

# Tipos de Inteligencia Artificial. Escuelas y clasificaciones.

## Caso práctico



[Wirestock \(CC BY-SA\)](#)

Después de la visita al centro logístico que Miguel ha hecho junto a su clase ha quedado fascinado por la cantidad de información y tareas que son controladas y supervisadas desde el ordenador central. La conversación que ha tenido al final de la misma con Dara, la directora del centro de control, le ha despejado algunas dudas pero al mismo tiempo le ha generado otras muchas cuestiones que hasta ahora no se planteaba. No puede evitar comentar con su

amiga Eva en el viaje de regreso al instituto.

— Eva ¿Te has dado cuenta de que estamos rodeados de dispositivos que usan Inteligencia Artificial? La mitad de las aplicaciones que tengo en el móvil llevan algo de IA — le dice a su compañera.

— ¡Ya te digo! Desde el texto predictivo del teclado, hasta el clasificador de Spam del correo electrónico. Mi padre flipa con la herramienta de desbloqueo de su móvil nuevo por reconocimiento de imagen. ¿Tú crees que esa también está hecha con Inteligencia Artificial? — pregunta Eva.

— Pues seguramente... — dice Miguel pensativo. Y añade — Es que, como nos ha dicho Dara en el centro logístico, hay muchos tipos de Inteligencia Artificial. Lo importante es conocerlos para poder entender cómo funciona y qué posibilidades tiene cada programa o dispositivo basado en ella.

— Hay que investigar, Miguel. Nos tenemos que poner al día. ¿Quedamos esta tarde para buscar información? — le propone entusiasmada Eva.

— ¡Venga, va! — confirma Miguel — A ver qué somos capaces de encontrar.

En la actualidad la evolución de la Inteligencia Artificial comprende un campo tan amplio, con tantas ramas, que es complicado poder atender a una única clasificación. De hecho encontramos diferentes clasificaciones según la diferente visión con la que se aborda dicha tecnología. Las hay más filosóficas, más técnicas, y según su aplicación. A lo largo de este tema vamos a conocer las más comunes y aceptadas.

Pero antes de entrar en materia es importante volver a insistir en que nos encontramos en un nivel de desarrollo de la Inteligencia Artificial bajo. Sabemos, o nos imaginamos, que en algún momento esta ciencia nos permitirá resolver problemas o ejecutar tareas que ahora mismo no puede. Así pues, dentro de las distintas clasificaciones que vamos a ver, hay algunas tipologías que están definidas "en teoría" aunque aún no exista ninguna aplicación IA de esa clase.

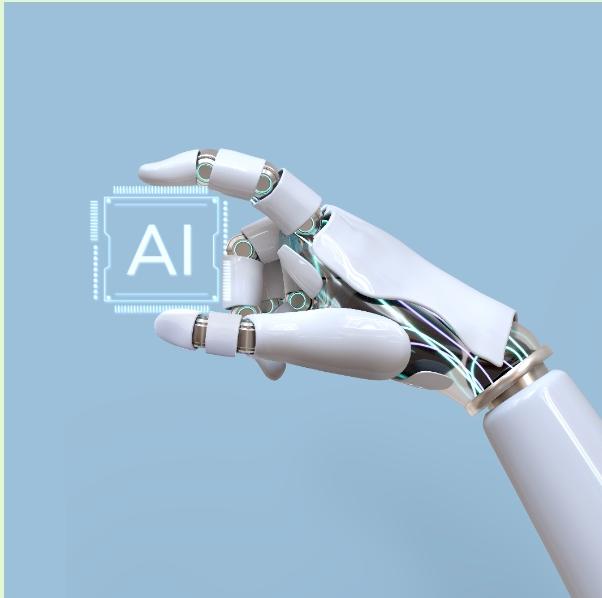
### Tipos de Inteligencia Artificial según distintas clasificaciones



Fran Bartolomé - Elaboración propia ([CC BY-SA](#))

# 1.- Según Tareas a resolver.

## Caso práctico



[Rawpixel \(CC BY-SA\)](#)

de la siesta.

Eva se puso a investigar y encontró una clasificación muy sencilla para diferenciar tipos de Inteligencia Artificial.

- ¡Según el tipo de tareas que resuelve una IA! – dijo triunfante. Inmediatamente llamó a Miguel.
- Escucha, Miguel. Se supone que las máquinas del futuro que serán capaces de hacer muchas tareas distintas se considerarán Inteligencia Artificial Fuerte. – Dijo en cuanto Miguel respondió a la llamada sin dejarle apenas decir "hola".
- ¿Cómo? ¿de qué hablas? – farfulló Miguel – Me has despertado

– Que según la cantidad de tareas que sea capaz de hacer una misma IA se consideran dos tipos: Débil o Fuerte – Dijo más calmada Eva – . Es la primera forma de clasificación de la Inteligencia Artificial que se ha hecho.

La primera clasificación de la Inteligencia Artificial que vamos a ver se basa en qué tipo de tareas nos ayuda a resolver o ejecutar.

Cuando hablamos de tarea a resolver nos referimos, por ejemplo, a si pretendemos que la IA sea capaz de jugar al ajedrez en un ordenador o si pretendemos que sea capaz de gestionar una cocina sin intervención humana (desde el abastecimiento de alimentos, procesado, decisión de qué cocinar en cada momento, limpieza y mantenimiento de utensilios, resolver imprevistos o accidentes...). Hay tareas sencillas, concretas, y puntuales que son relativamente sencillas de programar, mientras que hay tareas complejas en las que, además influyen muchos factores exteriores (sentimientos, contexto, moral, ética, creencia religiosa...).

Seguro que eres capaz de ir intuyendo que nuestros programas de Inteligencia Artificial actuales son más bien de los que resuelven tareas en un entorno muy concreto y acotado, mientras que hay que irse a las películas o series de ficción para poder hablar de algún caso de Inteligencia Artificial capaz de actuar en escenarios complejos, cambiantes y con "conciencia".

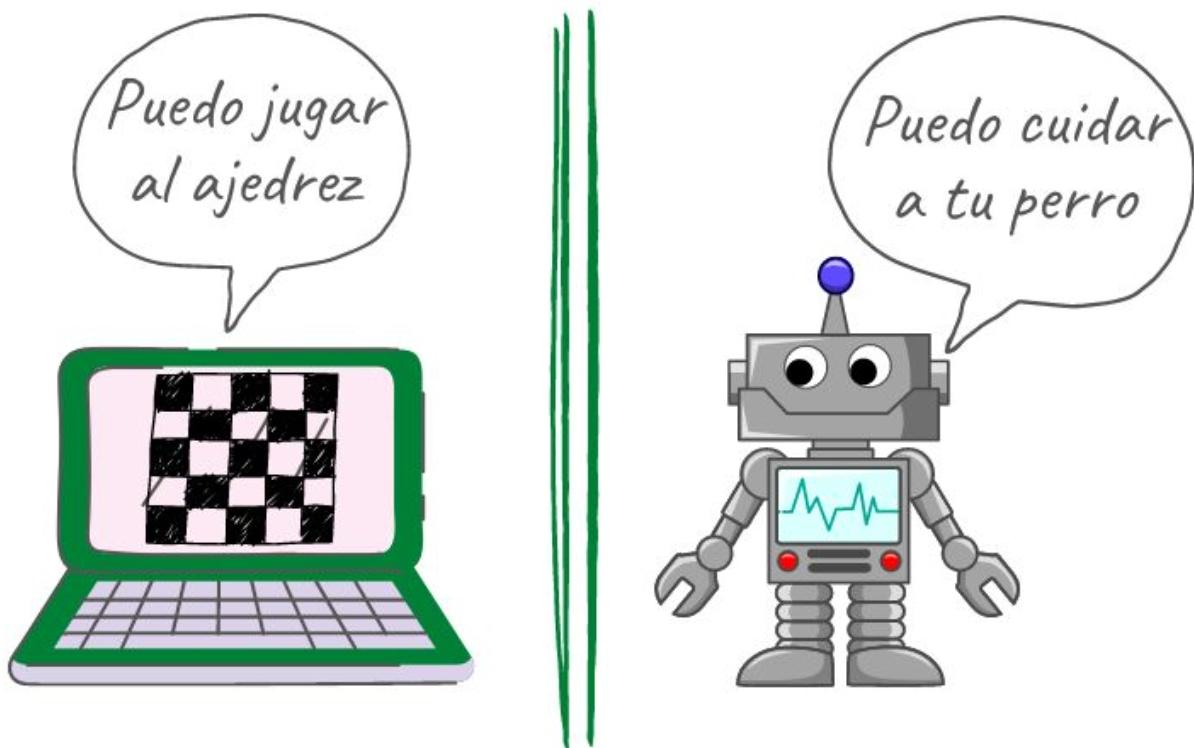
Esta clasificación según tareas es una aproximación bastante simple, con dos opciones:

- ✓ IA Débil.
- ✓ IA Fuerte.

Como veremos en la explicación más detallada de cada una de estas dos posibles categorías de Inteligencia Artificial, en la actualidad aún no hemos sido capaces de desarrollar ninguna IA Fuerte. Todo lo que conocemos en este momento quedaría dentro de la clasificación de IA Débil.

Vamos a ver qué características tienen (o tendrán) las IA según esta clasificación.

#### Tipos de IA según las TAREAS que pueden resolver



Fran Bartolomé - Elaboración propia ([CC BY-SA](#))

## 1.1- Inteligencia Artificial Débil (o estrecha).

centra típicamente en una tarea estrecha concretos, en un entorno perfectamente acotado.

la inteligencia artificial racional que se

Por tanto consideramos que este tipo de Inteligencia Artificial débil es **limitada**, pues no es capaz de adaptarse o asumir cambios respecto a lo que se le ha programado.

Los asistentes virtuales (Siri, Ok Google, Alexa, etc.) son un ejemplo bastante ilustrativo de hasta dónde es capaz de llegar la Inteligencia Artificial débil. Cualquiera de ellos opera dentro de un rango de respuestas limitado, definido en su base de datos. En realidad "la máquina" no tiene inteligencia genuina. No es capaz de aprender, ni de tener en cuenta el entorno o el contexto en el que se le realizan las preguntas. No tiene conciencia, ni mucho menos vida propia. Sin duda es un tipo bastante sofisticado de IA débil, pero llega un punto en el que no es capaz de responder cierta clase de preguntas, y, salvo que cambiáramos su base de datos y programación nunca sería capaz de llegar a encontrar por sí misma respuestas adecuadas a dichas preguntas.



Fran Bartolomé - Elaboración propia ([CC BY-SA](#))

De hecho, uno de los principales entretenimientos más habituales cuando nos ponemos "a charlar" con un asistente virtual es intentar llevarla al límite... A ver en qué tipo de pregunta, cada cual más compleja o absurda, es incapaz de responder. O, sin llegar al límite de no encontrar respuesta, puede llegar a dar respuestas molestas o inadecuadas.

Desde el punto de vista de esta clasificación (por tarea a resolver), las características de la Inteligencia Artificial débil son:

- ✓ Ya existen en la vida real: Como hemos comentado, los asistentes virtuales, programas como Watson o Alpha Go que vimos en la unidad anterior.
- ✓ Se orientan a resolver problemas muy concretos: El programa que "sabe" jugar al Go, no sabe hacer otra cosa. Ni tiene posibilidades de aprender a jugar a otra cosa, por muy similar que sea.
- ✓ Son reactivas: No tienen iniciativa, es necesario que se desencadene la acción que tienen programada para que se inicie su rutina. En el ejemplo del asistente virtual, tiene programado responder cuando le preguntas, y por tanto nunca tomará la iniciativa de ofrecerte nada sin que tú lo actives previamente.
- ✓ No son flexibles: Colapsan si se encuentran en un caso no previsto en su programación.
- ✓ Quedan limitadas por lo que programa un humano: Es el humano quien programa lo que "tiene que pensar" la máquina. Si el humano no programa deja sin considerar ciertas opciones o posibles situaciones, la IA nunca será capaz de suplirlo o aprenderlo sobre la marcha por sí misma.
- ✓ Se programan con pocas redes neuronales: Hablaremos más adelante sobre las redes neuronales. Por el momento es suficiente entender que el nivel de computación compleja que requieren este tipo de Inteligencias Artificiales es menor que otros casos.
- ✓ No razonan, solo computan: No tienen en cuenta ningún factor moral, contextual, circunstancial, emocional... que a un humano le haría reaccionar de manera diferente. La máquina está programada para alcanzar tal objetivo o funcionar de tal manera, y así lo hará sin "entender" lo que está haciendo. Por tanto: no tiene conciencia.
- ✓ Aprenden a base de ejemplos: Necesita conocer muchos ejemplos de lo que tiene que hacer (la base de datos), con todas las variantes posibles. Por ejemplo, en la máquina que juega al Go, se la "entrenó" con 50 millones de partidas de dicho juego.
- ✓ Son repetitivas: No se cansan nunca, son implacables, siempre la misma rutina.
- ✓ No salen de su marco de trabajo: Y esto supone que pueden ocuparse de tareas mecánicas, repetitivas, "aburridas" para sustituir al humano mejorando rendimiento y precisión. Pero necesitará siempre una supervisión humana que vaya decidiendo cómo adaptar el programa a las cambiantes circunstancias.

Según se mire la Inteligencia Artificial débil podría llegar a ser peligrosa. Porque este tipo de IA, al no tomar en consideración todo un contexto amplio, ni seguir las reglas sociales, éticas,... ejecuta las tareas para las que se le ha entrenado con **eficacia y contundencia**. No evalúa las consecuencias como lo hacemos los humanos, considerando un espectro amplio de efectos y relaciones. Por eso, es una opción incompleta, inestable y peligrosa si no se utiliza con prudencia, o si quien la programa pretende, precisamente, causar mal.

# Autoevaluación

Responde para cada afirmación si es verdadera o falsa

La Inteligencia Artificial Débil también se conoce como "estrecha".

Sugerencia

- Verdadero  Falso

Una Inteligencia Artificial Débil que sea capaz de jugar al ajedrez también podría aprender a jugar a las damas.

Sugerencia

- Verdadero  Falso

## 1.2.- Inteligencia Artificial Fuerte.

---

La Inteligencia Artificial fuerte (IAF) o general o (IAG), desde el punto de vista de la tarea a resolver, sería aquella que iguala o excede la inteligencia humana promedio. Sería capaz de realizar con éxito cualquier tarea intelectual del ser humano, teniendo en cuenta todos los factores y matices que pueden intervenir cuando una persona toma decisiones en cada momento mientras realiza una tarea.

No existe todavía ninguna máquina que se pueda considerar completamente IA fuerte. Salvo en algunas películas o series de televisión.

En comparación con la Inteligencia Artificial débil, las características de la fuerte son:

- ✓ No existe en la realidad: Si quieres "ver" cómo sería puedes recurrir a personajes de ficción como [T-800](#), [Wall-E](#) o [J.A.R.V.I.S.](#).
- ✓ Resolverán problemas abiertos: Deberían poder abarcar múltiples posibles tareas, distintas unas de otras (reparar una puerta, ir a recoger a los niños del colegio, regar las plantas, darte conversación...).
- ✓ Serán proactivas: En función de la misión u objetivo que tenga, y de las circunstancias, iniciará cualquier tipo de rutina sin esperar a que un humano se lo pida o esté pendiente.
- ✓ Serán flexibles: Podrán encontrar similitudes entre algo que conocen y algo que se le parezca un poco. Por ejemplo, aunque inicialmente solo haya sido programada para saber andar, será capaz de aprender a correr sin necesidad de intervenir en su programa.
- ✓ Se autoprogramarán: Serán capaces de detectar sus propios límites y se regularán a sí mismas para no excederlos.
- ✓ Usarán muchas redes neuronales: Y además podrán entrar en conflicto entre ellas en algunas ocasiones. Esto quiere decir que necesitarán una capacidad de almacenaje de información y cómputo que aún hoy no hemos llegado a alcanzar.
- ✓ Imitarán el comportamiento humano: Serán capaces de razonar, y por tanto, de alcanzar algún tipo de conciencia.
- ✓ Aprenderán como las personas: Podrán recordar datos, observar nuevas situaciones y encontrar relaciones entre diferentes acciones. Esto quiere decir que si saben jugar al ajedrez y "observan" el juego de las damas, podrán aprender a jugar a las damas basándose en lo que saben sobre jugar al ajedrez.
- ✓ Serán capaces de aprender nuevas tareas: Modificarán la tarea o cómo realizan la tarea para adaptarse a las circunstancias.
- ✓ Serán capaces de adaptarse a nuevos escenarios: Podrán adaptarse a cambios y nuevas situaciones para seguir cumpliendo su objetivo.

## El robot inteligente C3PO de la saga Stars Wars



[Phil Renaud \(CC BY-NC-SA\)](#)

Este tipo de IA es la que sería capaz de analizar cualquier situación y deducir el conjunto de acciones más adecuado para dicha situación y contexto. Lo mismo sabría conducir un coche, que resolver una ecuación matemática o mantener una conversación sobre un tema concreto.

Aunque aún no existe este tipo de IA, todas las empresas e instituciones dedicadas a la investigación y desarrollo de IA están buscando formas de avanzar hacia este tipo de Inteligencia Artificial. De momento, al menos, se está trabajando en conseguir una IAF en el campo de los asistentes virtuales.

Sin duda uno de los ámbitos más ambiciosos de aplicar esta IAF (los asistentes virtuales humanoides) necesitan contar también con otras ramas científicas como son la robótica y mecatrónica.

## 2.- Escuelas de Pensamiento.

### Caso práctico



[Rawpixel \(CC BY-SA\)](#)

— Esto de que haya una clasificación de Inteligencia Artificial que contenga solo dos tipos y que uno de ellos no se haya desarrollado todavía es un rollo — dijo Miguel de forma despectiva.

— Tiene que haber alguna otra forma de clasificar la Inteligencia Artificial algo más emocionante.

— Vaya, pues a mí lo de la IA Débil y Fuerte me parece "super-emocionante", qué quieras que te diga — respondió Eva inmediatamente — . Pero estoy de

acuerdo en que deberíamos profundizar un poco más en esto de la clasificación por si encontramos alguna otra manera de organizar la Inteligencia Artificial.

— Antes de la visita al centro logístico escuché hablar a nuestra profesora sobre las "Escuelas de Pensamiento" de la Inteligencia Artificial — dijo Miguel mirando al techo de la habitación como hablando para sí mismo — ¡Seguro que si investigamos por ahí encontramos algo más!

En el ámbito de la Inteligencia Artificial más moderna podemos encontrar dos escuelas de pensamiento:

- ✓ Inteligencia Artificial Convencional.
- ✓ Inteligencia Artificial Computacional.

Estas dos escuelas difieren en la ciencia que hay tras los procesos que siguen para llegar a los resultados esperados. Pero, con los avances que se están dando en los recursos que utiliza la segunda, muchas de las aplicaciones que tenía la primera, están siendo llevadas al campo computacional.

### Escuelas de Pensamiento en Inteligencia Artificial



Fran Bartolomé (elaboración propia) ([CC BY-SA](#))

## 2.1.- Inteligencia Artificial Convencional.

---

Se conoce también como Inteligencia Artificial simbólico-deductiva. Está basada en el análisis formal y estadístico del comportamiento humano ante diferentes problemas:

- ✓ Razonamiento basado en casos: Ayuda a tomar decisiones mientras se resuelven ciertos problemas concretos y, aparte de que son muy importantes, requieren de un buen funcionamiento.
- ✓ Sistemas expertos: Infieren una solución a través del conocimiento previo del contexto en que se aplica y ocupa de ciertas reglas o relaciones.
- ✓ Redes bayesianas: Propone soluciones mediante inferencia probabilística.
- ✓ Inteligencia artificial basada en comportamientos: Esta inteligencia contiene autonomía y puede auto-regularse y controlarse para mejorar.
- ✓ Smart process management: Facilita la toma de decisiones complejas, proponiendo una solución a un determinado problema al igual que lo haría un especialista en dicha actividad.

Esta rama de la Inteligencia Artificial ha sido la que ha proporcionado la mayoría de algoritmos que conocemos como “**automatización**”, y básicamente se sirven de **sistemas con reglas condicionales y estadística avanzada**.



[Vector Juice \(CC BY-SA\)](#)

# Autoevaluación

Responde si las siguientes afirmaciones son Verdaderas o Falsas

Ya existen máquinas con Inteligencia Artificial Fuerte

Sugerencia

- Verdadero  Falso

La clasificación de la Inteligencia Artificial según las Escuelas de Pensamiento se fijan en los procesos por los que la máquina adquiere dicha inteligencia.

Sugerencia

- Verdadero  Falso

La Inteligencia Artificial Convencional es la más utilizada en la actualidad

Sugerencia

- Verdadero  Falso

## **2.2.- Inteligencia Artificial Computacional.**

---

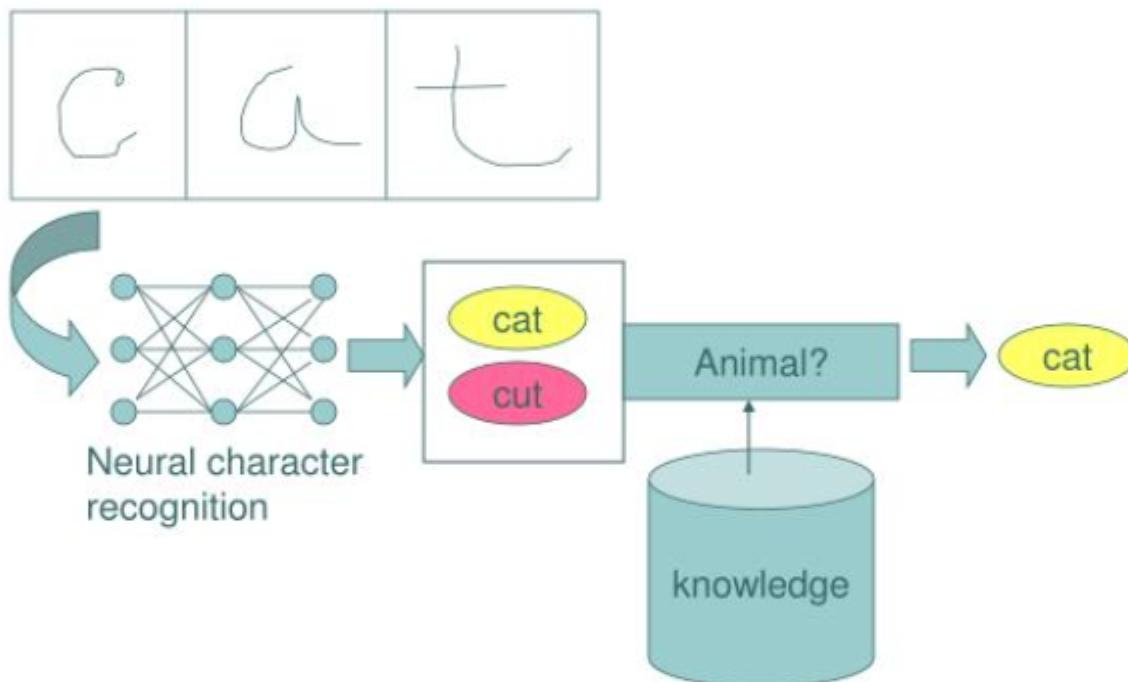
La Inteligencia Computacional (también conocida como IA subsimbólica-inductiva) implica desarrollo o aprendizaje interactivo (por ejemplo, modificaciones interactivas de los parámetros en sistemas de conexiones). El aprendizaje se realiza basándose en datos empíricos, utilizando métodos computacionales inspirados en procesos de la naturaleza, que permiten alcanzar soluciones aptas a problemas complejos que los modelos tradicionales no pueden resolver por no existir una solución analítica, por no contar con todos los parámetros necesarios o porque el problema es en sí estocástico y precisa de una aproximación envolvente en vez de convergente.

Esta corriente ha sido la que impulsó hace pocos años lo que conocemos como “**Aprendizaje Automático**” o “**Machine Learning**”, que es la técnica que más se está utilizando actualmente en desarrollos de IA.

Algunas técnicas de esta escuela son:

- ✓ **Máquina de vectores soporte**: sistemas que permiten reconocimiento de patrones genéricos de gran potencia.
- ✓ **Redes neuronales**: sistemas basados en redes de unidades de computación lineal para simular computación no lineal
- ✓ **Modelos ocultos de Markov**: aprendizaje basado en dependencia temporal de eventos probabilísticos.
- ✓ **Sistemas difusos**: técnicas para lograr el razonamiento bajo incertidumbre
- ✓ **Computación evolutiva**: también conocidos como algoritmos genéticos, aplica conceptos inspirados en la biología, tales como población, mutación y supervivencia del más apto para generar soluciones sucesivamente mejores para un problema.

### Esquema de cómo trabaja una IA Computacional



Carmen Bartolome ([CC BY-SA](#))

### 3.- Hintze.

## Caso práctico



[Michigan State University](#) (Dominio público)

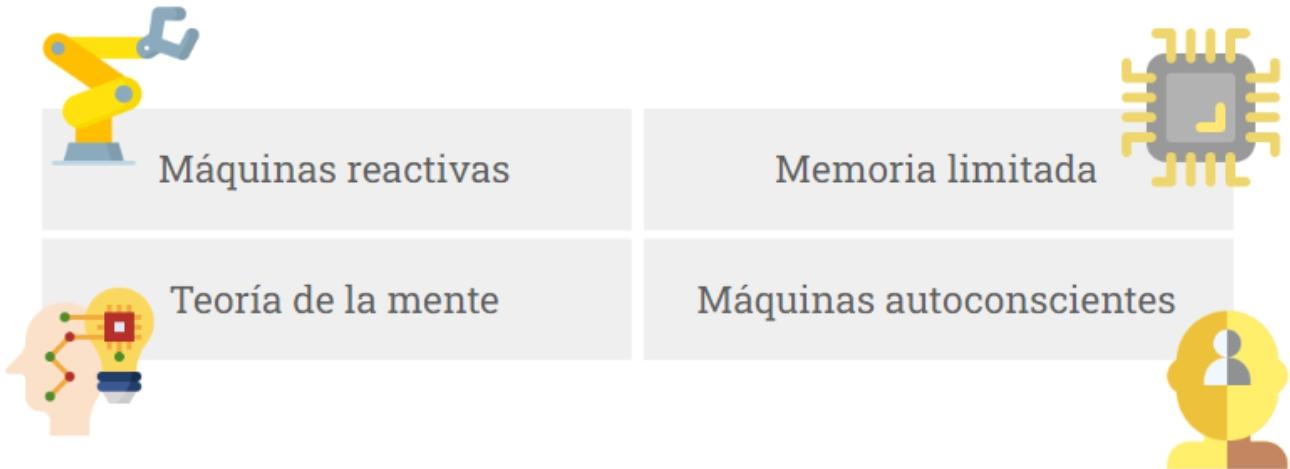
— Tiene que haber algo un poco más actual para poder clasificar la Inteligencia Artificial — sentenció Eva después de un buen rato enfrascada en Internet.

— A lo mejor tenemos que buscar entre los científicos y profesores que están desarrollando y enseñando ahora mismo Inteligencia Artificial — dijo Miguel después de pensar un rato en silencio — . Si están haciendo Inteligencia Artificial actualmente, o enseñándola, no creo que se limiten a la clasificación simple de Débil o Fuerte, o a la ya prácticamente obsoleta de Convecional o Computacional.

— ¡Eso es! — Exclamó Eva — Busquemos cómo están explicando y clasificando la Inteligencia Artificial los profesores que en este momento imparten clases de Inteligencia Artificial en universidades de prestigio. ¡Seguro que encontramos algo "super-actual" y más práctico que lo que hemos visto hasta ahora!

En noviembre de 2016, Arend Hintze, profesor de la Universidad de Michigan e investigador en el campo de la Inteligencia Artificial, escribió un artículo titulado: “[Understanding the four types of AI, from reactive robots to self-aware beings](#)” (Comprendiendo los cuatro tipos de IA, desde los robots reactivos a los seres auto-conscientes), en el que sintetizaba toda la evolución de los últimos desarrollos y avances en materia de Inteligencia Artificial para aportar una clasificación más realista y concreta para los tipos de entidades que existen o que se aspira a crear.

### Tipos de Inteligencia Artificial según Hintze



Carmen Bartolome ([CC BY-SA](#))

# Para saber más

Stuart J. Russell y Peter Norvig, investigadores informáticos, publicaron en 1995 su libro “[Artificial Intelligence: A Modern Approach](#)”, que se ha convertido en el libro de texto fundamental en cientos de universidades a nivel mundial (ya lleva varias ediciones publicadas). Plantean cuatro categorías básicas partiendo de un enfoque de génesis del acto inteligente, del origen y proceso por el cual se llega al comportamiento inteligente.

- ✓ **Sistemas Cognitivos:** Piensan como humanos, intentan emular el proceso humano → Proceso de toma de decisiones, resolución de problemas, y el propio paradigma del aprendizaje.
- ✓ **Test de Turing:** Actúan como humanos, intentan emular el comportamiento humano (sin pasar por el pensamiento o razonamiento que conduce a dicho comportamiento) → A nivel práctico se aplica en la robótica y sistemas de actuadores en el mundo físico.
- ✓ **Leyes del pensamiento:** Piensan con razonamientos. Cumplimiento exacto de las leyes del razonamiento lógico, teniendo en cuenta todos los factores que afectan a la cuestión. No hemos llegado a esto aún. Sería el caso de los sistemas expertos. Solo son posibles aproximaciones para campos de investigación muy especializados y acotados.
- ✓ **Agentes inteligentes:** Actúan racionalmente (sin pasar por el proceso de razonamiento lógico).

Los dos últimos requieren una capacidad de cómputo muy importante, a veces, aún, inaccesible.

## Tipos de IA según Russel y Norvig



Fran Bartolomé - Elaboración propia ([CC BY-SA](#))

## 3.1.- Máquinas reactivas.

Los tipos más básicos de sistemas de IA son puramente reactivos. No tienen la capacidad de formar recuerdos. Tampoco pueden utilizar experiencias pasadas en las que basar las decisiones actuales.



Carmen Bartolome (elaboración propia) ([CC BY-SA](#))

**Deep Blue** (descrita en el tema anterior) fue una supercomputadora creada por **IBM**. Fue capaz de vencer al ajedrez al gran maestro internacional Garry Kasparov. Ocurrió a fines de la década de 1990 y es el ejemplo perfecto de este tipo de máquina.

Puede **identificar las piezas** en un tablero de ajedrez y **saber cómo se mueve cada una**. Puede **realizar predicciones** sobre los mejores movimientos y elegir el mejor de todas las posibilidades.

Pero **no tiene ningún concepto del pasado**. Tampoco posee recuerdos de lo que ha sucedido antes. Aparte de una regla de ajedrez, *Deep Blue* ignora todo antes del momento presente. Todo lo que hace es enfocar las piezas del tablero en tiempo real y elegir entre los siguientes movimientos posibles.

En el caso de que se trate de una Inteligencia Artificial reactiva aplicada a una máquina capaz de "conversar" es importante que el usuario sepa que está tratando con una máquina, pues de lo contrario se suelen crear falsas expectativas sobre lo que puede esperar de dicha conversación.

### Autoevaluación

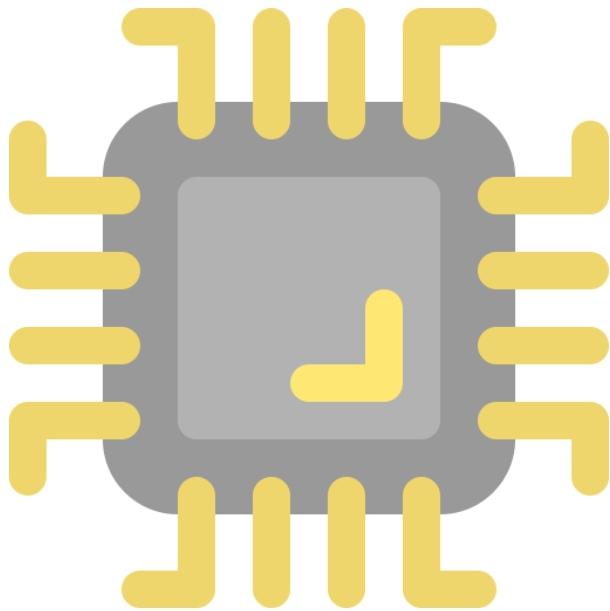
Completa la frase rellenando los huecos con las palabras correctas.

Los tipos más básicos de sistemas de IA son puramente  . No tienen la capacidad de formar  . Tampoco pueden utilizar experiencias pasadas en las que basar las decisiones  .

## 3.2.- Memoria limitada.

---

El segundo tipo de Inteligencia Artificial que contempla la clasificación de Hintze se caracteriza por que sí maneja máquinas que **pueden mirar hacia el pasado**. Los vehículos autónomos ya hacen algo parecido. Por ejemplo, observan la velocidad y dirección de otros autos, y en base a la memorización de dicha información toman decisiones en el futuro inmediato.



Carmen Bartolome (Elaboración Propia) ([CC BY-SA](#))

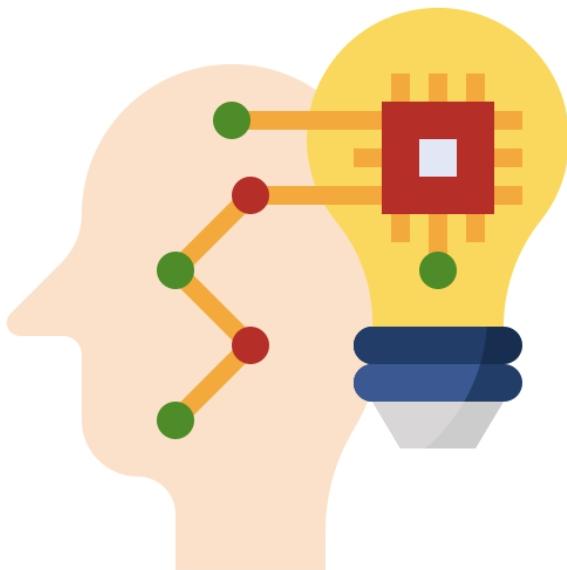
Digamos que estas observaciones se agregan a las representaciones preprogramadas para la memoria de estos coches. Se incluyen marcas de carril, semáforos y otros elementos importantes, como curvas en la carretera.

También se añaden experiencias como cuando el automóvil decide en qué momento cambiar de carril para evitar interrumpir a otro conductor o ser embestido por un automóvil cercano.

Pero estas simples piezas de información sobre el pasado son solo transitorias. No se guardan como parte de la biblioteca de experiencias del automóvil. En estos tipos de inteligencia artificial, la máquina no puede compilar la experiencia durante años, como lo hace un humano.

### 3.3.- Teoría de la mente.

---



Carmen Bartolome (Elaboración Propia) ([CC BY-SA](#))

¿Podemos construir sistemas de Inteligencia Artificial que construyan representaciones completas, recordar sus experiencias y aprender cómo manejar situaciones nuevas?

Llegamos a un punto en el que nos acercamos más a los tipos de Inteligencia Artificial que deseamos en un futuro. Las máquinas de esta clase son más avanzadas. No solo **forman representaciones sobre el mundo**, también sobre **otros agentes o entidades**.

En psicología, esto se denomina '**teoría de la mente**'. Implica la comprensión de que las personas, las criaturas y los objetos en el mundo pueden tener **pensamientos y emociones que afectan a su propio comportamiento**. Esto es crucial para la forma en que los humanos formamos sociedades, porque nos permite la interacción social.

Si las máquinas van a andar entre nosotros, deberán tener una comprensión sobre cómo pensamos y cómo sentimos. Además deberán llegar a saber qué esperamos y cómo queremos que nos traten. Tendrán que ajustar su comportamiento en consecuencia.

Como habrás podido intuir, este tipo de máquinas aún no existen. Igual que la IA Fuerte es aún algo que se sabe cómo funcionará pero que no hemos llegado a desarrollar todavía.

## 3.4.- Autoconciencia.



Carmen Bartolome  
(Elaboración Propia) ([CC BY-SA](#))

El paso final del desarrollo de la IA es construir sistemas que puedan **formar representaciones sobre sí mismos**. En última instancia, los investigadores de la Inteligencia Artificial tendrán que **comprender no solo la conciencia**, sino también construir máquinas que la tengan.

Los seres conscientes son conscientes de sí mismos, conocen sus estados internos y pueden predecir los sentimientos de los demás. Es probable que **estemos lejos de crear máquinas que sean conscientes de sí mismas**. Sin embargo, los esfuerzos se enfocan hacia la comprensión de la memoria, el aprendizaje y la capacidad de basar las decisiones en experiencias pasadas.

Este es un paso importante para entender la inteligencia humana por sí misma. Es crucial para diseñar o desarrollar máquinas que sean más excepcionales para clasificar lo que ven frente a ellas.

Los cuatro tipos de inteligencia artificial dan una idea sobre las intenciones que el hombre tiene acerca del futuro de la máquina. Puede que estemos muy lejos de la Inteligencia Artificial autoconsciente. No obstante, está claro que eso es lo que se persigue en última instancia.

### Autoevaluación

Completa la siguiente frase rellenando los huecos con las palabras correctas.

El paso  del desarrollo de la Inteligencia Artificial es construir sistemas que puedan formar representaciones sobre  mismos. En última instancia, los investigadores de la IA tendrán que comprender no solo la , sino también construir máquinas que la .

# Condiciones y términos de uso de los materiales

Materiales desarrollados inicialmente por el Ministerio de Educación, Cultura y Deporte y actualizados por el profesorado de la Junta de Andalucía bajo licencia Creative Commons BY-NC-SA.



MINISTERIO  
DE EDUCACIÓN  
Y FORMACIÓN PROFESIONAL



Antes de cualquier uso leer detenidamente el siguiente [Aviso legal](#)

## Historial de actualizaciones

Versión: 01.00.00

Fecha de actualización: 05/07/23

Versión inicial de los materiales.