

BIG DATA E INTELIGENCIA ARTIFICIAL

Lali Soler
Big Data and Data Science Unit Director
Eurecat

Innovant amb les empreses

QUÉ ES BIG DATA?

Big Data



Big Data

Estamos convirtiendo en datos muchos aspectos del mundo que nunca se han cuantificado antes:

business networks

books I'm reading

location

physical activity

consumed food

purchases

physiological signals

straight thoughts

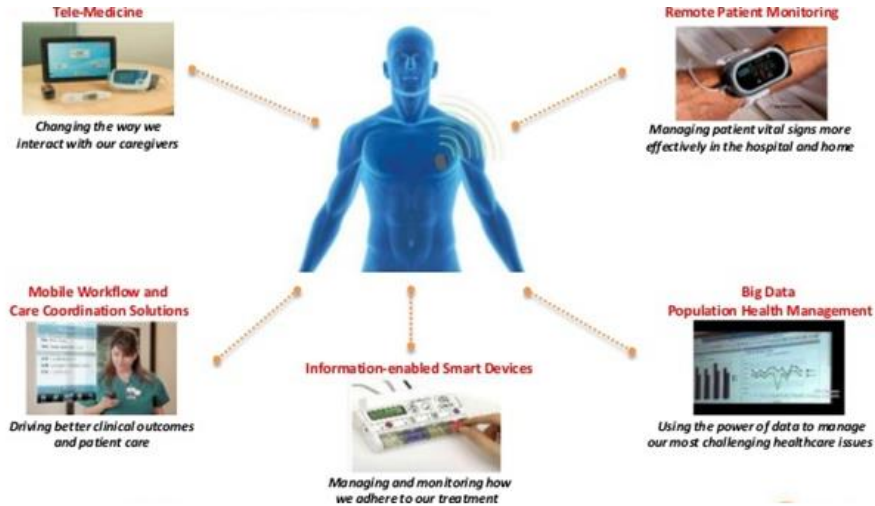
friendship

gaze

driving behavior

Big Data

Wal-Mart maneja más de un millón de transacciones de clientes por hora, la información se almacena en una base de datos con un tamaño superior a 2,5 Petabytes ($2,0 \times 10^{16}$ bits).



En 2016 un hospital típico generaba 665 terabytes (5.32×10^{15} bits) de datos por año.



Big Data

Big Data es más que volumen
Generalmente, se caracteriza por las 4 Vs



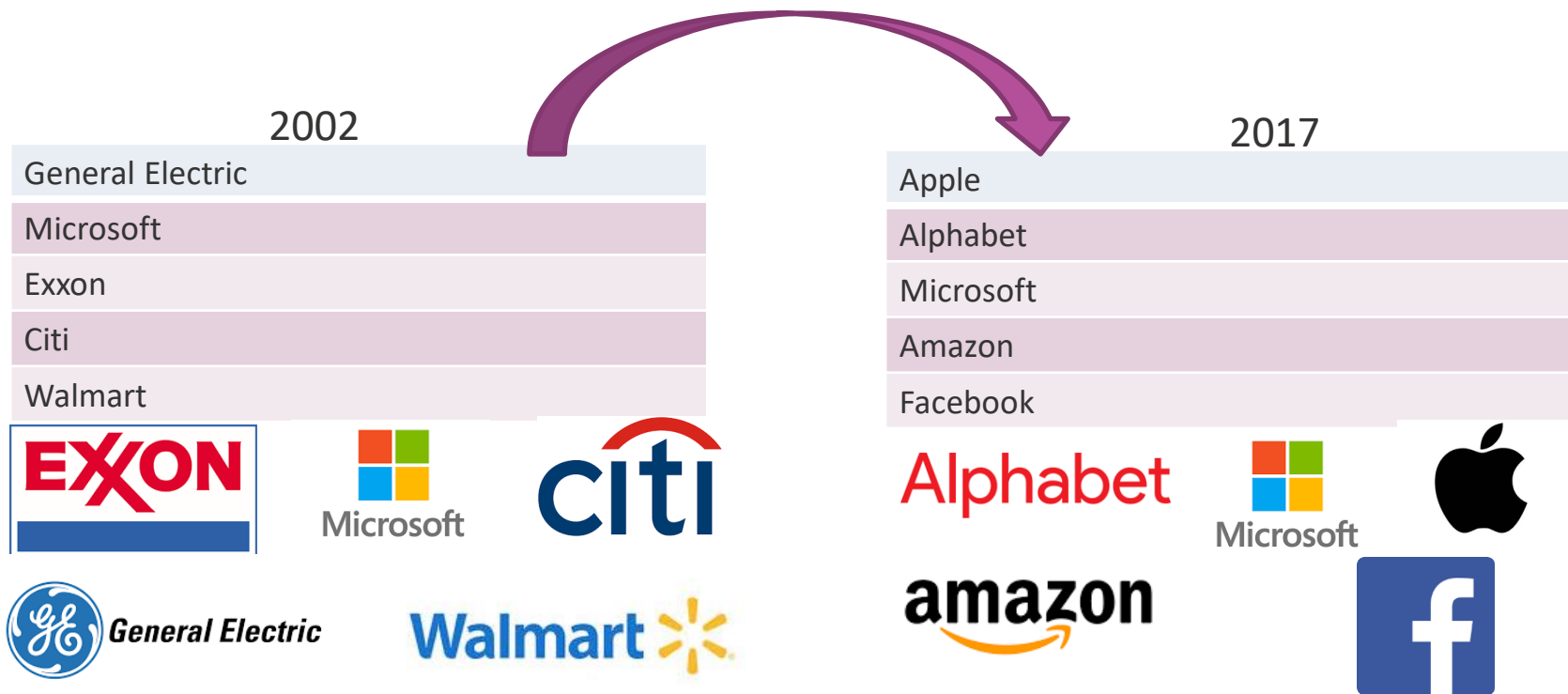
Big Data

La generación masiva de datos en formato digital es lo que se llama **Datificación**. Este fenómeno, unido a las nuevas capacidades tecnológicas de distribución, paralelización y los aceleradores de computación ha propiciado el uso y explotación del Big Data



Mundo Data Driven

Las 5 compañías mejor valoradas en bolsa des de 2002 a 2017:



Data Science



Data Science
Big Data

Data Science



Big Data es una tecnología clave para procesar cantidades masivas de datos (por ejemplo, para contar elementos).



Data Science es una metodología para definir qué queremos hacer con los datos, cómo evaluamos nuestras acciones, qué decisiones pueden basarse en datos, cómo combinamos evidencias de varias fuentes, etc.

QUÉ ES LA INTELIGENCIA ARTIFICIAL

Terminología Básica en Inteligencia Artificial

Conceptos que las empresas deben conocer

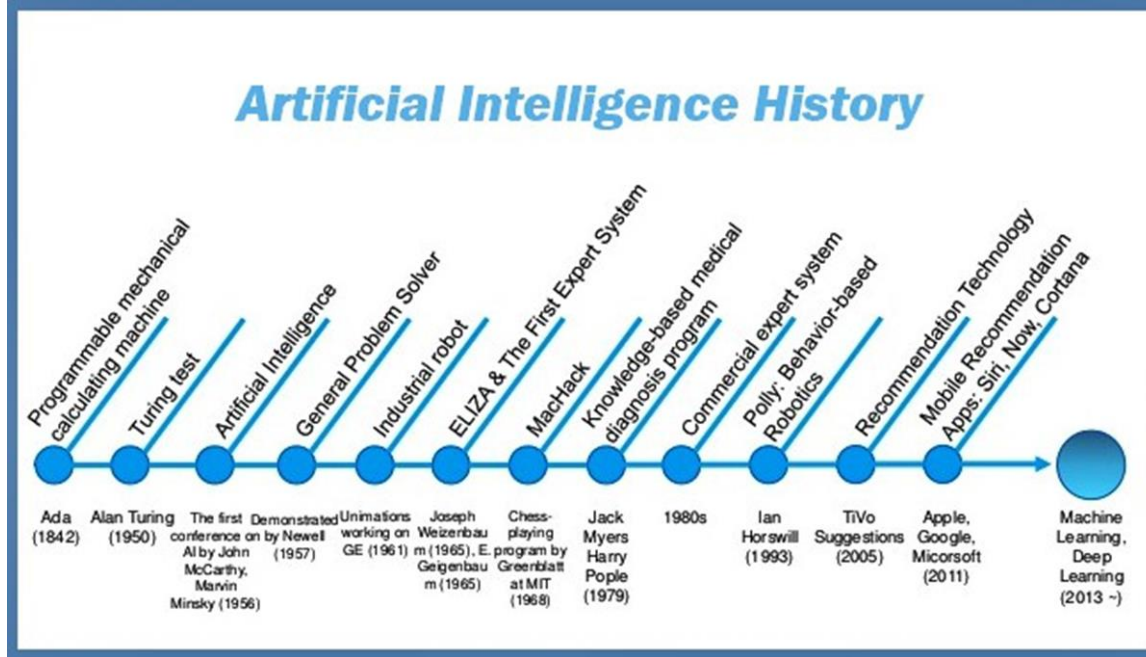
La **inteligencia artificial (IA)**, es la inteligencia exhibida por máquinas. En ciencias de la computación, una máquina «inteligente» ideal es un agente racional flexible que percibe su entorno y lleva a cabo acciones que maximicen sus posibilidades de éxito en algún objetivo o tarea. Coloquialmente, el término inteligencia artificial se aplica cuando una máquina imita las funciones «cognitivas» que los humanos asocian con otras mentes humanas, como por ejemplo: «aprender» y «resolver problemas».



https://www.youtube.com/watch?v=nsPQvZm_rgM

La inteligencia artificial, hoy

La Inteligencia Artificial no es nada nuevo...



AHORA

La historia de la inteligencia artificial no está exenta de altibajos y momentos de olvido. No es hasta la pasada década cuando los avances en capacidad e computación y el fenómeno de **datificación** (o generación de datos digitales) confluyen para crear el entorno óptimo para su desarrollo y aplicación en ámbitos inusuales hasta entonces, como la gestión empresarial.

Machine Learning

Conceptos que las empresas deben conocer: Machine Learning

Atribuir a la máquina la capacidad de aprender es uno de los pilares fundamentales de la inteligencia artificial. El **Machine Learning** es el campo de la ciencia de la computación que se ocupa de enseñar a los ordenadores a desarrollar determinadas tareas sin necesidad de hacer una programación expresa o ad-hoc. Existen varias maneras de abordar el problema:

Aprendizaje Supervisado

Tiene lugar cuando el ordenador recibe datos "etiquetados" que consisten en pares de inputs y outputs (ejemplo: una imagen de un gato correctamente etiquetada como "gato"). De este modo, la máquina aprende reglas generales que le permiten otorgar a cada input su etiqueta correcta. Es un mecanismo muy eficaz para clasificar o predecir valores futuros.

Aprendizaje No Supervisado

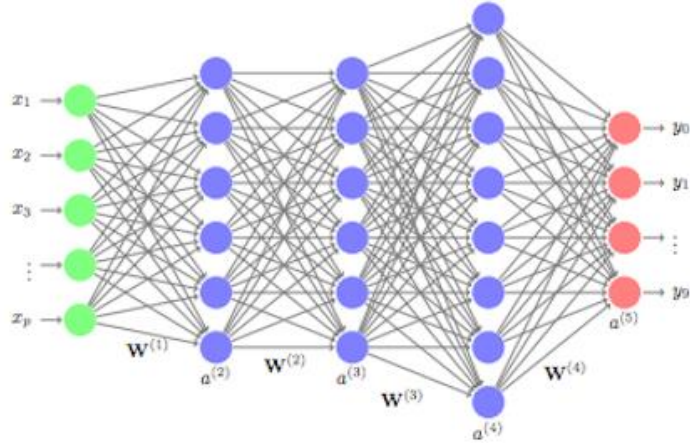
Tiene lugar cuando el ordenador recibe datos desestructurados y no etiquetados y desarrolla la capacidad de descubrir estructuras y patrones inherentes a los datos. Una de las aplicaciones más importantes es el *clustering* cuyo resultado es la obtención de grupos de datos que comparten características similares.

Aprendizaje Reforzado

Se trata de una disciplina que otorga a la máquina la capacidad de aprender mediante una metodología de prueba y error. En este caso, el ordenador es instruido para alcanzar un determinado objetivo en un entorno dinámico. El programa aprende a medida que realiza acciones repetidamente, midiendo el feedback recibido y mejorando su comportamiento de manera iterativa.

Deep Learning

Conceptos que las empresas deben conocer: Deep Learning



El **deep learning** es un sub-campo del **machine learning** cuyos algoritmos de aprendizaje se basan en redes neuronales multicapa. Se trata de estructuras matemáticas que emulan el mecanismo biológico de las conexiones neuronales.

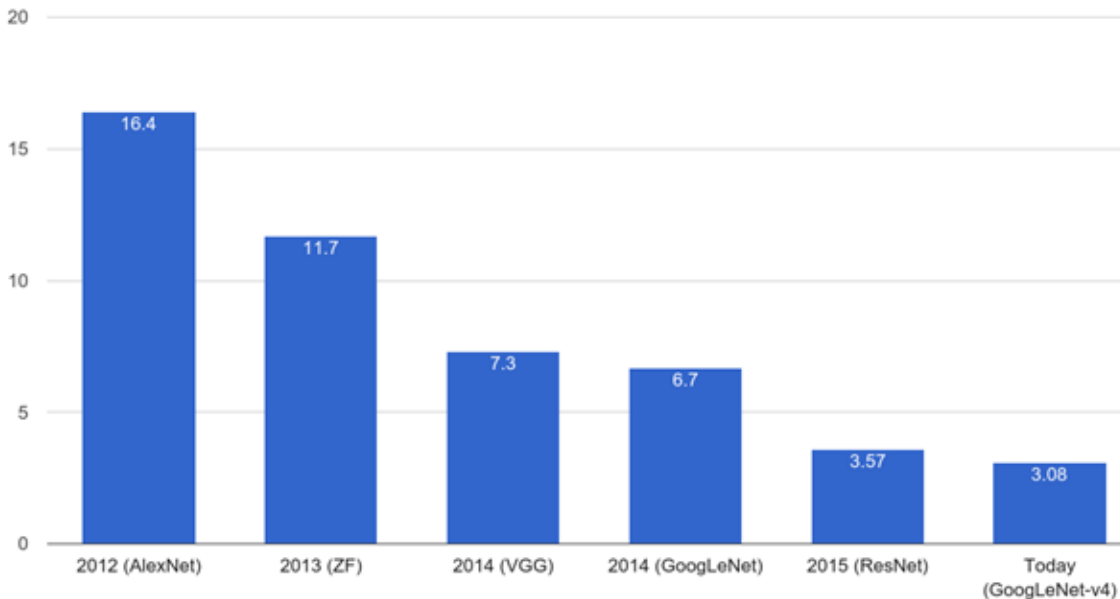
Las redes neuronales aparecieron en 1950, pero los avances tecnológicos en capacidad de computación, unidos al fenómeno del **Big Data** han hecho posible que los algoritmos basados en **deep learning** se aproximen o mejoren el rendimiento humano en campos como el reconocimiento de voz o de imágenes.



El Reto ImageNet

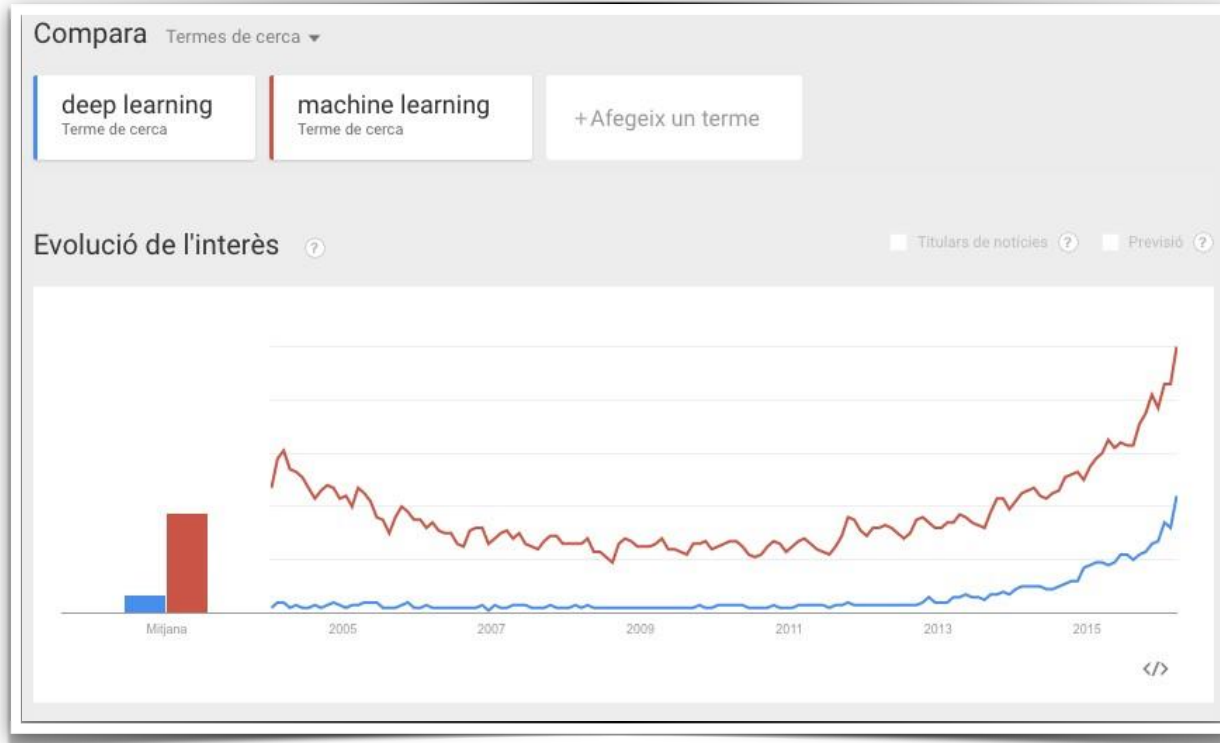
La clasificación y reconocimiento de objetos en imagen ha dado un salto cualitativo enorme en los últimos 10 años, periodo durante el cual el deep learning se ha instalado como técnica hegemónica en sustitución a las técnicas clásicas de tratamiento de imagen (segmentación, extracción de contornos, etc.)

ImageNet Classification Error (Top 5)



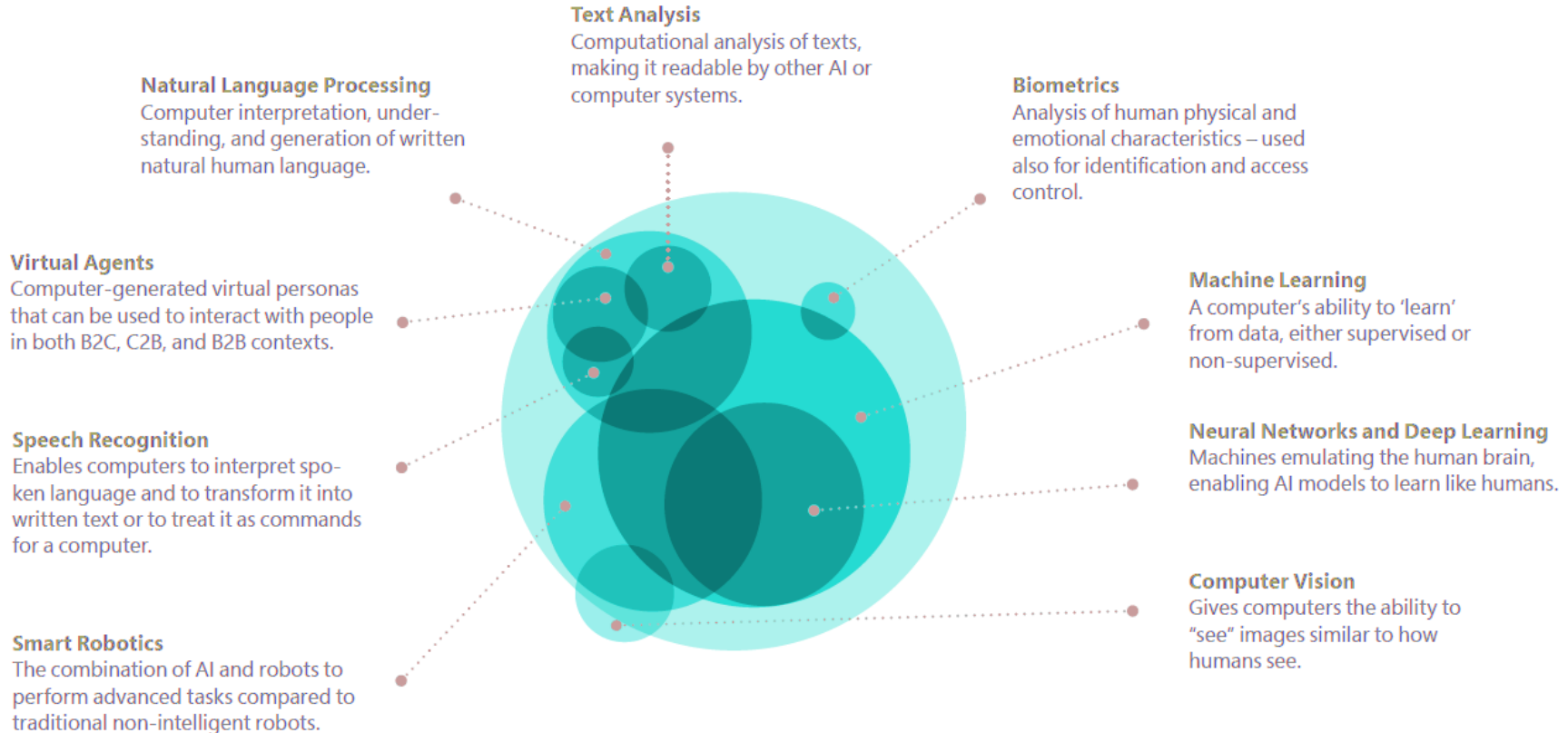
La inteligencia artificial, hoy

Una tecnología de interés global



Google Trends

Campos de la IA



APLICACIONES PIONERAS

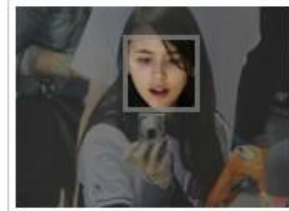
Reconocimiento de imágenes

La inteligencia artificial da solución automática, actualmente, de manera casi perfecta a problemas como el reconocimiento de voz, de imágenes y el procesamiento de lenguaje natural.



Reconocimiento de caras

Reconocimiento de caras



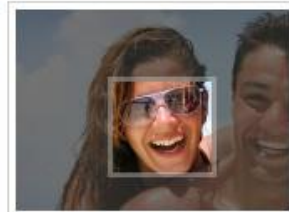
Who is this?



Who is this?



Who is this?



Who is this?



Who is this?

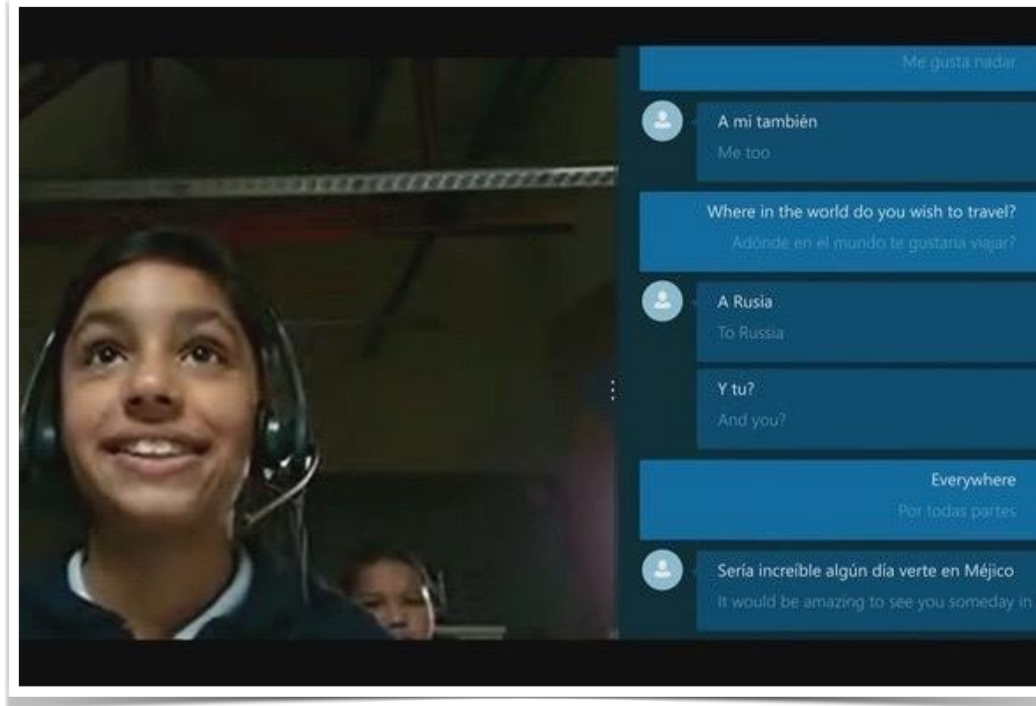


Who is this?

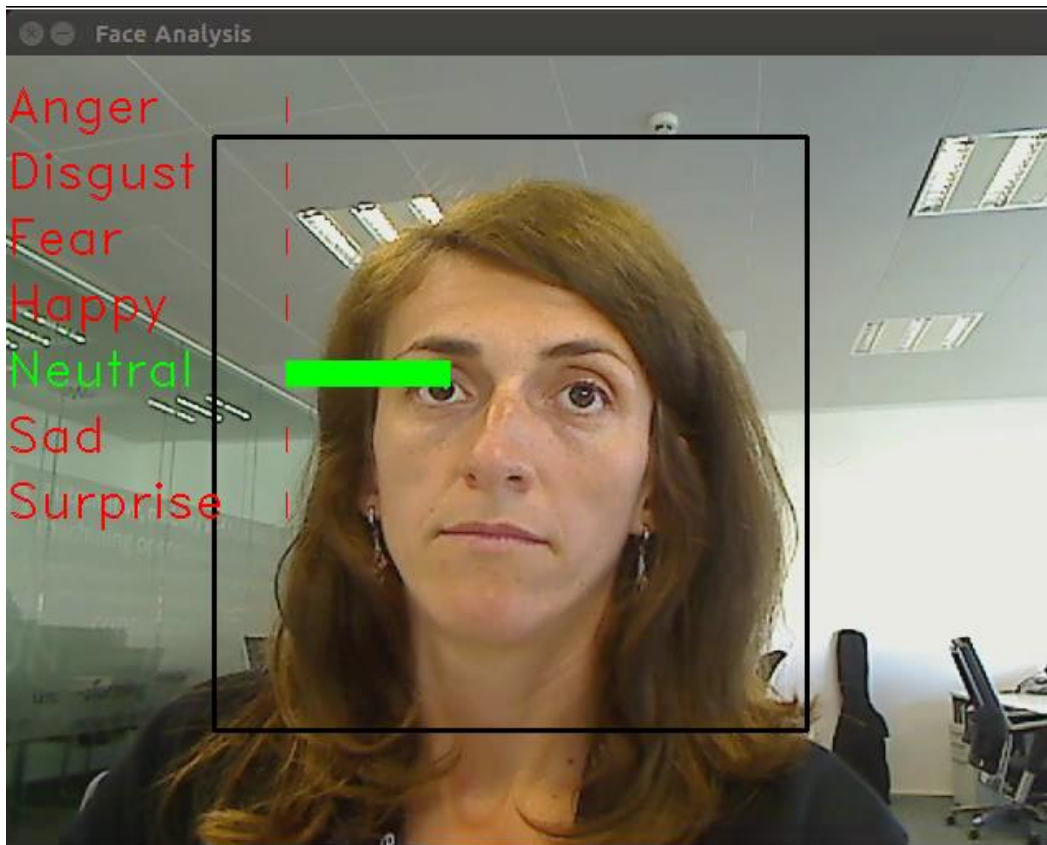
DeepFace (Facebook): Porcentaje de acierto del 97.35%, superior al de un humano entrenado para ello.

2
1

Traducción automática



Reconocimiento de emociones



Lo que la Inteligencia Artificial puede hacer

What Machine Learning Can Do

A simple way to think about supervised learning.

INPUT A	RESPONSE B	APPLICATION
Picture	Are there human faces? (0 or 1)	Photo tagging
Loan application	Will they repay the loan? (0 or 1)	Loan approvals
Ad plus user information	Will user click on ad? (0 or 1)	Targeted online ads
Audio clip	Transcript of audio clip	Speech recognition
English sentence	French sentence	Language translation
Sensors from hard disk, plane engine, etc.	Is it about to fail?	Preventive maintenance
Car camera and other sensors	Position of other cars	Self-driving cars

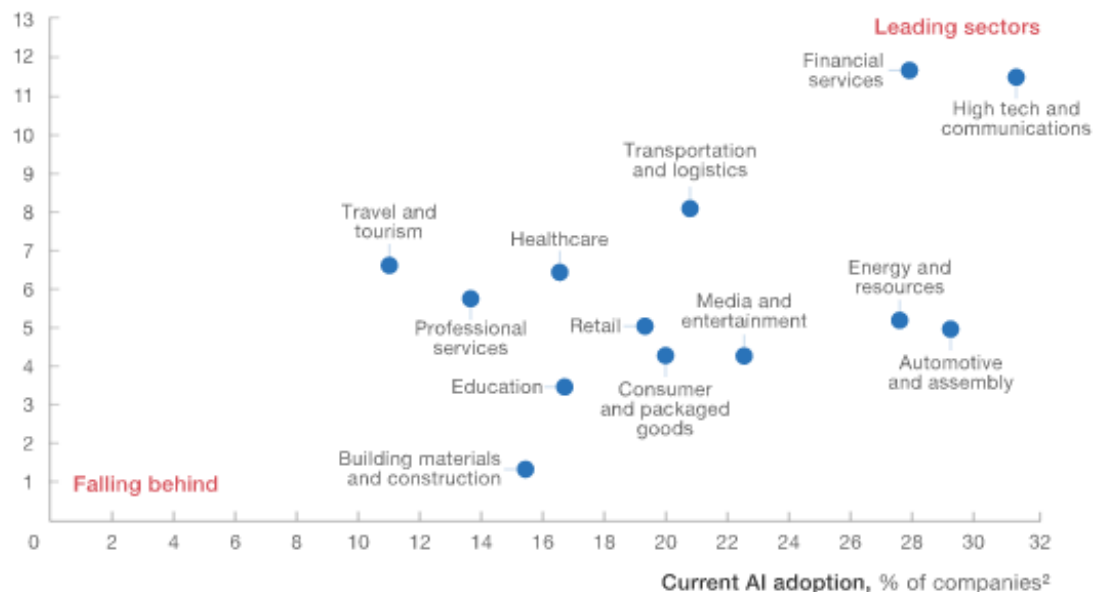
Andrew Ng is the founder and CEO of Landing AI, the former VP & Chief Scientist of Baidu, Co-Chairman and Co-Founder of Coursera, the former founder and lead of Google Brain, and an Adjunct Professor at Stanford University.

Sorprendentemente, a pesar de la amplitud del impacto de la IA, los tipos de despliegue que se están implementando todavía son extremadamente limitados. Casi todo el progreso reciente de la IA es a través de un tipo, en el que algunos datos de entrada (A) se utilizan para generar rápidamente alguna respuesta simple (B). Esto es, **aprendizaje supervisado**.

El software de aprendizaje supervisado de hoy tiene **un talón de Aquiles: requiere una gran cantidad de datos**. Se debe mostrar al sistema muchos ejemplos de A y B. Por ejemplo, crear un etiquetador de fotos requiere de decenas a cientos de miles de imágenes (A), así como etiquetas o etiquetas que le informen si hay personas en ellas. La construcción de un sistema de reconocimiento de voz requiere decenas de miles de horas de audio (A) junto con las transcripciones (B).

Sectores pioneros

Future AI demand trajectory, % change in AI spending over next 3 years¹



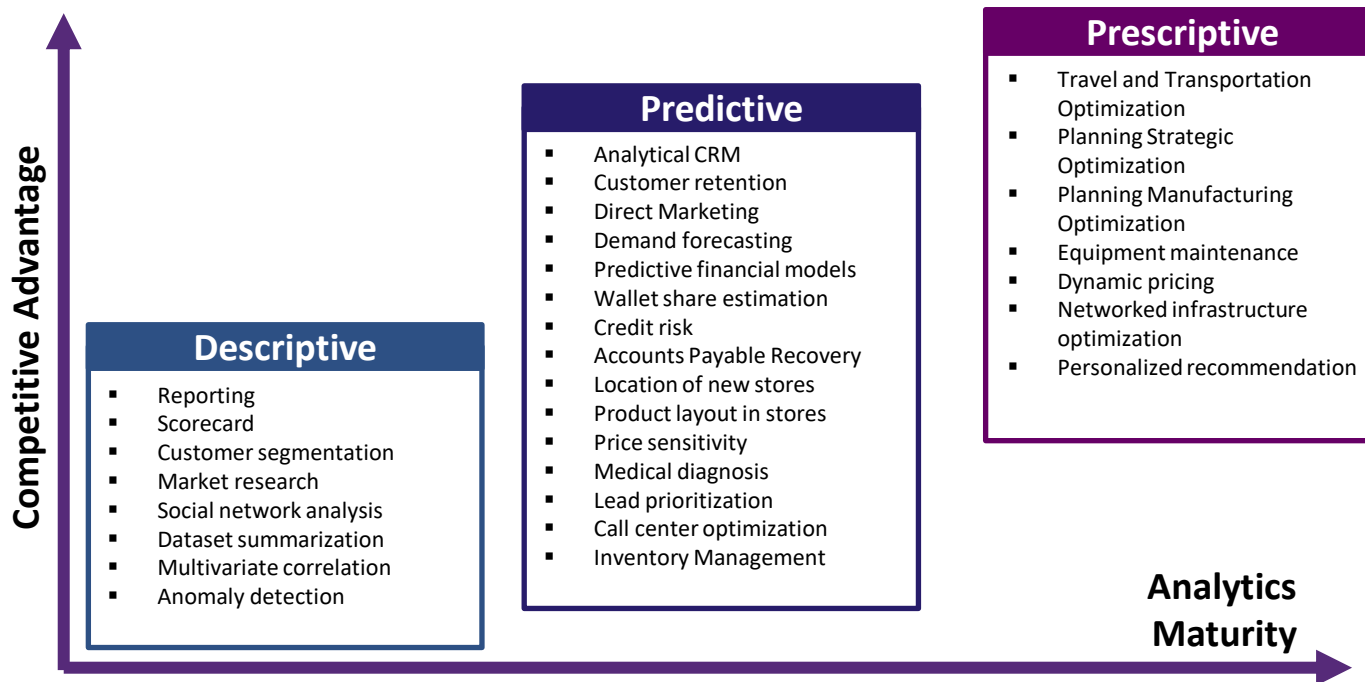
¹Estimated average, weighted by company size; demand trajectory based on midpoint of range selected by survey respondent.

²Adopting 1 or more AI technologies at scale or in business core; weighted by company size.

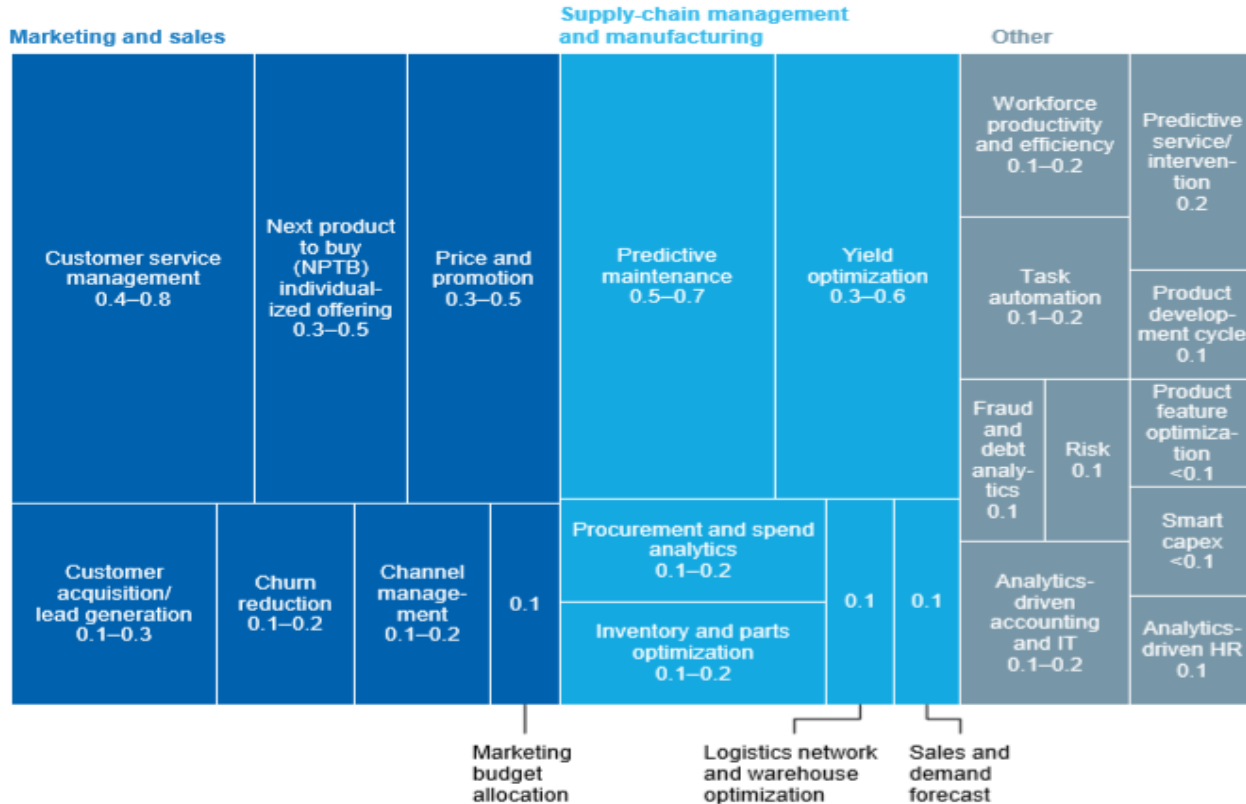
Source: McKinsey Global Institute AI adoption and use survey; McKinsey Global Institute analysis

APLICACIONES DEL BIG DATA Y LA INTELIGENCIA ARTIFICIAL EN LAS EMPRESAS

Los 3 estadios del análisis de datos



Ámbitos de aplicación



SOURCE: McKinsey
Global Institute analysis,
April 2018

GRÀCIES

