas.
- Fuente https://tinyurl.com/ycq2gk52



#### Vehículos autónomos

26 de Septiembre, 2012



# Driverless car bill is signed in California at Google headquarters

A bill to bring driverless cars to roads in California has been signed.

State Governor Jerry Brown backed legislation on Tuesday, and said: "Today we're looking at science-fiction becoming tomorrow's reality".

The bill was signed at the headquarters of Google, which has been testing a fleet of 12 autonomous computer-controlled vehicles for several years.



Governor Brown arrived in one of Google's driverless vehicles

#### Vehículos autónomos

1 de Julio, 2016

INICIO I TECNOLOGÍA I EL COCHE QUE CONQUES SOLO ARRIESGA SU ELTURO TRAS EL PRIMER MUERTO

# El coche que conduce solo arriesga su futuro tras el primer muerto

EEUU investigará 25.000 vehículos después de que un conductor falleciese mientras usaba el piloto automático del automóvil



#### Vehículos autónomos

#### Opinion

**GRAY MATTER** 

# Whose Life Should Your Car Save?

By Azim Shariff, Iyad Rahwan and Jean-François Bonnefon

Nov. 3, 2016











- New York Times, 3 de Noviembre, 2016.
- Fuente: https://tinyurl.com/ychzvccd

# Asuntos legales

¿Y si la Inteligencia Artificial decide que eres culpable?

SIDEBAR

# Sent to Prison by a Software Program's Secret Algorithms



- New York Times, 1 de Mayo, 2017.
- Fuente: https://tinyurl.com/lzwshya

# Investigación en redes sociales



#### The Ethics of Internet-based and Social Media Research

2016

- ¿Análisis de sentimiento? ¿Qué es un comentario positivo o negativo?
- ¿Privacidad? ¿Consentimiento?
- https://tinyurl.com/yc3vlxoo



#### Discriminación

#### ¿Y si la Inteligencia Artificial es sexista o racista?

#### Intelligent Machines

# Forget Killer Robots— Bias Is the Real Al Danger

John Giannandrea, who leads AI at Google, is worried about intelligent systems learning human prejudices.

by Will Knight October 3, 2017



- MIT Technology Reviews, 3 de Octubre, 2017.
- Fuente: https://tinyurl.com/y8jgcmgz
- Fuente: https://tinyurl.com/ya7g2p8f



#### Discriminación



- ¿Entrenamos nuestras redes con sesgo racista?
- https://tinyurl.com/md6ecwf



#### Discriminación

# Weapons of Math Destruction

- https://weaponsofmathdestructionbook.com/
- ¿Y si el acceso a una beca, el precio de un seguro de vida, las pruebas de selección para un trabajo vienen determinadas mediante un algoritmo computacional a partir de mi código postal, mi religión y mi raza?



# Diagnóstico médico

¿Y si la Inteligencia Artificial contradice el Diagnóstico Médico?



- The New Yorker, 3 de Abril, 2017.
- Fuente: https://tinyurl.com/y82nzems



# Diagnóstico médico

24 de Noviembre, 2017

No es necesario que haya contradicción con el diagnóstico ...

#### The Doctor Just Won't Accept That!

Zachary C. Lipton
Carnegie Mellon University,
Amazon AI
zlipton@cmu.edu

"I work with medical data. We work with doctors and they're interested in predicting risk of mortality, recognizing cancer in radiologic scans, and spotting diagnoses based on electronic health record data. We can train a model, and it can even give us the right answer. But we can't just tell the doctor "my neural network says this patient has cancer!" The doctor just won't accept that! They want to know why the neural network says what it says. They want an explanation. They need interpretable models!"

- Prototypical call to arms

Fuente: https://arxiv.org/pdf/1711.08037.pdf



# Nueva Ley de Protección de Datos

OPINIONS

# Is Deep Learning Going To Be Illegal In Europe?

RICHA BHATIA - JAN 30, 2018



- General Data Protection Regulation (GDPR) (EU) 2016/679. En vigor desde el 25 de mayo de 2018.
- ¿Qué es un dato personal?
- https://tinyurl.com/ydenkkm3



#### Derecho al olvido



- ¿Puedo evitar que se difunda un dato o una foto en internet?
- Regulado por ley
- https://tinyurl.com/ybmtuth9



# Otras cuestiones éticas

Desempleo

#### La Inteligencia Artificial nos quitará el trabajo a todos.

#### Predictable physical work



For example, welding and soldering on an assembly line, food preparation, or packaging objects

#### Unpredictable physical work



For example, construction, forestry, or raising outdoor animals

• Fuente: https://tinyurl.com/yb5qpvp4



# Otras cuestiones éticas Singularidad



- 23 de Febrero, 2014
- Fuente: https://tinyurl.com/ybuohyww
- Elon Musk is wrong. The AI singularity won't kill us all https://tinyurl.com/ycrrm88v



#### Otras cuestiones éticas

Con más o menos consenso ...

- Desigualdad. ¿Cómo distribuimos la riqueza creada por las máquinas?
- Convivencia. ¿Cómo afectan las máquinas a nuestro comportamiento y nuestras interacciones?
- Estupidez artificial. ¿Cómo podemos protegernos de los errores?
- Seguridad. ¿Cómo mantenemos la inteligencia artificial a salvo de terroristas?
- Genios del mal. ¿Cómo protegemos contra las consecuencias no deseadas?
- Derechos de los robots. ¿Cómo definimos el tratamiento humanitario de la IA?



# Otras aplicaciones a tener en cuenta

- Reproducir la voz de alguien
- https://tinyurl.com/y95kgyd3



# Otras aplicaciones a tener en cuenta

- Crear un entorno visual completo
- Artículo: https://tinyurl.com/ybxmkscm
- Video: https://tinyurl.com/yanyw993



Fake Obama created using AI video tool - BBC News

358.936 visualizaciones







# Otras aplicaciones a teneren cuenta

- ¿Podemos distinguir videos reales y falsos?
- https://tinyurl.com/yaqjl3pr

Current technology can allow anyone to make fake video of other people, including politicians and celebrities—which makes the future of fake news even more troubling.

#### **Intelligent Machines**

# The US military is funding an effort to catch deepfakes and other Al trickery

But DARPA's technologists admit that it might be a losing battle.

by Will Knight May 23, 2018

# ¿Se está haciendo algo?

 The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems https://perma.cc/MB4W-6GWB

Al at Google: our principles
 https://blog.google/topics/ai/ai-principles/

 The Media Lab's new Ethics initiative (MIT) https://tinyurl.com/ybsqcgk2

Code of Business Conduct and Ethics (Tesla)
 https://tinyurl.com/yafmtfpv

• ...





- 4 de junio, 2018
- Fuente: https://tinyurl.com/yadyqe4f



# Inteligencia Artificial y Ética

# Miguel A. Gutiérrez Naranjo

http://www.cs.us.es/~naranjo/

