



UNIVERSIDAD  
COMPLUTENSE  
MADRID

**Desarrollo de Aplicaciones y Sistemas Inteligentes (DASI)**

## **Introducción a la IA**

**Juan Pavón Mestras**

Dep. Ingeniería del Software e Inteligencia Artificial UCM

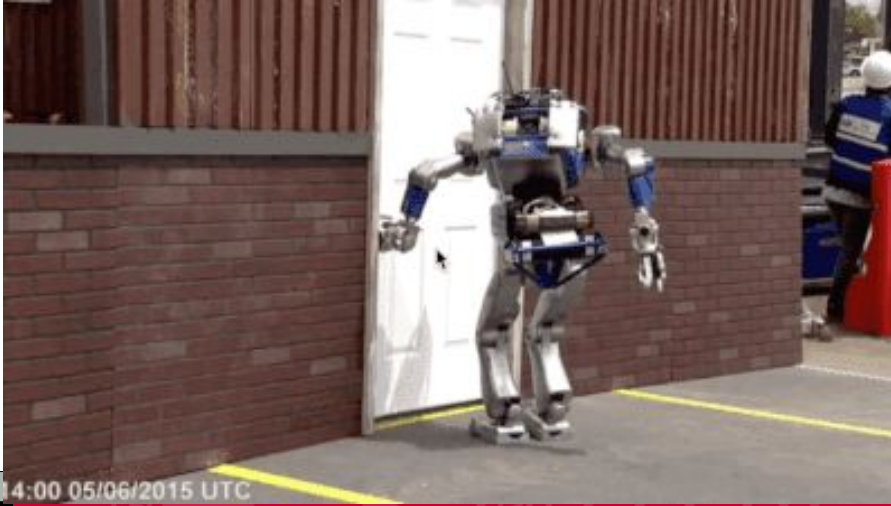
**DASI**

**¿Qué es la Inteligencia Artificial?**

## La paradoja de Moravec

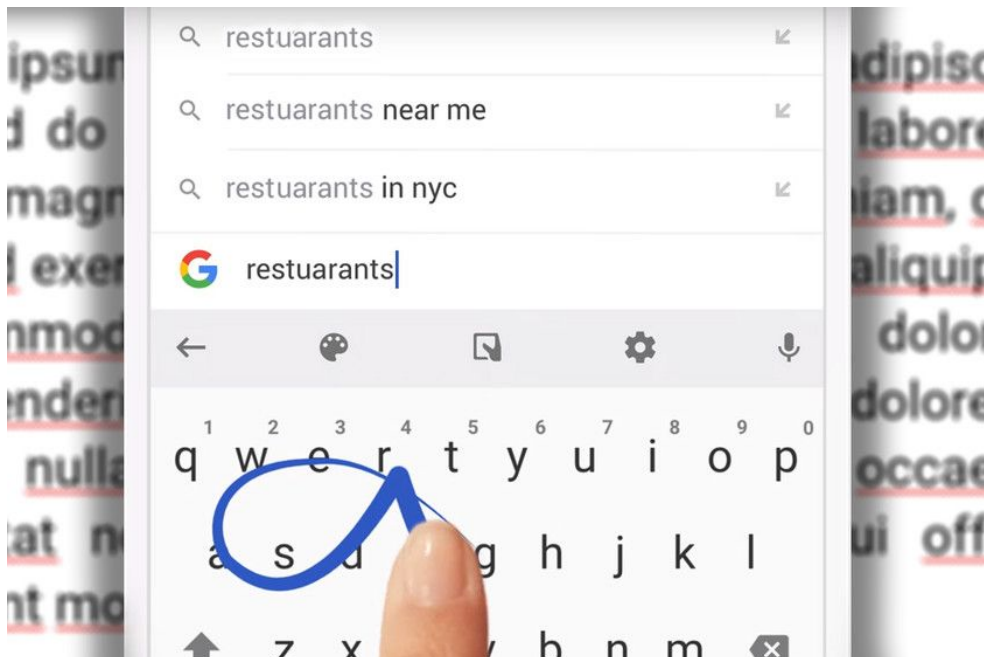
*It is comparatively easy to make computers exhibit adult level performance in intelligence tests or playing checkers, and difficult or impossible to give them the skills of a one-year-old when it comes to perception and mobility*

Hans Moravec, *Mind Children*, Harvard University Press (1988)



3

## ¿En qué es diferente la IA del software tradicional?



## Ubicua e invisible...

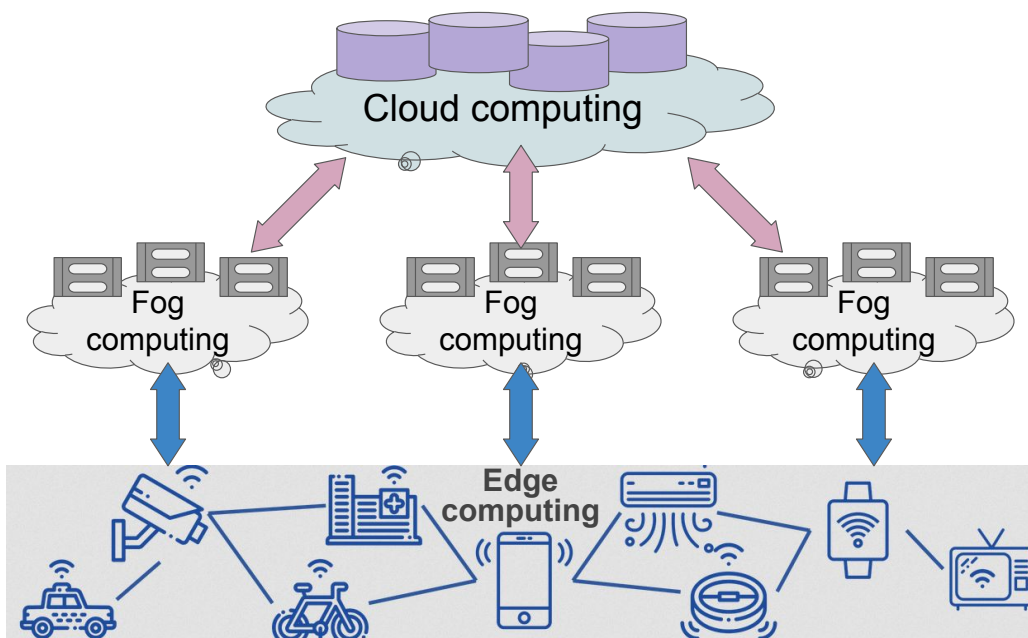


J Pavón (UCM)

DASI - Introducción a la IA

5

## Distintos niveles de computación distribuida



J Pavón (UCM)

DASI - Introducción a la IA

6

## IA en la Internet de las Cosas (Internet of Things, IoT)



## Ramas de la Inteligencia Artificial

- Representación del conocimiento
- Razonamiento y resolución de problemas
- Planificación automática
- Procesamiento de lenguaje natural (lenguaje humano)
- Aprendizaje máquina  $\Rightarrow$  Reconocimiento de patrones
- Percepción artificial
- Visión artificial
- Reconocimiento del habla y síntesis de voz
- Robótica
- Computación afectiva
- Creatividad computacional



# Inteligencia General Artificial

- Existe otra rama conocida como **Inteligencia General Artificial** (**IA fuerte**) persigue el desarrollo de sistemas que, integrando las facultades anteriores de una forma similar a como lo hace el hombre, sean capaces de obrar de forma inteligente
  - En otras palabras, su objetivo es hacer dispositivos con una inteligencia global igual o superior a la del hombre
- El propósito de los sistemas inteligentes actuales es mucho más modesto: especialización en áreas de aplicación ⇒ **IA débil**

**DASI**

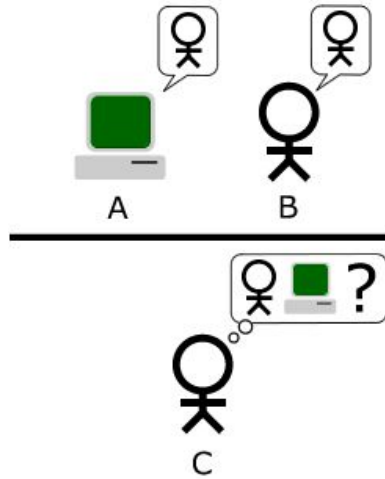
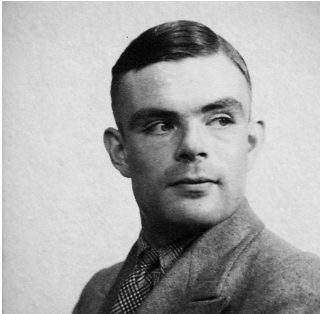
**Evolución histórica**

## ¿Pueden pensar las máquinas?

- **1950: The Turing Test (*imitation game*)**

[Alan M. Turing "Computing machinery and intelligence." Mind 59.236 (1950): 433]

Available at <http://aleph.calculemus.org/lect/07szt-intel/turpap.htm>



De Bilby - Own work, public domain,  
<https://commons.wikimedia.org/w/index.php?curid=3777581>



J Pavón (UCM)

DASI - Introducción a la IA

11

## El test de Turing

P: ¿Puedes escribir un soneto sobre el amanecer en Aranjuez?

R: Lo siento, pero nunca fui capaz de escribir poemas.

P: Suma 34957 y 70764

R: (Pausa de unos 30 segundos antes de dar la respuesta) 105621.

P: ¿Sabes jugar al ajedrez?

R: Sí.

P: Tengo el Rey en R1, y ninguna otra pieza. Tú tienes el Rey en R3 y la Dama en TD8. Te toca jugar. ¿Cuál sería tu movimiento?

R: (Tras una pausa de 15 segundos) D-TD1 mate.



J Pavón (UCM)

DASI - Introducción a la IA

12

## ¿Pueden pensar las máquinas?

- Objeciones

- **Objeción teológica**

*Thinking is a function of man's immortal soul. God has given an immortal soul to every man and woman, but not to any other animal or to machines. Hence no animal or machine can think.*

- **Objeción "Esconder la cabeza como el avestruz"**

*The consequences of machines thinking would be too dreadful. Let us hope and believe that they cannot do so.*

- **Objeción matemática**

Hay límites a las cuestiones que una máquina puede responder (Church, Kleene, Rosser, Turing, etc.)

- Pero, ¿y los humanos?

## ¿Pueden pensar las máquinas?

- Objeciones sobre los **límites de las máquinas**

- Las máquinas **no cometen errores**: pero los pueden simular
  - Una máquina **no puede tener comportamientos variados**: realmente se le pueden añadir tantos como queramos
  - **Falta de originalidad** (objeción de *Lady Lovelace*): *a machine can never do anything really new*, esencialmente porque no puede aprender sola
  - Una máquina **no tiene consciencia de sí misma**: pero un depurador puede informar de su estado interno

⇒ Turing propone la hipótesis de construir una *child machine* elemental capaz de aprender y que vaya adquiriendo nuevas competencias hasta que pueda competir con los humanos en todos los campos del intelecto

## Competiciones del test de Turing

**Mitsuku:** *Loebner Prize* winner (2013, 2016, 2017, 2018, 2019)

- <https://pandorabots.com/mitsuku/>



### Computer game bot

- Unreal Tournament bots

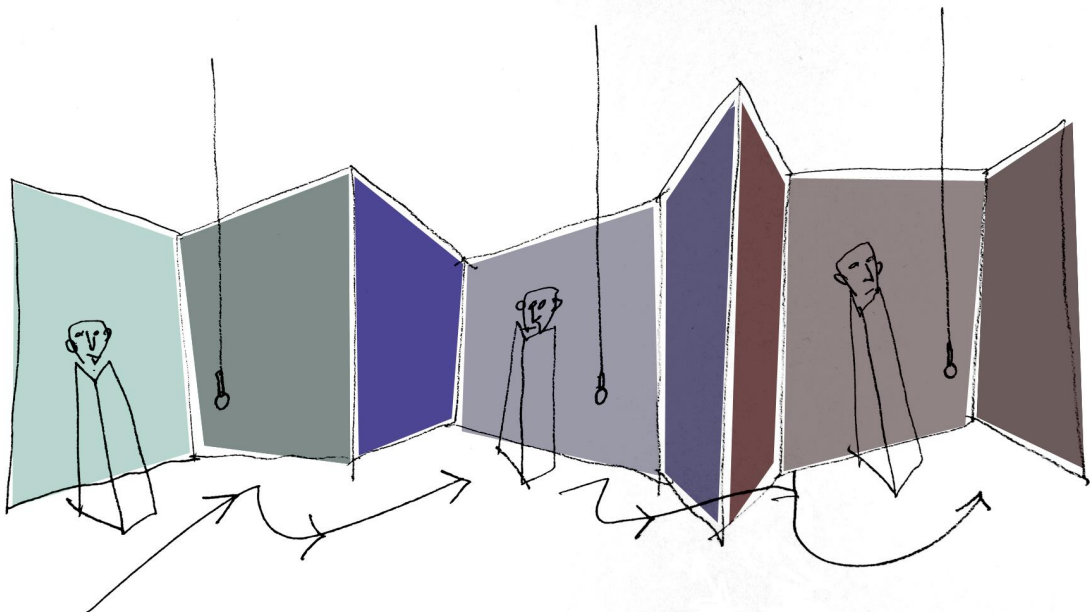


J Pavón (UCM)

DASI - Introducción a la IA

15

## Talking Agents - Oráculos (MdM 2008)



JM Fernández, J Pavón. *Talking Agents: A distributed architecture for interactive artistic installations.* Integrated Computer-Aided Engineering 17 (3), 243-259. Drawing by **Eduardo Gómez**, Masdelomismo



J Pavón (UCM)

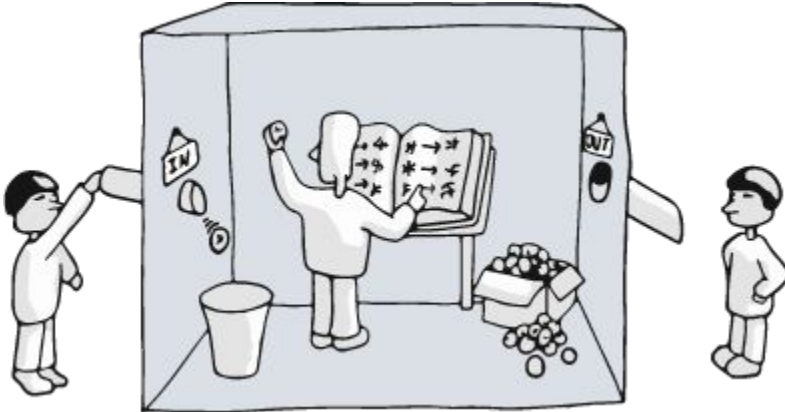
DASI - Introducción a la IA

16



## ¿Pueden pensar las máquinas?

- **El argumento de la consciencia:** los computadores no pueden tener experiencias conscientes  
⇒ John Searle (1980), la habitación china  
<https://www.youtube.com/watch?v=TryOC83PH1g>  
[https://www.youtube.com/watch?v=j\\_OPQgPIIdKg](https://www.youtube.com/watch?v=j_OPQgPIIdKg) (Consciousness & the Brain: John Searle at TEDxCERN)



## Inteligencia Artificial y el Test de Turing

- Para que una máquina supere esta prueba tendría que estar dotada de al menos las siguientes capacidades:
  - Debe ser capaz de **comunicar** satisfactoriamente en la lengua del juez
  - Debe ser capaz de almacenar y **recordar** lo que conoce
  - Debe utilizar la información almacenada para **razonar** sobre ello y responder preguntas
  - Debe ser capaz de **aprender** y extraer conocimiento adaptándolo a nuevas circunstancias
- Estas facultades junto a otras relacionadas con la **percepción** de estímulos físicos y la capacidad de **moverse** y de manipular objetos constituyen hoy en día áreas de estudio de la IA

## Inteligencia Artificial y el Test de Turing

⇒ Hoy en día, se considera que no tiene sentido reproducir la inteligencia humana en una máquina:

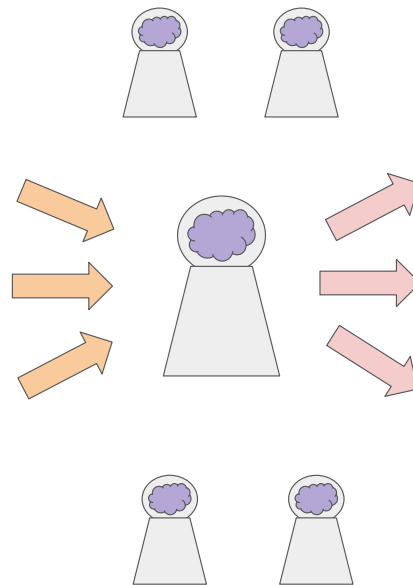
- Sin embargo, sí se pretende dotar a las máquinas de facultades asociadas con la inteligencia
- Para ello, se debe profundizar en los principios en los que se basa la inteligencia
- Algo parecido sucedió en la aviación que tuvo éxito cuando se dejó de intentar imitar a los pájaros y se comprendieron mejor las leyes de la aerodinámica
- Por ello, el estudio de la IA se ha ramificado en distintos campos que trabajan sobre facultades que caracterizan la inteligencia

## Historia de la IA

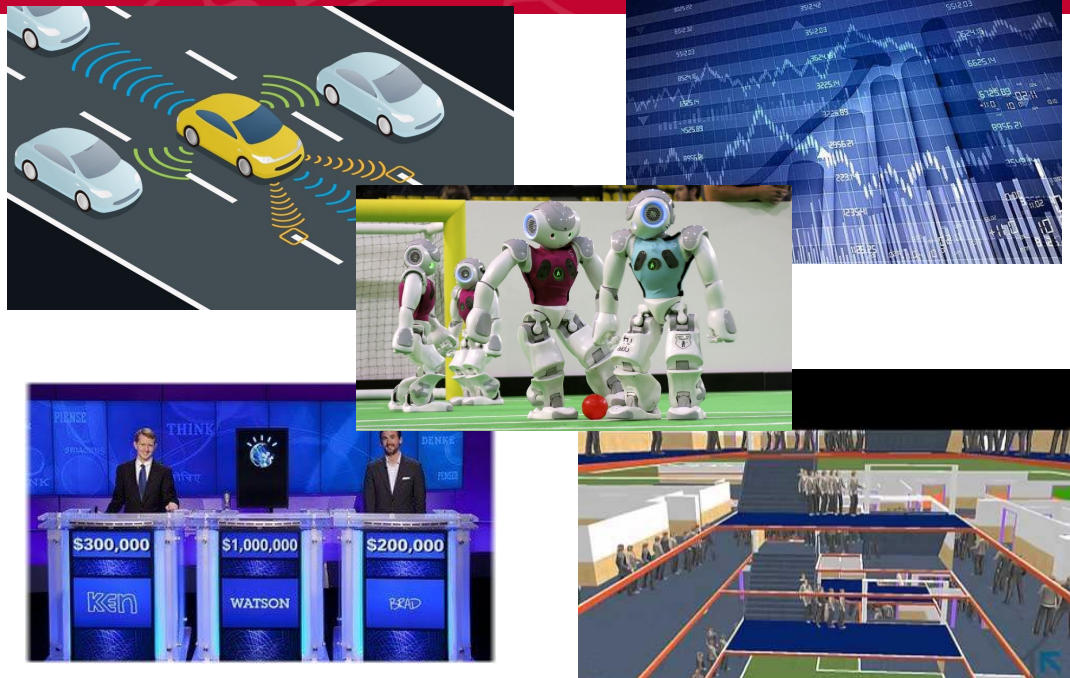
- **1943:** Warren McCulloch y Walter Pitts: primer modelo de redes neuronales artificiales
- **1956:** Dartmouth Artificial Intelligence (AI) conference: nacimiento de la IA (McCarthy, Minsky, Newell, Simon, etc.)
  - LISP, General Problem Solver, Perceptron, Fuzzy sets, ...
  - Expectativas muy optimistas para 10 años, ...que no se cumplieron
- **Mediados-1970:** Desilusión
  - Algún avance: Prolog
- **1982** Japón: Fifth Generation Computer (FGC)
  - Evolución del Prolog (Kernel Language) en computación paralela
  - Sistemas expertos, comercialización

⇒ Pero vuelven a fallar las expectativas
- **1995:** Sistemas multi-agentes
  - Base para la IA distribuida
  - Robots, agentes en bolsa, asistentes personales (Cortana, Siri, Alexa), vehículos autónomos, etc.

**A**utonomía  
**E**ntorno  
**I**nteracciones  
**O**rganización  
**U**suario



## Ejemplos de sistemas multi-agentes



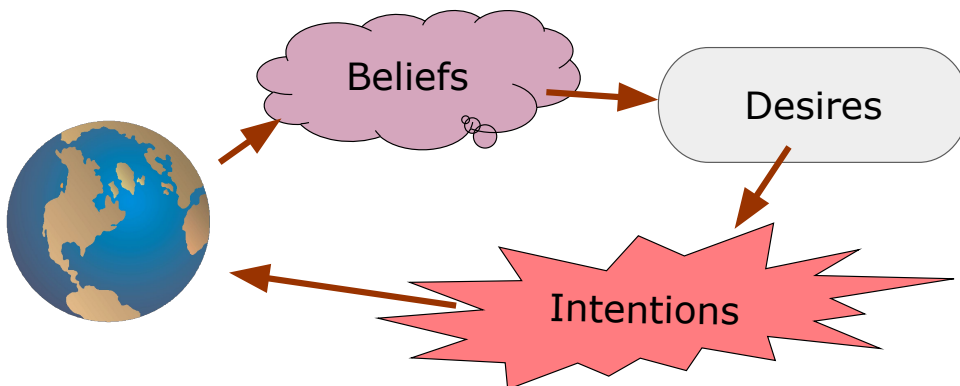
## Sistemas inteligentes

- Principio de Racionalidad (Allen Newell (1982). *The knowledge level*. Artificial Intelligence **18**: 87-127):  
*If an agent has knowledge that one of its actions will lead to one of its goals, then the agent will select that action.*
  - ... but this does not imply that it will take the best decision
- ⇒ El entorno es dinámico, complejo, incierto, y solo prácticamente conocido
- Razonamiento
  - Un agente puede decidir:
    - qué objetivo perseguir o a qué evento reaccionar
    - qué acciones (plan) realizar para lograr un objetivo
    - suspender o abandonar un objetivo para perseguir otro
- Aprendizaje
  - Un agente puede adaptarse progresivamente a cambios en entornos dinámicos



## ¿Cómo piensan los agentes?

- Modelo BDI (Bratman, 1987)



## Sociabilidad

## Interacción

## Diálogo

## Delegación

## Cooperación

## Coordinación

## Negociación

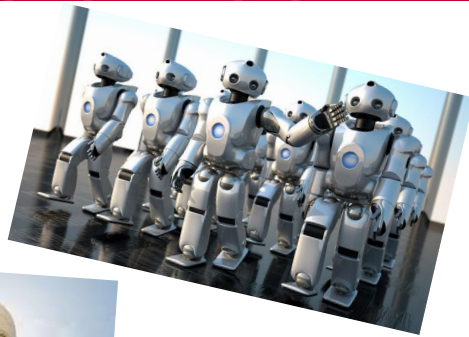


# SMART everything...





## Inteligencia Artificial Hoy



J Pavón (UCM)

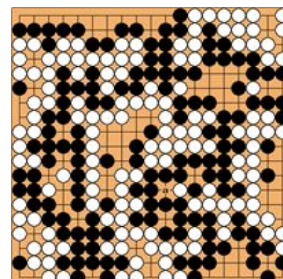
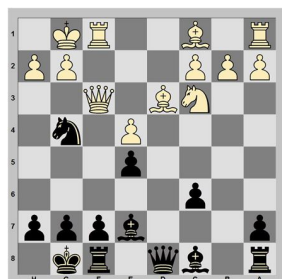
DASI - Introducción a la IA

27

## Inteligencia Artificial Hoy

- 1997: IBM Deep Blue derrota al campeón de ajedrez G. Kasparov
- 2011: IBM Watson gana el concurso *Jeopardy!* contra los dos mejores jugadores en la historia de este juego de preguntas y respuestas
- 2014: El programa Eugene (Rusia) pasa el test de Turing simulando un niño ucraniano de 13 años
- 2016: AlphaGo (Google DeepMind) derrota al campeón mundial de Go

$10^{120}$



$10^{360}$

átomos en un humano:  $10^{27}$   
átomos en la Tierra:  $10^{50}$   
átomos de H en el universo:  $10^{79}$



J Pavón (UCM)

DASI - Introducción a la IA

28

## AlphaGo Zero

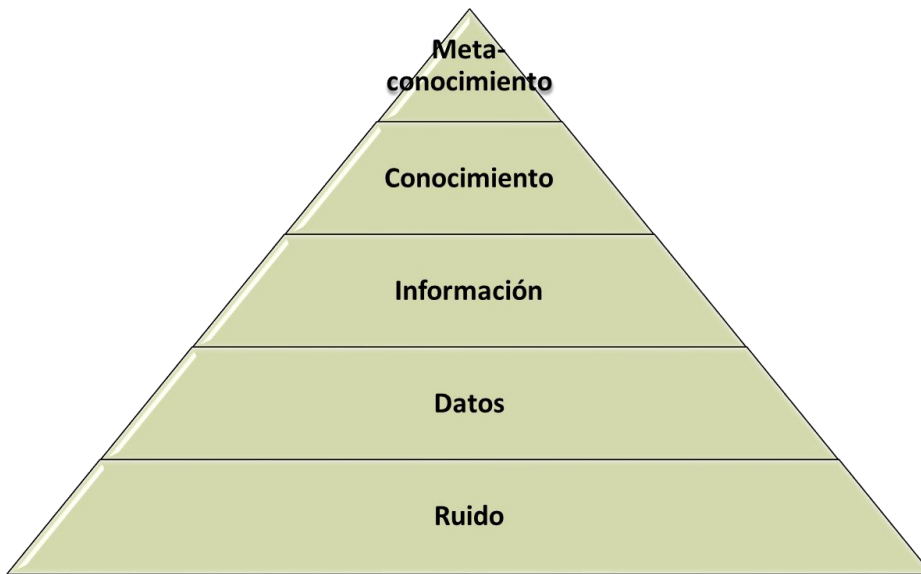
- AlphaGo: Full Documentary - <https://youtu.be/WXuK6gekU1Y>
- AlphaGo Zero: Starting from scratch - <https://youtu.be/WXHFqTvfFSw>
  - Reportaje de El País: <https://youtu.be/tCpf5wDr0UE>
  - Reportaje de El Independiente: [https://youtu.be/Pn\\_zGDJwkcU](https://youtu.be/Pn_zGDJwkcU)



**DASI**

**Métodos y Técnicas**

## Jerarquía del conocimiento



## Gestión del conocimiento

- **Representación del conocimiento**
  - De forma explícita  $\Rightarrow$  mediante **símbolos**
    - Puede ser estructurado, almacenado, distribuido, procesado
  - Sin representación explícita  $\Rightarrow$  **representación sub-simbólica**
    - Ejemplos: Redes neuronales, algoritmos genéticos
- **Razonamiento**
  - Manipulación de símbolos para producir **nueva información**
    - Los símbolos representan una colección de proposiciones en las que se cree
    - La nueva información son representaciones de nuevas proposiciones
- **Aprendizaje**
  - Adquisición de **nuevo conocimiento**
    - A partir de **nueva información del entorno**
    - **Algoritmos** de aprendizaje

## Tipos de sistemas inteligentes

- **Sistemas basados en conocimiento (IA simbólica)**
  - La inteligencia se puede reducir a un mero problema de manipulación de símbolos, cuyas relaciones se pueden organizar en estructuras como listas, jerarquías o redes
  - El conocimiento se representa de manera explícita y formal (mediante símbolos) para que pueda ser procesada por un algoritmo de razonamiento e inferir soluciones a problemas o preguntas del dominio
  - Ejemplos: Sistemas expertos, algoritmos de búsqueda, inferencia, ontologías, planes
- **Sistemas dirigidos por datos (*Machine learning*)**
  - No hay una representación explícita del conocimiento
  - El sistema aprende a partir de ejemplos o de la experiencia en el uso del sistema
  - Los datos observados proporcionan información incompleta sobre fenómenos y los algoritmos de aprendizaje intentan generalizar la información para hacer predicciones sobre fenómenos
  - Ejemplos: Redes neuronales artificiales, deep learning, SVM, aprendizaje bayesiano, etc.

## IA simbólica vs. ML

### IA simbólica

- Manipulación de símbolos que representan conceptos o entidades del mundo real
- Usando lógica y algoritmos de búsqueda
- Más útil para razonamiento
- Más fácil de depurar
- Explicabilidad
- Más control
- Mejor para problemas abstractos
- Problemas bien definidos
- Mal adaptabilidad para entornos dinámicos
- Conocimiento explícito (hay que codificarlo)

### Aprendizaje automático

- Más robusto ante ruido
- Mejor eficiencia
- Menor necesidad de conocimiento a priori
- Más escalable
- Big Data
- Más relacionado con la neurociencia
- Mejor para problemas de percepción y reconocimiento de patrones
- Falta de transparencia y explicabilidad
- No apropiado para toma de decisiones de alto riesgo

## Deducción vs. inducción

- Razonamiento **deductivo**

- Infiere hechos/verdades particulares a partir de una verdad general, basado en un sistema de lógica
- El razonamiento será válido o inválido

*Todos los humanos son mortales*

*Sócrates es humano*

*⇒ Sócrates es mortal*

- Razonamiento **inductivo**

- Las premisas sirven de evidencia (fundamento) para concluir algo
- Se infieren verdades generales a partir de verdades particulares
- El razonamiento no es válido o inválido, solo probable (mejor o peor)

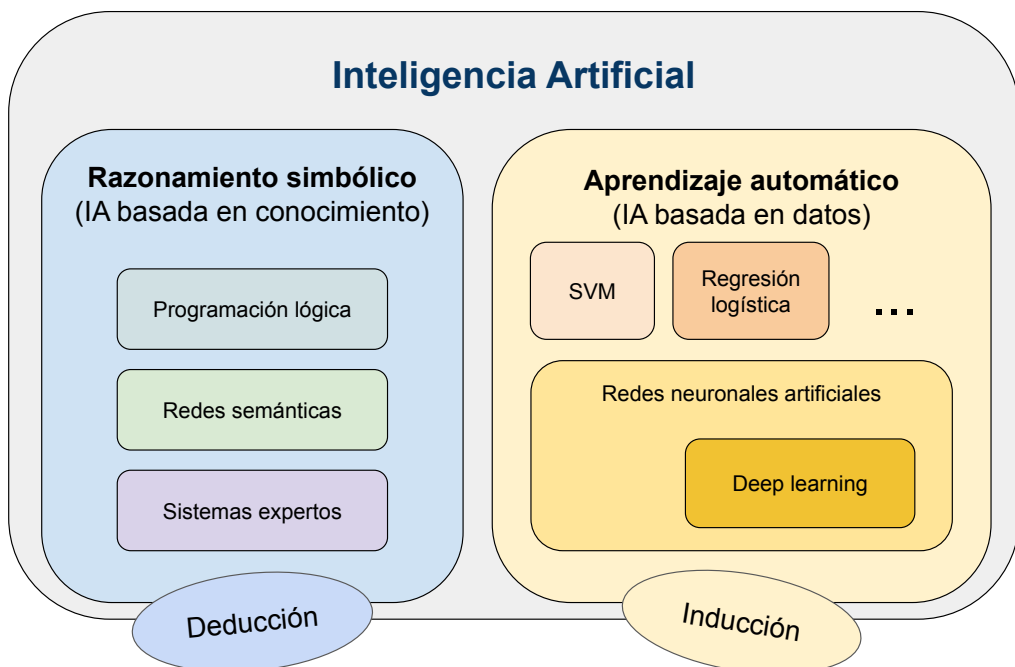
*Sócrates es humano y mortal*

*Platón es humano y mortal*

*Aristóteles es humano y mortal*

*⇒ Probablemente, todos los humanos son mortales*

## Métodos de IA





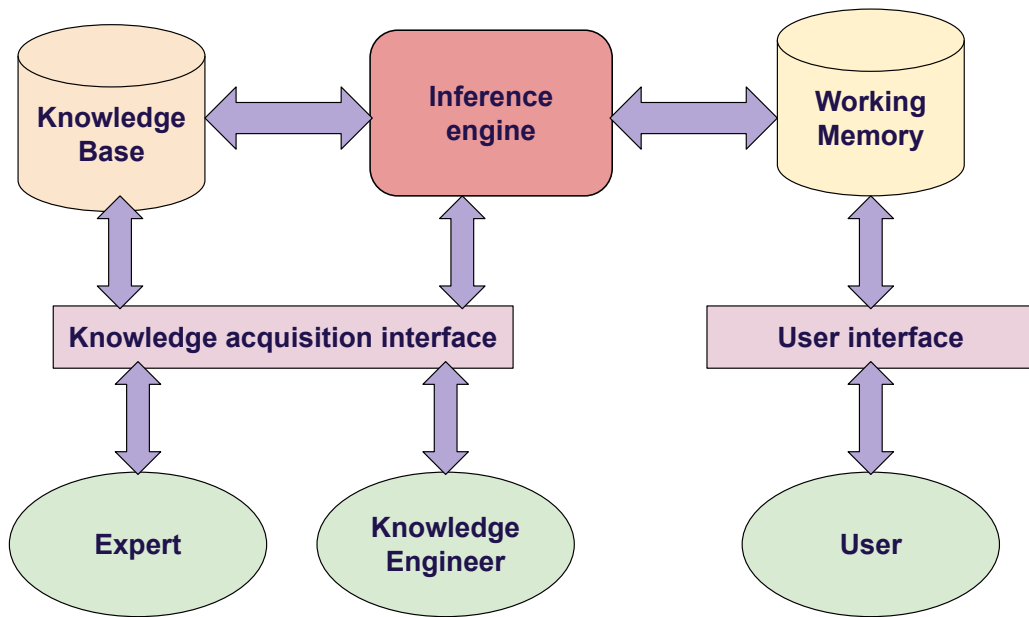
# DASI

## Razonamiento simbólico IA basada en el conocimiento

### Tipos de conocimiento en un sistema inteligente

- Conocimiento **normativo**: conocimiento general del dominio que se almacena en la base de conocimientos
- Conocimiento **factual**: Datos concretos de un problema a resolver, y a las conclusiones intermedias y finales
  - Basado en **observaciones**: datos que al interpretarlos el sistema cobran un significado para él, convirtiéndose en información, que al ser integrada se convertirá en conocimiento
  - Resultado de **razonamientos** y que puede ser o no verificable por observación o experimentación
- Conocimiento **estratégico** (o de **control**): se encarga de la manipulación del conocimiento factual y normativo.
  - Por ejemplo, si el normativo está expresado mediante reglas, se refiere a la estrategia de búsqueda que decide en cada momento qué regla se aplica

## Sistema de gestión de conocimiento



## Razonamiento basado en reglas

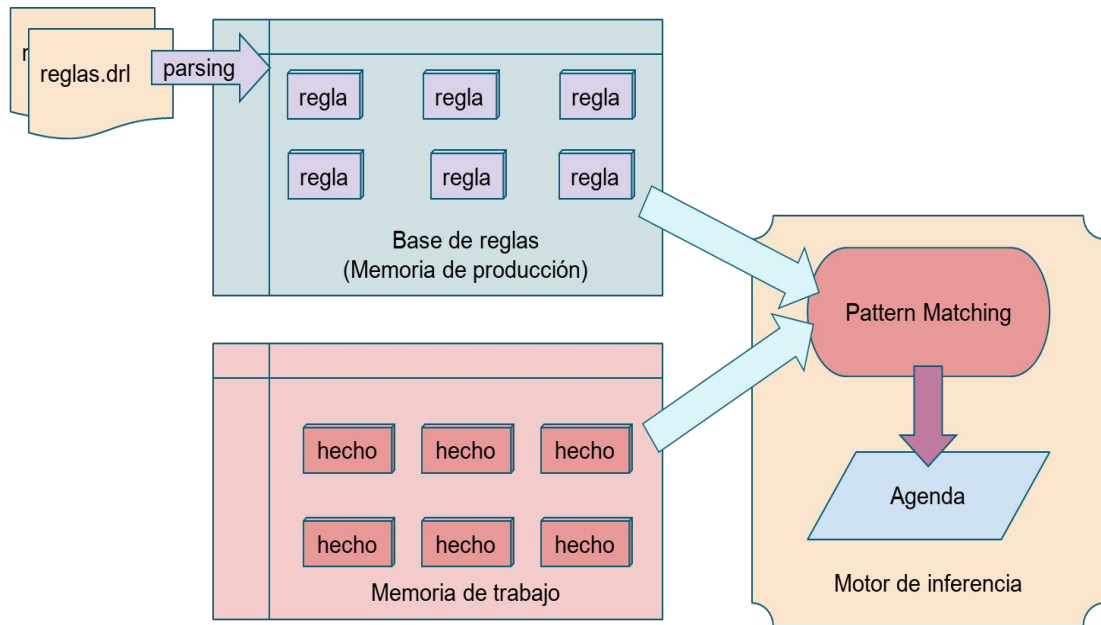
- El Razonamiento Basado en Reglas, también llamados **sistemas de producción**, es el esquema de representación de conocimiento procedimental más popular
  - El conocimiento procedimental se centra en las acciones
- Las reglas representan un conjunto de habilidades de resolución de problemas en un dominio
  - Adecuadas para representar conocimiento heurístico
  - Establecen patrones que indican cómo reaccionar ante una situación
  - Reglas explícitas estímulo → respuesta

***SI situación ENTONCES acción***

***SI premisas ENTONCES conclusión***

- Ejemplos de software para razonamiento basado en reglas: CLIPS, JESS, Drools

## Estructura de un sistema basado en reglas



## Ejemplo de sistema basado en reglas

- Sistema de predicción del tiempo
  - **Reglas:**
    - R1:** IF temperatura > 25 THEN hace-calor
    - R2:** IF humedad-relativa > 75% THEN atmósfera-húmeda
    - R3:** IF hace-calor AND atmósfera-húmeda THEN probabilidad-tormentas-alta
  - **Hechos:**
    - F1:** humedad-relativa = 80%
    - F2:** temperatura = 32
  - **Ejecución ⇒ Nuevos hechos:**
    - R2 ⇒ F3:** atmósfera-húmeda
    - R1 ⇒ F4:** hace-calor
    - R3 ⇒ F5:** probabilidad-tormentas-alta

## Sistemas basados en reglas

- Dominios de aplicación
  - Aquellos con conocimiento difuso, incompleto y no muy bien definido, pero con muchos hechos
    - Ej. diagnóstico médico o de averías
  - Dominios que separan el conocimiento de la forma de uso
    - Ej. reconocimiento de moléculas o asesoramiento jurídico
  - Dominios con conocimiento experto en forma de heurísticas
    - Ej. exploración de grandes espacios de alternativas
- Desventajas
  - Ineficiencia: Hay que comprobar las múltiples combinaciones posibles de hechos y condiciones de reglas
  - Dificultad de verificación de la consistencia de la base de conocimiento:  
¿Son consistentes las reglas? ¿Los hechos? ¿Ambos?
  - Opacidad y dificultad de depuración
    - Dificultad para seguir el flujo de ejecución: Las reglas representan pasos muy pequeños en la resolución del problema completo, se pierde la visión global
  - Dificultad para cubrir todo el conocimiento:
    - Hay que extraer y organizar grupos de reglas del conocimiento del dominio
    - Cuello de botella del desarrollo en la adquisición de conocimiento

## Razonamiento basado en casos *Case based reasoning (CBR)*

- El ciclo del CBR



## Razonamiento basado en casos *Case based reasoning (CBR)*

- Aplicaciones
    - Tareas de clasificación: diagnóstico en medicina o mecánica, pronóstico de fallos o del resultado de operaciones financieras, análisis de riesgos de seguros, recuperación de casos legales
    - Tareas de síntesis: creación de un sistema o de un producto, elaboración de planes de trabajo, de recetas, argumentación legal
  - Desventajas
    - La base de casos puede ser pobre ⇒ dificultad para plantear soluciones
    - La base de casos puede estar sesgada
    - Al aumentar el tamaño y la variedad de la base de casos, se produce un conflicto entre:
      - El ahorro que se produce en la fase de adaptación
      - El deterioro del tiempo de recuperación de los casos relevantes
      - Aunque con cada caso nuevo el tiempo de adaptación disminuye, la mejora es menor a cada caso, mientras que el tiempo de recuperación empeora.
- ⇒ Debe evitarse almacenar casos redundantes o que aporten poco en cuanto a operaciones de adaptación requeridas



**DASI**

**Aprendizaje automático**  
***Machine Learning***



## ¿Qué puede aprender un agente inteligente?

1. La correspondencia de condiciones del estado actual a acciones
2. Los medios para inferir propiedades relevantes del mundo a partir de una secuencia de percepciones
3. Información sobre el modo en que el mundo evoluciona y los resultados de las posibles acciones
4. Información (función) de utilidad que indique lo deseable que son los estados del mundo
5. Información sobre el valor de las acciones que indique cuán deseables son las acciones
6. Objetivos que describen clases de estados cuya consecución maximiza la utilidad del agente

## ¿Qué puede aprender un agente inteligente?

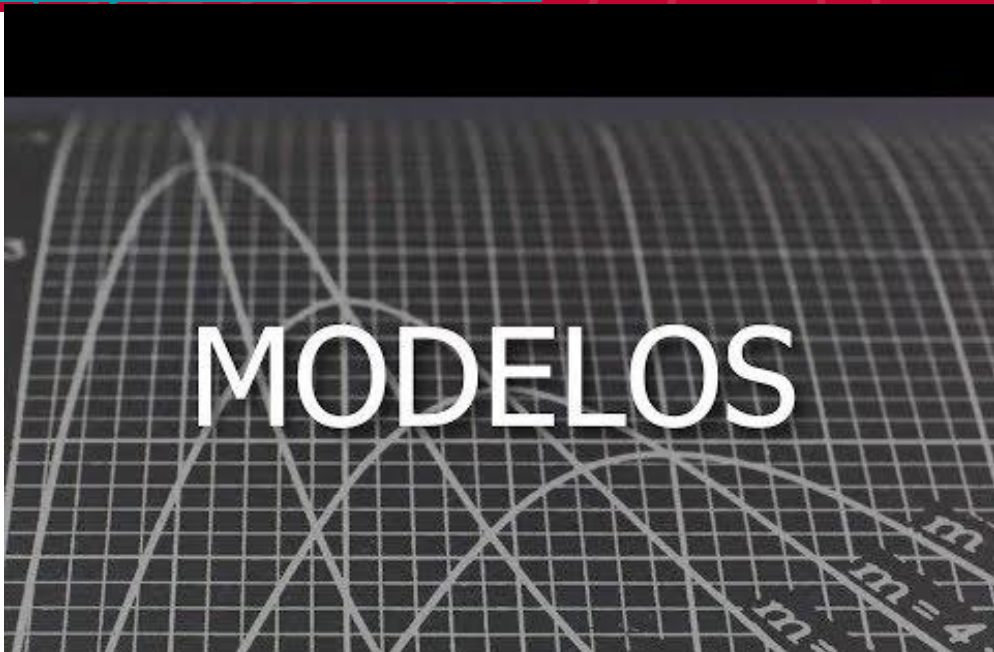
- Ejemplo: Asistente de conducción de un taxi [Russel y Norvig, 2016]
  - Cuando el usuario grite «Frena» el agente debe aprender una regla condición→acción para saber cuándo debe frenar (1)
  - El agente también aprende cuando el usuario no grita
  - Viendo muchas imágenes que se indica que tienen autobuses, aprenderá a reconocerlos (2)
  - Intentado realizar acciones y observando los resultados, por ejemplo al frenar sobre firme mojado, aprenderá el efecto de sus acciones (3)
  - Si no recibe propina de los pasajeros que han sufrido varias sacudidas durante su viaje aprenderá un componente útil para mejorar su función de utilidad (4)

## Tipos de problemas que se resuelven con ML

- **Clasificación:** indicar a qué clase predefinida pertenece un objeto
- **Agrupamiento (*clustering*):** dividir un conjunto de objetos heterogéneos en grupos homogéneos
- **Diagnóstico:** inferir problemas en el funcionamiento y sugerir soluciones
- **Selección:** recomendar la mejor opción de entre varias alternativas
- **Predicción:** predecir el comportamiento futuro en base al pasado
- **Optimización:** mejorar una solución hasta obtener la óptima
- **Control:** regular el comportamiento de un objeto en tiempo real de acuerdo a una serie de requisitos

Para ello el objetivo es construir **MODELOS**

<https://youtu.be/Sb8XVheowVQ>



## Tipos de aprendizaje máquina

- Aprendizaje **supervisado**
  - Se le presenta al sistema un conjunto de entrenamiento (pares de problemas *entrada-solución*)
    - O se corrigen las soluciones que genera a partir de las entradas
  - El sistema almacena y generaliza estas soluciones
  - Requiere preparar (*etiquetar*) las entradas (y soluciones) apropiadas
  - Suele tener problemas de adaptación en entornos cambiantes  
⇒ Necesita nuevo entrenamiento
- Aprendizaje **no supervisado**
  - El sistema no recibe información (feedback) acerca de la salida esperada
    - Aprende de su propia observación
    - A partir de las propiedades de las entradas realiza una agrupación o caracterización (clasificación, clustering) de las entradas según la similitud de sus propiedades
  - Le permite adaptarse mejor a entornos cambiantes
  - Pero las soluciones aprendidas pueden no ser las deseadas
    - Depende de la experiencia con la que aprendió

## Más tipos de aprendizaje máquina

- Aprendizaje **semi-supervisado**
  - Una combinación de los dos anteriores
  - Utiliza ejemplos clasificados (etiquetados) y no clasificados (sin etiquetar)
  - Útil cuando algunos ejemplos no están clasificados o lo están erróneamente
- Aprendizaje **por refuerzo**
  - El agente aprende observando el entorno y las consecuencias de sus acciones (recompensas o castigos)
  - Proceso de prueba y error, reforzando aquellas acciones que reciben una respuesta positiva en el mundo

## Aprendizaje supervisado

- Si tenemos esta serie:

$5 \rightarrow 7$

$12 \rightarrow 14$

$3 \rightarrow 5$

$123 \rightarrow 125$

- ¿Cuál será la solución para los siguientes casos ?

$26 \rightarrow ?$

$2 \rightarrow ?$

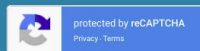
$7 \rightarrow ?$

⇒ Mediante observación generalizamos el conocimiento



## Google Recaptcha

### Google reCAPTCHA v3



Type the text

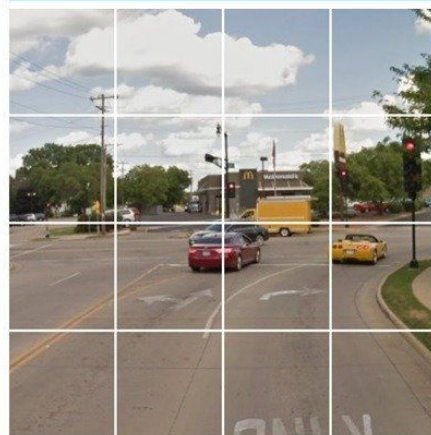
lights squares

Verify

I'm not a robot

reCAPTCHA Privacy - Terms

Select all squares with  
**traffic lights**  
If there are none, click skip

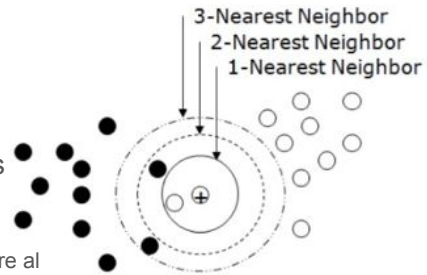


SKIP



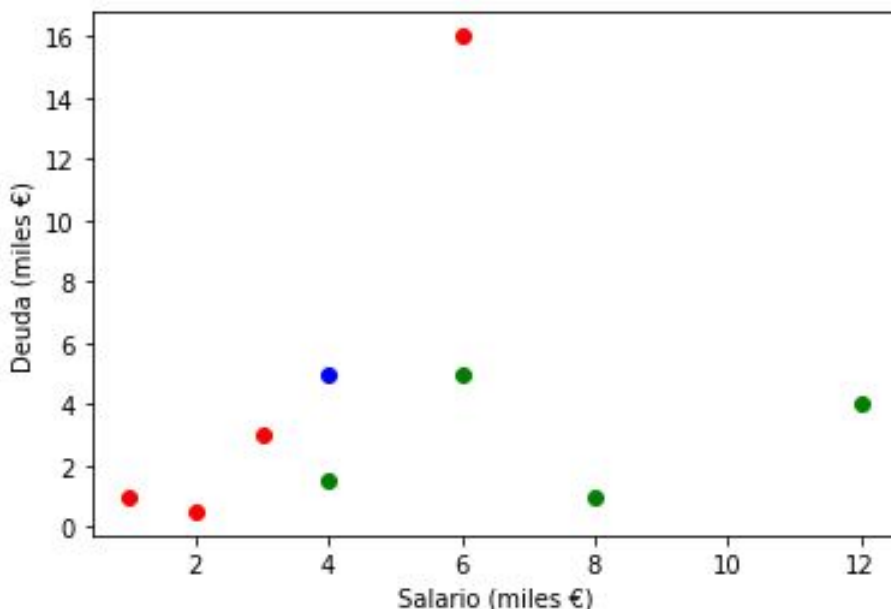
## Los k-vecinos más cercanos (k-NN)

- Los k-vecinos más cercanos (o k-NN por *k-nearest neighbor*) es uno de los algoritmos más sencillos
  - Es supervisado: parte de un conjunto de datos (ejemplos) de entrenamiento, esto es, que para cada ejemplo se conoce su resultado
  - Se dice basado en instancias (ejemplos) porque no aprende explícitamente un modelo (como ocurre con la regresión logística), sino que utiliza los ejemplos de entrenamiento para la predicción
- Considera la distancia de un punto a los k ejemplos (instancias) conocidos más cercanos para determinar a qué clase pertenece o el valor que le corresponde
  - Utiliza una medida de distancia: euclídea, L1, L, Mahalanobis, ...
  - Devuelve como valor solución la media de los valores solución de los k más cercanos
    - Puede ser una media aritmética estándar o una media aritmética ponderada, donde se dé más peso a un vecino según lo cerca que se encuentre al valor para el que queremos determinar la solución



## Ejemplo

- Determinar si dar un préstamo a un cliente





## Regresión lineal

- Modelo matemático que permite aproximar la relación de dependencia entre una variable dependiente  $Y$  y variables independientes  $X_i$

$$y = w_0 + w_1x_1 + w_2x_2 + \dots$$

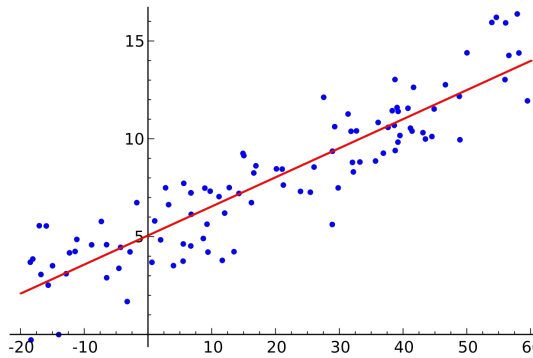
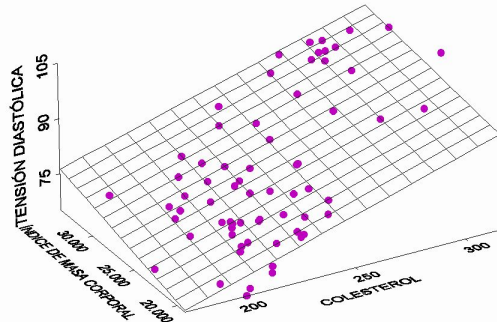
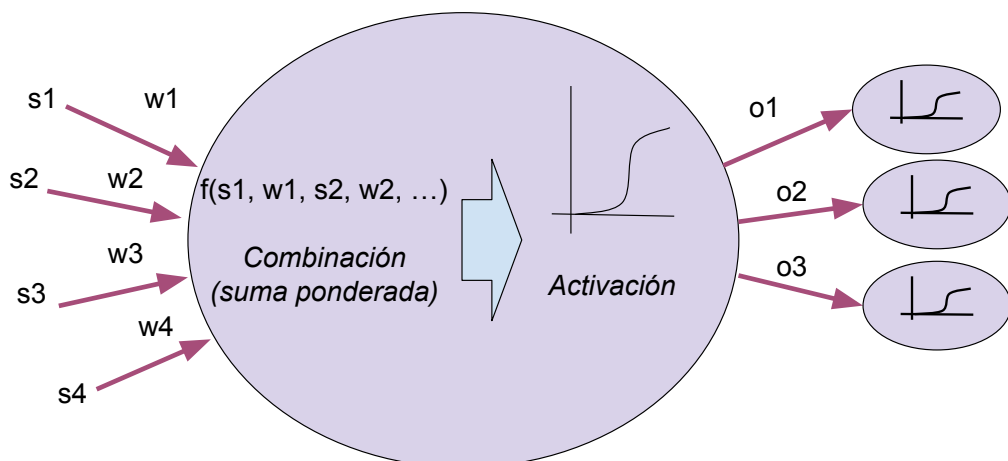


Figura 1.- Plano de regresión para la tensión arterial diastólica ajustando por colesterol e índice de masa corporal.



## Artificial Neural Networks

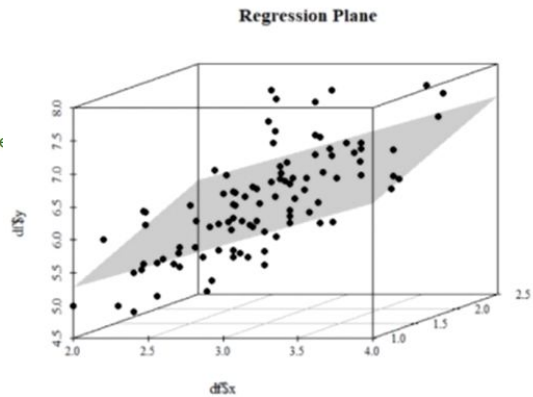
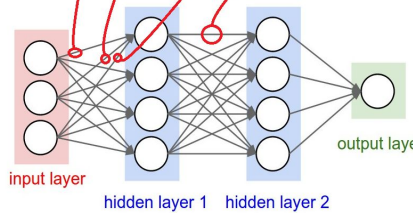


# Neurona artificial y regresión lineal

Multiple Regression Formula :

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n = Y$$

Some additional bias



$$\hat{y} = w_1 x_1 + w_2 x_2 + b$$



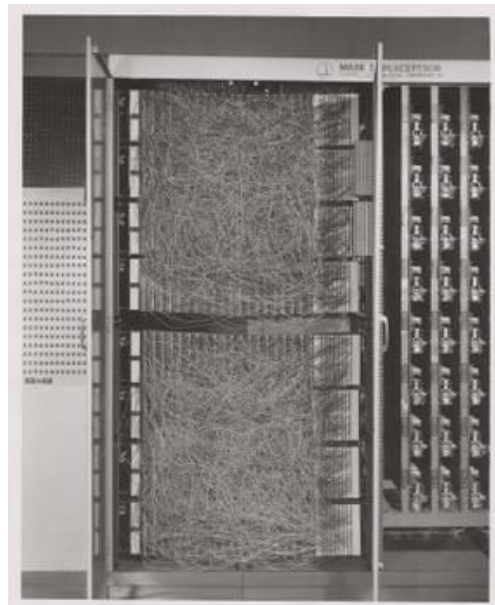
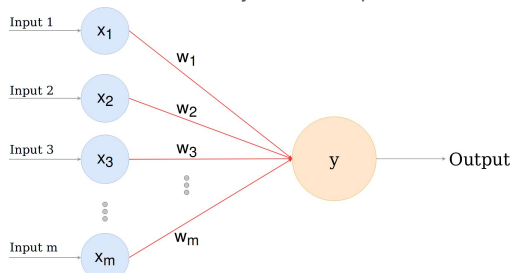
J Pavón (UCM)

DASI - Introducción a la IA

59

## Perceptron

- Frank Rosenblatt (Cornell Aeronautical Laboratory, 1958)
  - Mark I Perceptron machine
    - Image recognition
    - Camera with 400 photocells, randomly connected to the artificial neurons
    - Weights encoded in potentiometers, and weight updates during learning performed by electric motors
    - Not many classes of patterns

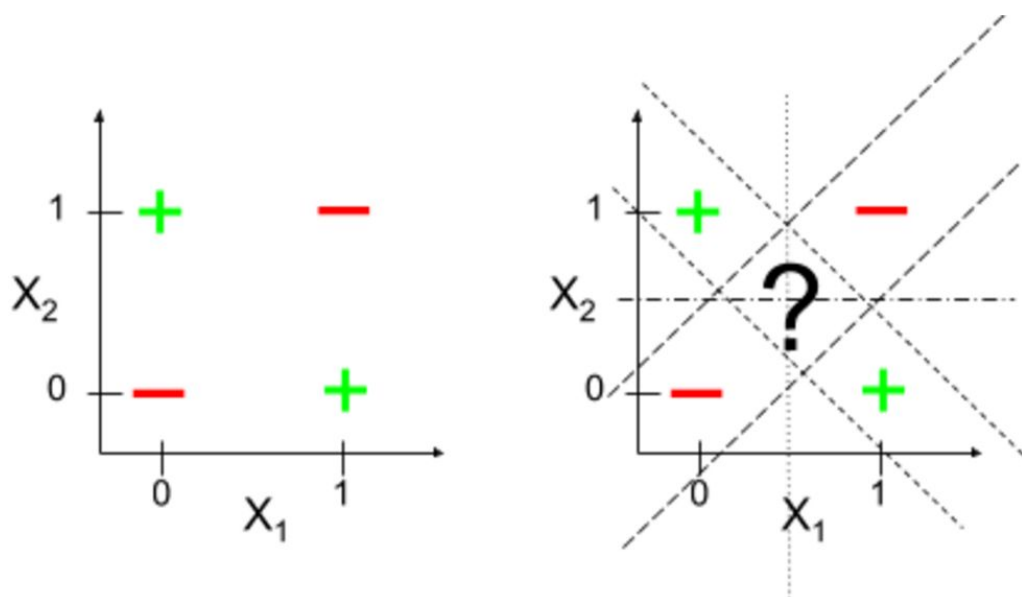


J Pavón (UCM)

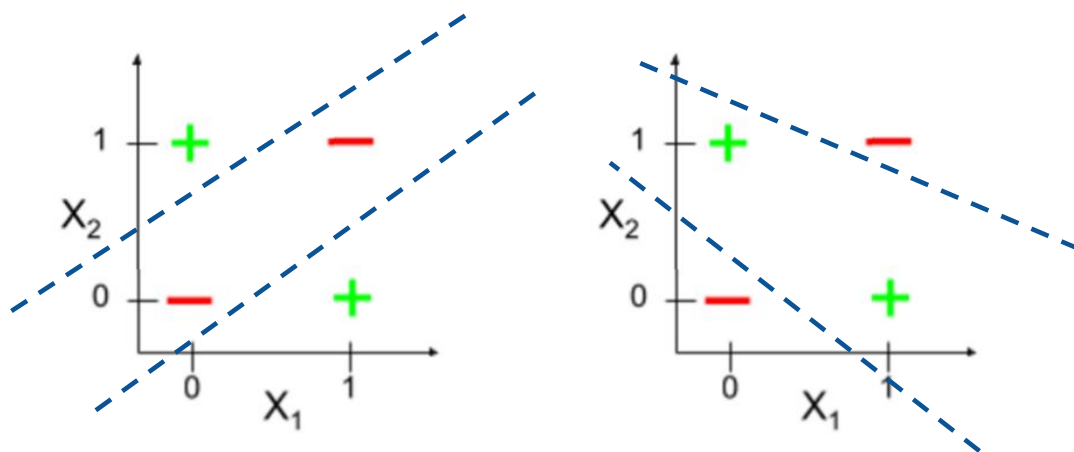
DASI - Introducción a la IA

60

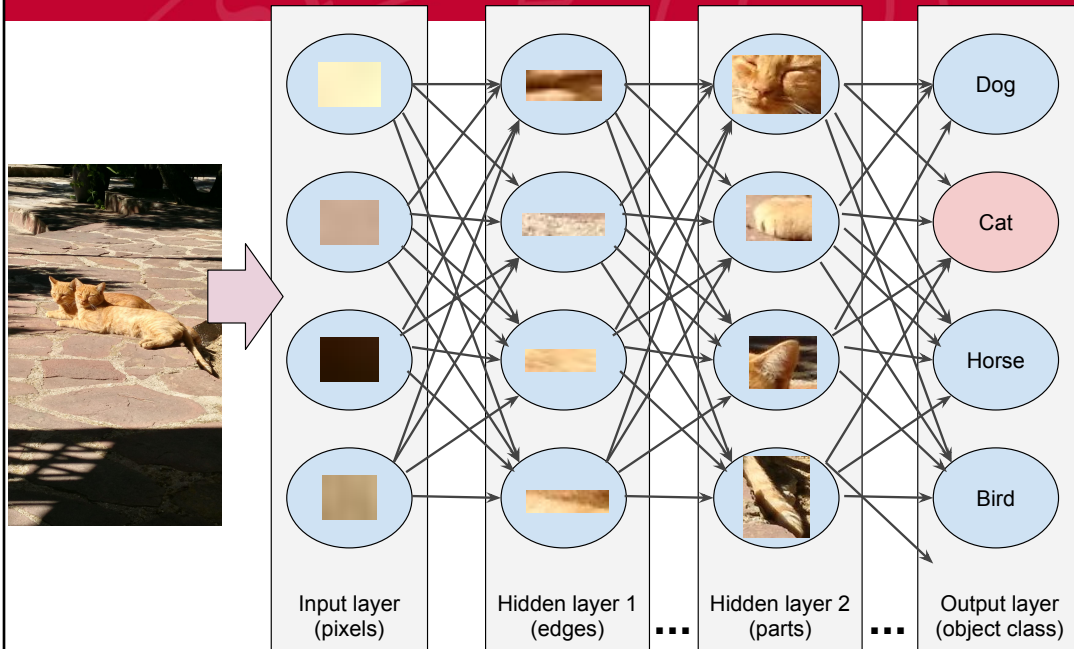
## Problema XOR



## Problema XOR - Solución con varias neuronas



## Deep Learning

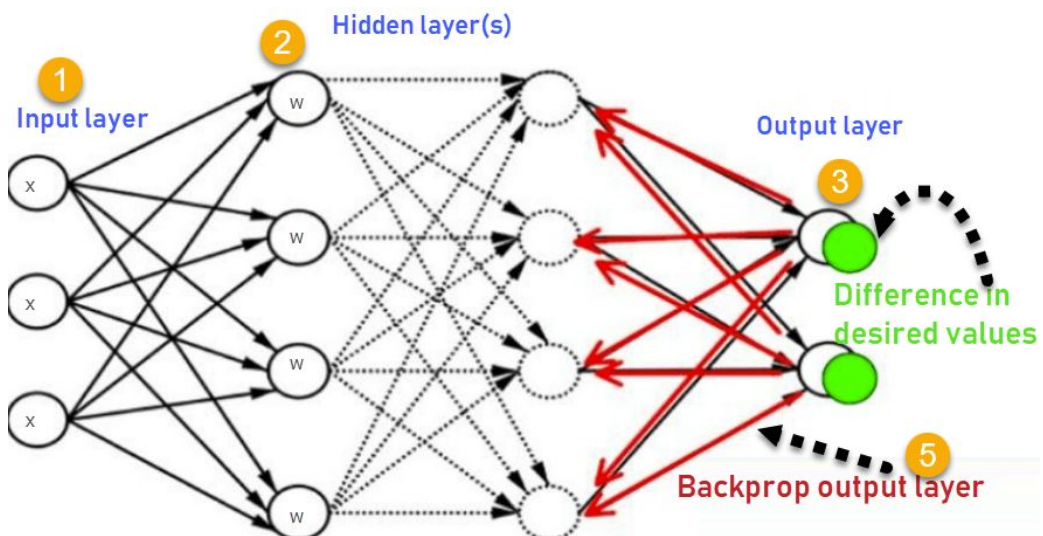


Several layers of Artificial Neural Networks

## Backpropagation

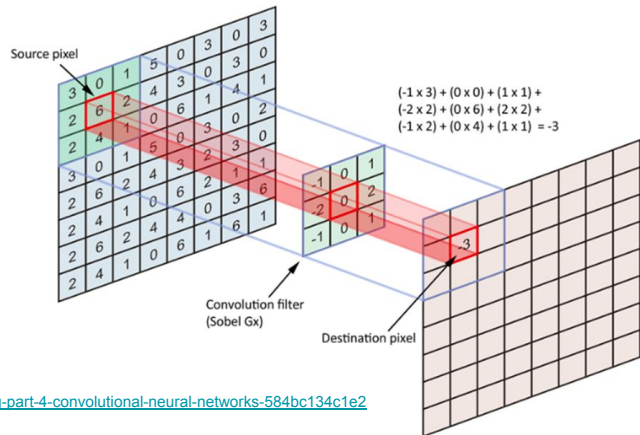
- Propagación hacia atrás de errores o retropropagación

David E. Rumelhart, Geoffrey E. Hinton & Ronald J. Williams. *Learning representations by back-propagating errors*. Nature volume 323, pages533–536(1986)



## Redes convolucionales

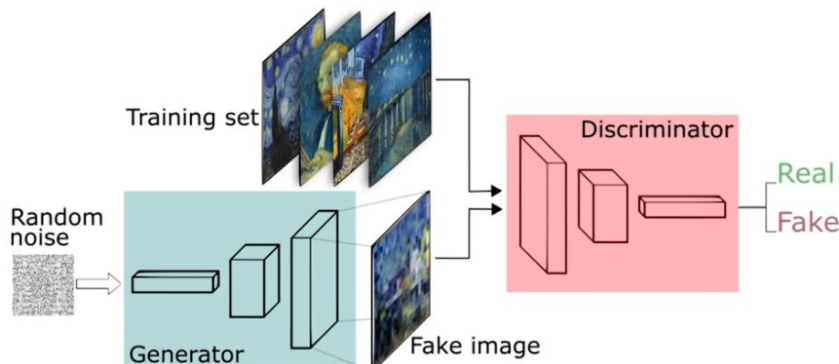
- Convolutional Neural Networks (CNN): Capas con operaciones de convolución
  - Ventana de percepción sobre la que actúan filtros convolucionales para extraer mapas de características y patrones locales
  - Una neurona trata una pequeña “ventana” de la capa anterior
  - Las neuronas de una capa hacen todas el mismo tratamiento y con los mismos parámetros (peso y sesgo/umbral) pero cada una sobre su ventana



<https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2>

## Redes Neuronales Generativas Adversarias

- Generative Adversarial Networks (GAN) permiten reducir aún más la participación humana
  - Una red neuronal artificial como oponente que entrena (juzga) a otra
  - Ej. Generación de pinturas de estilo impresionista a partir de una foto
    - Una red generadora crea cuadros impresionistas a partir de imágenes
    - Otra red, discriminadora, juzga si lo que ha creado la otra es o no un cuadro impresionista
    - Al principio no son muy impresionistas los cuadros pero a medida que la generadora va aprendiendo, la discriminadora va teniendo más difícil su tarea de corregir





## Tensor Flow

- Google Machine Learning Tools (2ª generación)
- Biblioteca de métodos de ML para construir aplicaciones inteligentes
  - Deep learning y otros algoritmos de ML (clasificación, regresión, SVM)
  - Otras: Theano, Torch, CNTK (Microsoft), MXNet (Apache)
- Sistema de programación que representa las computaciones como grafos de flujos de datos, y que se puede ejecutar en sistemas de computación distribuida y paralela
  - Los **nodos** en el grafo realizan operaciones
  - Las conexiones, **tensores**, representan datos estructurados como arrays n-dimensionales

- Se puede probar (sin necesidad de programar):

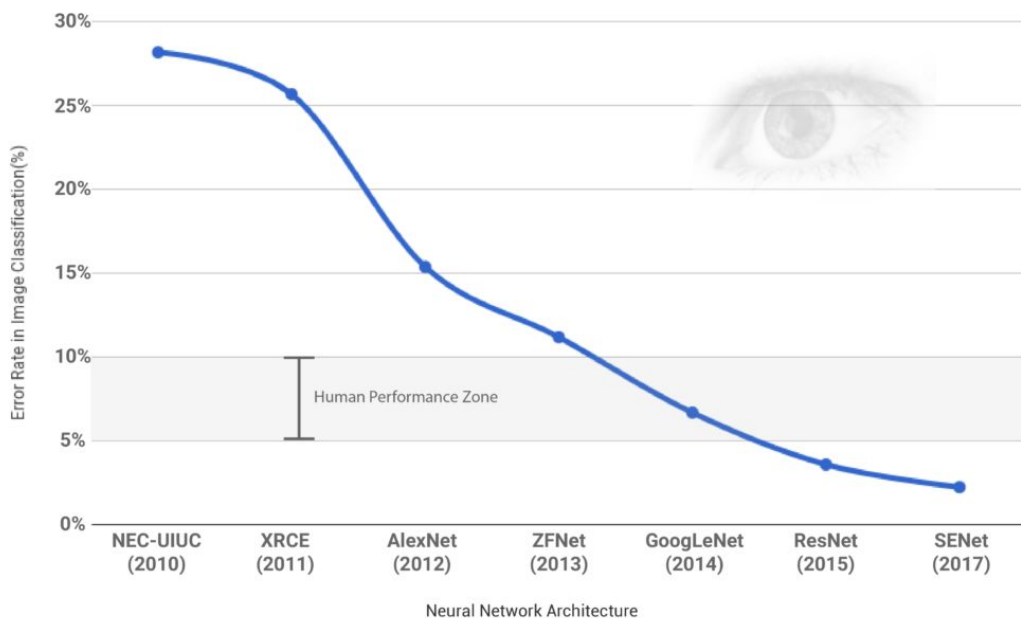
<https://playground.tensorflow.org/>



TensorFlow

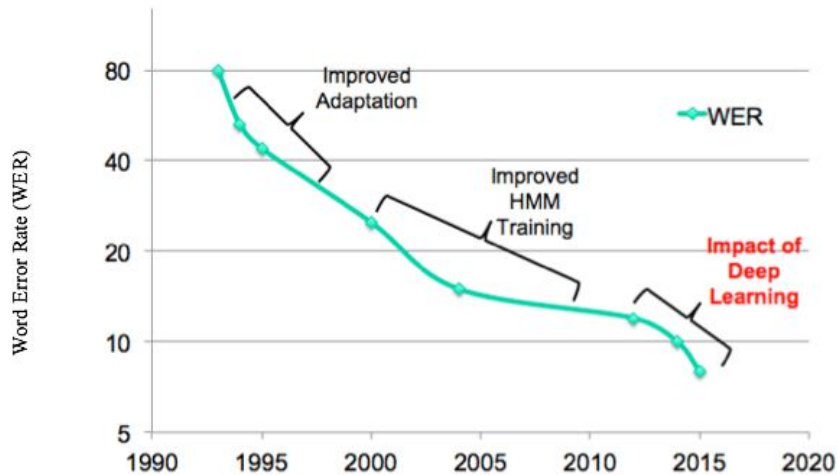


## Reconocimiento de imágenes



## Reconocimiento de voz

- Word Error Rate (%)
  - Google Assistant < 5%
  - Microsoft ~ 5%
  - Human error 2-4%



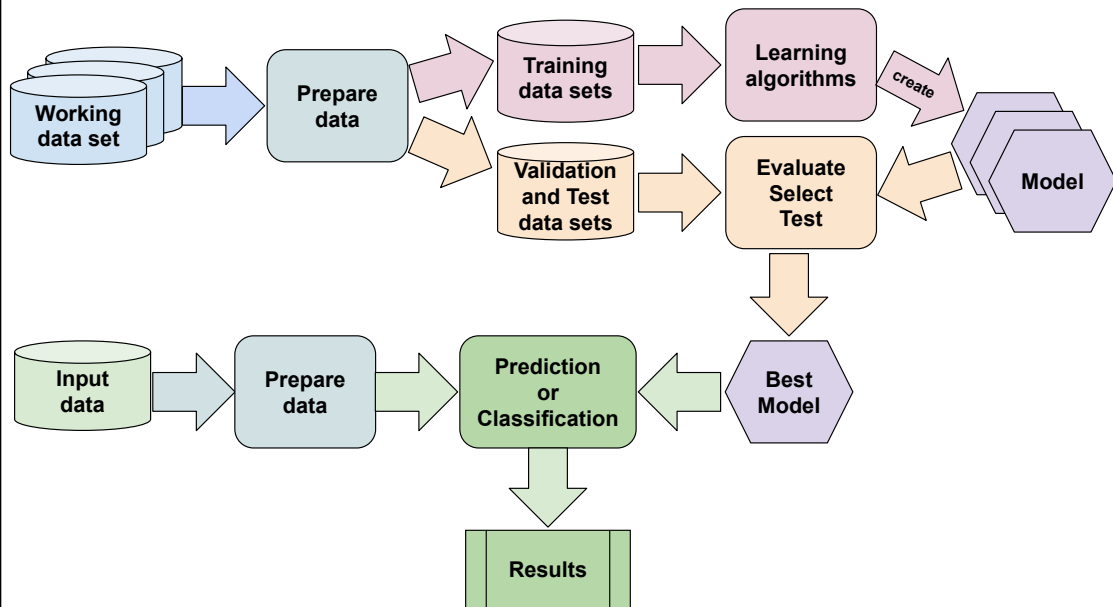
## Deep learning - El caso del bot de Microsoft



What capacity of ethical judgment or reflection on knowledge and the context should have artificial agents?

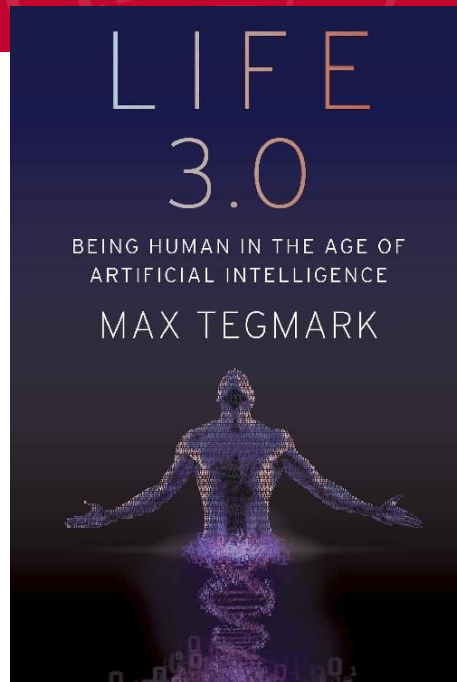
*Si Tay intentó hacerse "colega" de otros intervinientes en las redes sociales y supongamos por un momento que no hubo errores de diseño y Tay se limitó a seguir los comentarios más populares "de la manada". La pregunta es más incómoda: ¿Qué aprendió Tay de nosotros? ¿Qué tipo de opinión social - al menos la que se expresa en redes sociales - estamos construyendo? ¿Qué vemos en el espejo de Tay?* [Javier Busquets, 2016]

## Desarrollo de aplicaciones de Machine Learning



**DASI**

**Futuro**



*Artificial intelligence may enable us to launch Life 3.0 this century*

Can it design its hardware?			 
Can it design its software?		 	 
Can it survive & replicate?	 	 	 
	Life 1.0 (simple biological)	Life 2.0 (cultural)	Life 3.0 (technological)

Max Tegmark (2017). Life 3.0: being human in the age of artificial intelligence

## Life 3.0

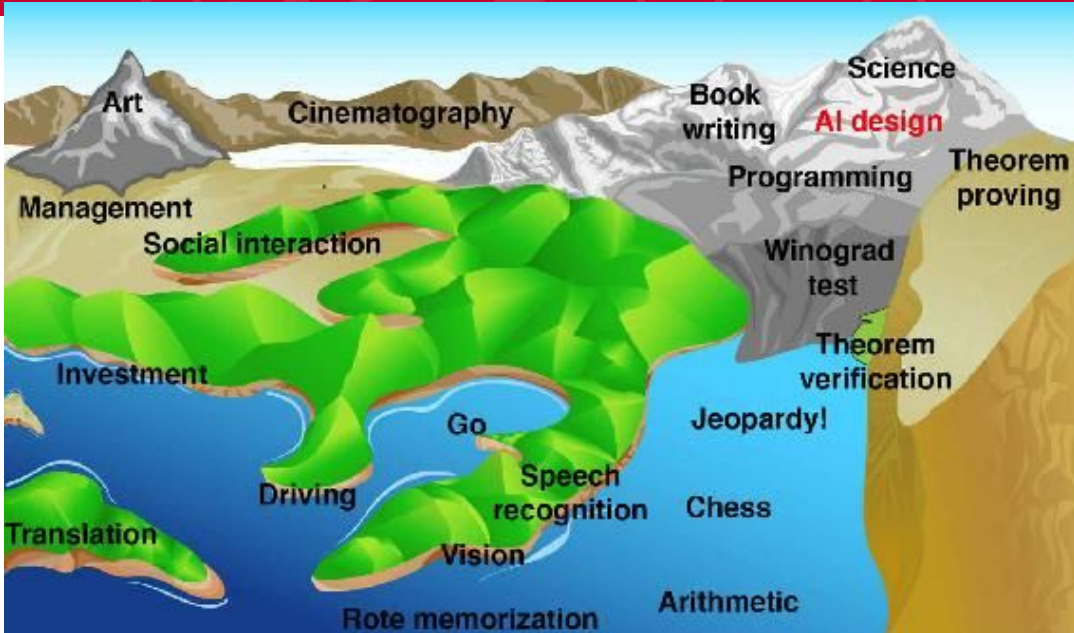


Illustration of Hans Moravec's "landscape of human competence," where elevation represents difficulty for computers, and the rising sea level represents what computers are able to do  
Max Tegmark (2017). Life 3.0: being human in the age of artificial intelligence



## Niveles de IA

- **Inteligencia**
  - Habilidad para realizar tareas
  - **Inteligencia General:** Habilidad para lograr virtualmente cualquier objetivo, incluido el aprendizaje
- **IA débil (narrow AI)**
  - Objetivos concretos: jugar al ajedrez o al go, conducir un vehículo autónomo como un coche o un dron, clasificar imágenes, reconocer la voz, etc.
- **IA fuerte (strong AI)**
  - [nivel humano] **IA General (Artificial General Intelligence, AGI):** realizar cualquier tarea cognitiva al menos tan bien como los humanos
- **Superinteligencia**
  - Inteligencia general más allá del nivel humano
  - **Singularidad:** Explosión de inteligencia (una auto-mejora recursiva que llevaría rápidamente a una superinteligencia)





## Super-inteligencia

*¿Qué ocurrirá cuando los sistemas inteligentes sean más inteligentes que los humanos?*

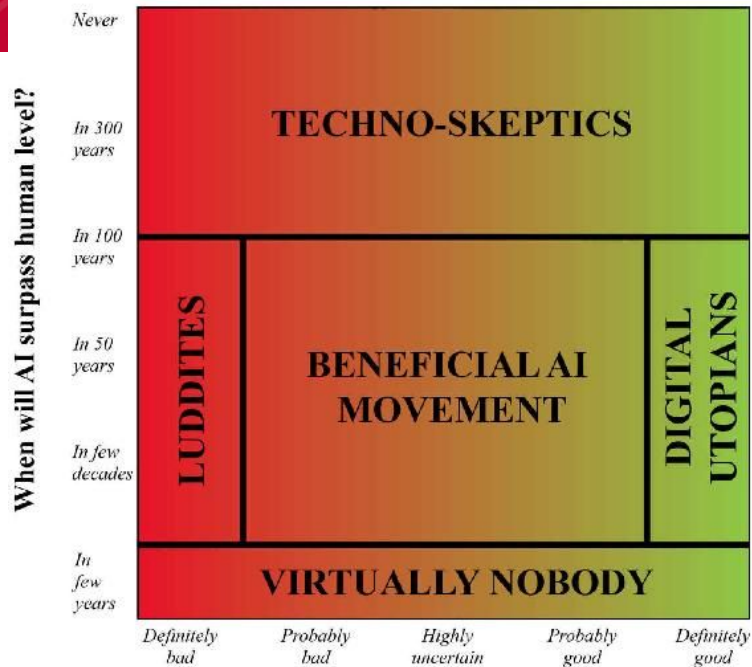


J Pavón (UCM)

DASI - Introducción a la IA

77

## Life 3.0



**If superhuman AI appears, will it be a good thing?**

Max Tegmark (2017). Life 3.0: being human in the age of artificial intelligence



J Pavón (UCM)

DASI - Introducción a la IA

78

## Super-inteligencia

¿Qué ocurrirá cuando los sistemas inteligentes sean más inteligentes que los humanos?

- Sam Harris: Can we build AI without losing control over it?  
[https://www.ted.com/talks/sam\\_harris\\_can\\_we\\_build\\_ai\\_without\\_losing\\_control\\_over\\_it](https://www.ted.com/talks/sam_harris_can_we_build_ai_without_losing_control_over_it)
- Nick Bostrom: What happens when our computers get smarter than we are? [TED, 2015]  
[https://www.ted.com/talks/nick\\_bostrom\\_what\\_happens\\_when\\_our\\_computers\\_get\\_smarter\\_than\\_we\\_are](https://www.ted.com/talks/nick_bostrom_what_happens_when_our_computers_get_smarter_than_we_are)
- Jeremy Howard: The wonderful and terrifying implications of computers that can learn [TEDxBruussels, 2014]  
[https://www.ted.com/talks/jeremy\\_howard\\_the\\_wonderful\\_and\\_terrifying\\_implications\\_of\\_computers\\_that\\_can\\_learn](https://www.ted.com/talks/jeremy_howard_the_wonderful_and_terrifying_implications_of_computers_that_can_learn)



## Lethal autonomous weapons

<http://autonomousweapons.org/slaughterbots/>



## Iniciativas éticas y legales

- *Barcelona declaration for the proper development and usage of AI in Europe* <https://www.iiia.csic.es/barcelonadeclaration/>
  - Code of conduct for AI practitioners:
    - Prudence
    - Reliability
    - Accountability
    - Responsibility
    - Constrained autonomy
    - Human role
- *Proposal for a Regulation laying down harmonised rules on artificial intelligence*
  - The Commission is proposing the first ever legal framework on AI, which addresses the risks of AI and positions Europe to play a leading role globally.
  - <https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence>

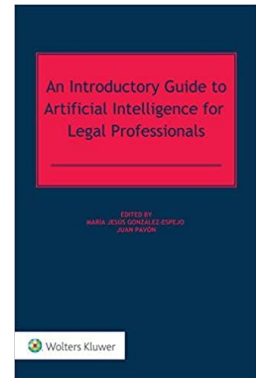
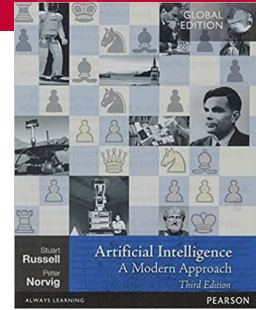
## Algunas conclusiones

- La IA es ubicua
  - Móviles, entornos inteligentes (IoT), industria 4.0, transporte, etc.
- La frontera de lo que es IA y simple computación es difusa
- El aprendizaje automático (machine learning) es el paradigma predominante hoy día, pero plantea nuevas cuestiones
  - Transparencia
  - Explicabilidad
  - Validación de datos y algoritmos
- Cada vez dependemos más de sistemas inteligentes
- La ética y la legislación como requisitos fundamentales en la ingeniería
- Faltan muchos muchísimos profesionales (inter-disciplinares) para desarrollar, mantener y evolucionar los sistemas inteligentes
  - Aunque también decaerán otras profesiones, o las transformarán

## Bibliografía básica

- S. J. Russell y P. Norvig (2016). **Artificial Intelligence: A Modern Approach. Third Edition.** Prentice Hall, 2016
- M.J. González Espejo y J. Pavón (eds., 2020). **An Introductory Guide to Artificial Intelligence for Legal Professionals.** Wolters kluwer.
- Max Tegmark (2017). **Life 3.0: Being Human in the Age of Artificial Intelligence.** KNOFF.
- Canal Dot CSV:  
<https://www.youtube.com/channel/UCy5znSnfMsDwaLIROnZ7Qbg>

⇒ El mejor en divulgación de IA, y en español



## Bibliografía

- ANA MAS: Agentes software y sistemas multiagente: Conceptos, arquitecturas y aplicaciones, Pearson – Prentice Hall, 2005.
- Sridhar Alla & Suman Kalyan Adari: *Beginning MLOps with MLFlow. Deploy Models in AWS SageMaker, Google Cloud, and Microsoft Azure.* Apress, 2021.
- Stuart Russell, Peter Norvig: *Artificial Intelligence: A Modern Approach, 3rd edition.* Prentice Hall, 2016. Teik Toe Teoh & Zheng Rong: *Artificial Intelligence with Python.* Springer Nature, 2022.
- Yuxi (Hayden) Liu: *Python Machine Learning By Example, Third Edition.* Packt Publishing, 2020.
- Documentación en línea de las distintas herramientas

**Podéis encontrar mejor bibliografía y publicarla en el campus virtual**

