



AGENTES

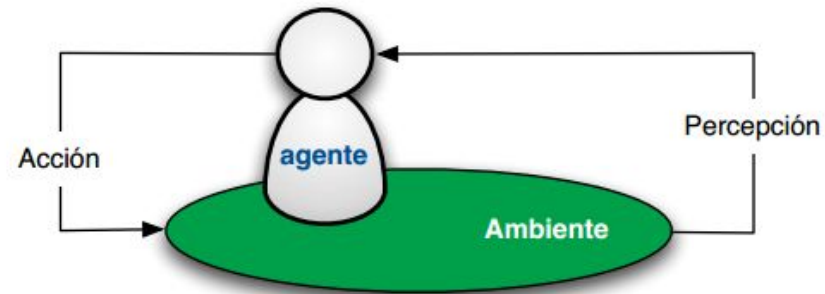
Introduction



Bloque 1: Principios de los Agentes Autónomos

1. ¿Qué es un Agente Autónomo?

- Un **Agente** es cualquier entidad que percibe su entorno a través de **sensores** y actúa sobre ese entorno a través de **actuadores**.
- **Agente Autónomo:** Es un agente inteligente que puede operar **independientemente** para lograr objetivos complejos sin la intervención humana constante.
- **Objetivo:** Maximizar una medida de rendimiento (racionalidad) dentro de su entorno.



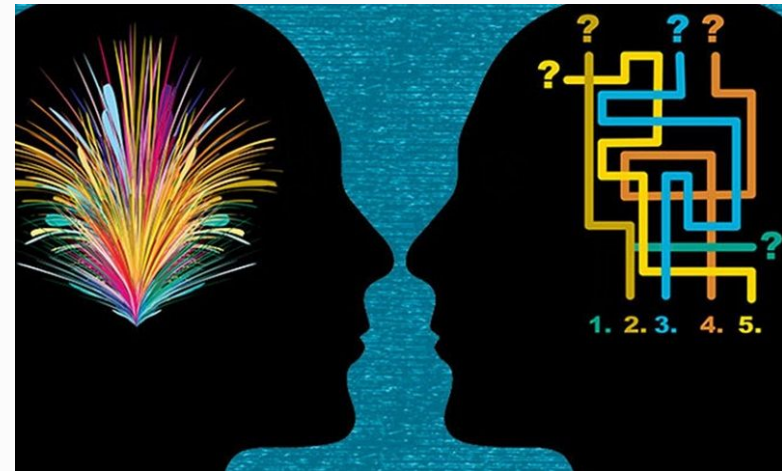
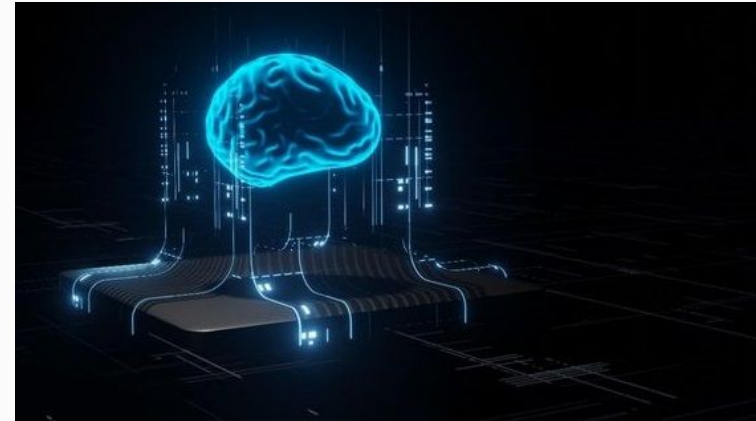
2. Principio 1: Autonomía

- **Definición:** La capacidad de operar sin control directo o constante de un humano.
- **Implicación:** El agente tiene la capacidad de **tomar sus propias decisiones** sobre qué hacer a continuación, basándose en su estado interno y percepción.
- **Evolución:** Va más allá de la simple automatización; implica una toma de decisiones dinámica y adaptativa.
- **Ejemplo:** Un dron de entrega que ajusta su ruta y velocidad en tiempo real para evitar obstáculos imprevistos.



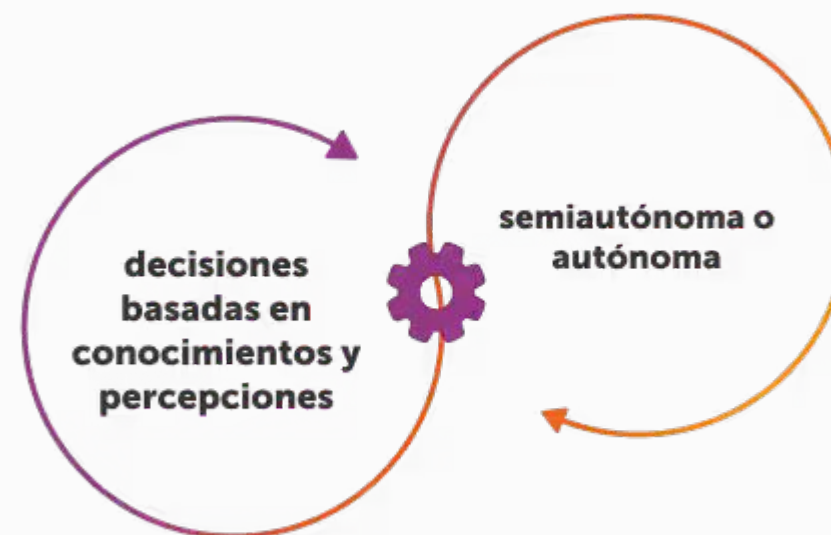
3. Principio 2: Racionalidad

- **El Agente Racional:** Siempre elige la acción que se espera que maximice su **medida de rendimiento** (función de utilidad), dadas las percepciones que ha visto hasta ese momento.
- **No es solo hacer lo correcto:** Es hacer lo **más óptimo** según la información disponible, incluso si esa información es incompleta o incierta.
- **Base de Decisión:** Percepciones + Conocimiento Interno --> Acción Racional.



4. Principio 3: Reactividad y Proactividad

- **Reactividad:** Responde rápidamente a los cambios inmediatos del entorno. (Ej. Frenar ante un obstáculo).
- **Proactividad:** Inicia acciones dirigidas a metas a largo plazo, en lugar de solo reaccionar. (Ej. Planificar una ruta completa y buscar un destino).
- **Agente Ideal:** Equilibra ambas; debe ser capaz de reaccionar a peligros inmediatos mientras avanza hacia sus metas generales.



5. Otros Principios Clave

- **Aprendizaje (Learning):** Capacidad de mejorar su rendimiento y adaptar su comportamiento a lo largo del tiempo y la experiencia (e.g., ajustando su modelo del mundo o sus prioridades).
- **Persistencia:** El agente funciona continuamente y mantiene un estado interno o memoria a lo largo del tiempo.
- **Sociabilidad (Sistemas Multiagente):** Capacidad de interactuar y colaborar con otros agentes o humanos para lograr objetivos compartidos.

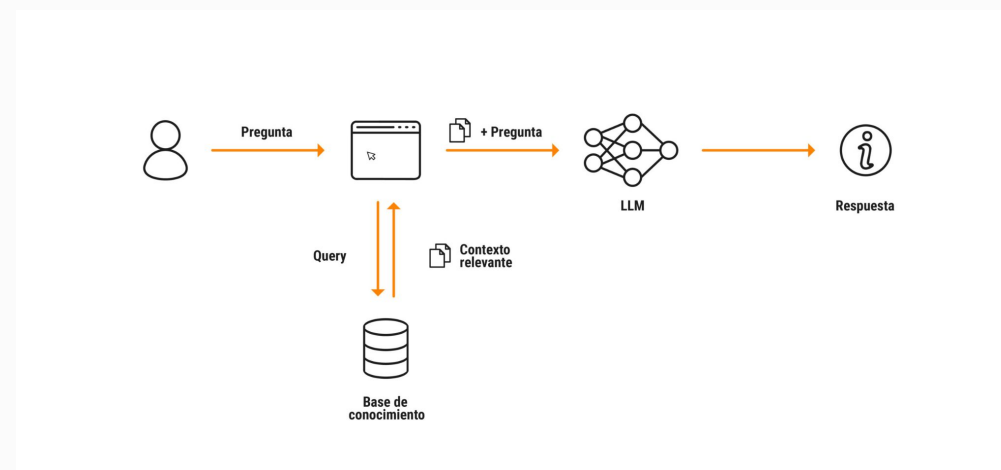
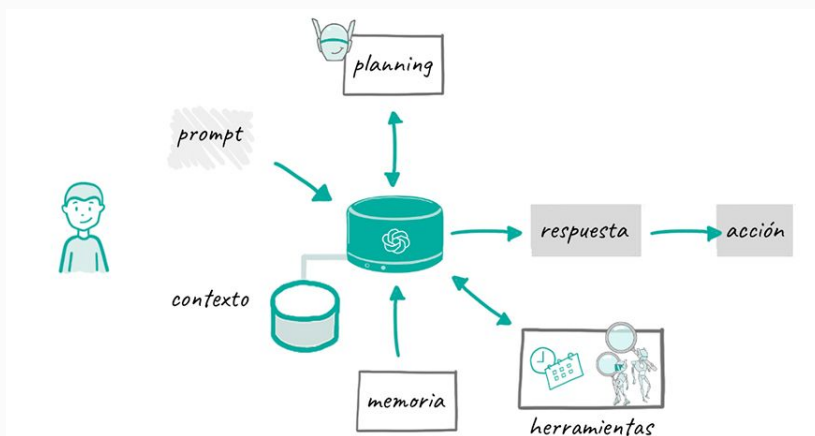




Bloque 2: Arquitecturas de Agentes

1. Arquitecturas de Agentes

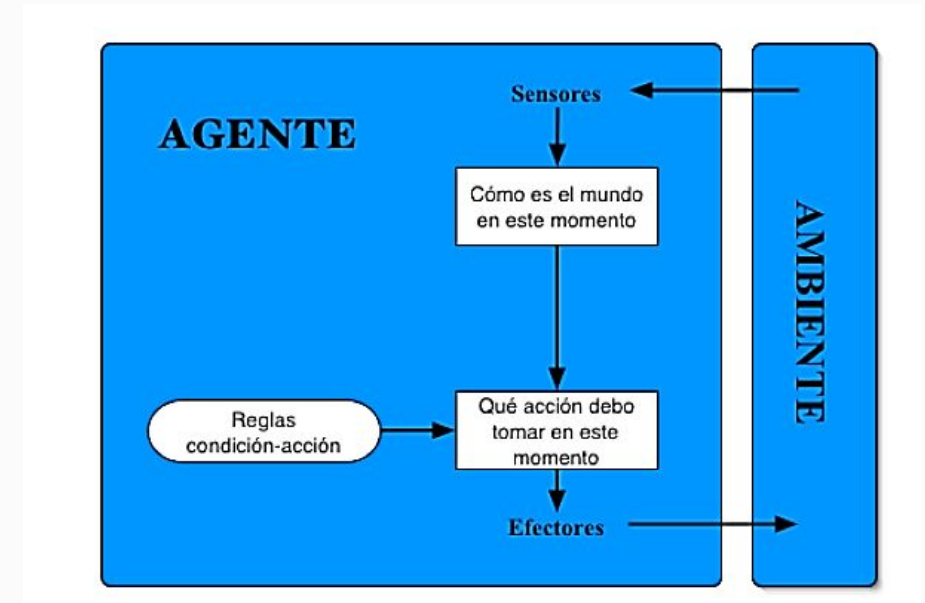
- La **Arquitectura** es el diseño interno del agente; cómo se organiza el programa para mapear las percepciones a las acciones.
- Las arquitecturas varían según la complejidad del **razonamiento** y el uso de un **Modelo Interno del Mundo**.
- **Tipos Principales:** Reactivos, Deliberativos (Basados en Objetivos) y Basados en Utilidad.



2. Arquitecturas Reactivas (Simples)

- **Mecanismo:** Estímulo ->Acción.
- **Características:**
 - No tienen **memoria explícita** ni modelo interno del mundo.
 - Toman decisiones basándose únicamente en la **percepción actual**.
- Son **rápidos** y eficientes, pero limitados a entornos simples o a decisiones urgentes.

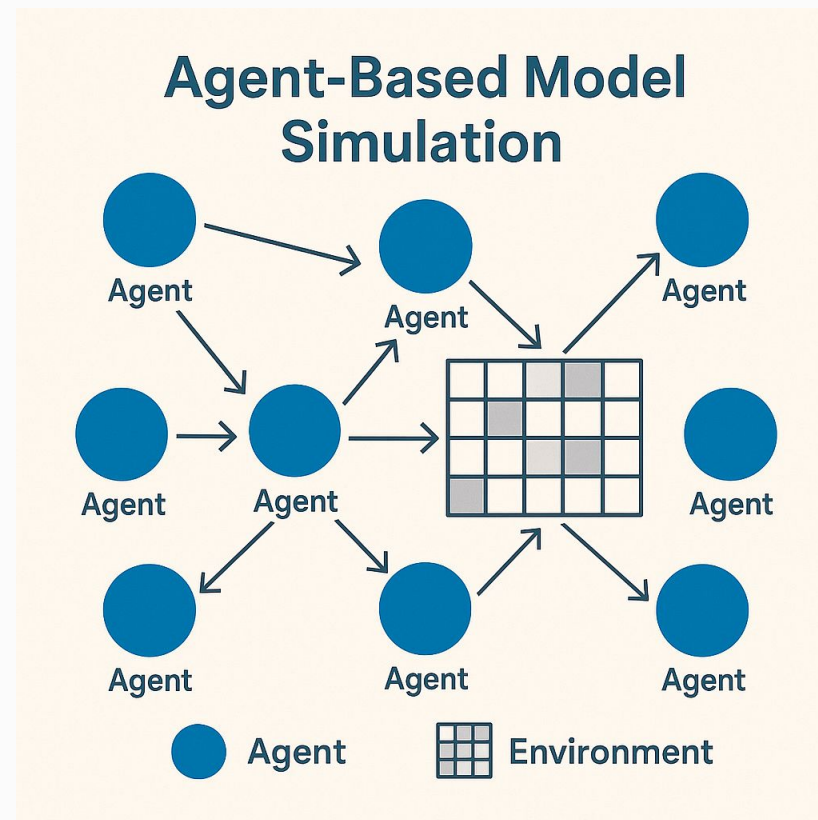
Ejemplo: Un filtro de correo no deseado que bloquea correos solo por palabras clave predefinidas en el asunto.



3. Arquitecturas Reactivas Basadas en Modelo

- **Mecanismo:** Percepción + Modelo Interno -> Acción.
- **Modelo Interno:** Es una representación de cómo funciona el mundo y cómo las acciones del agente lo afectan.
- **Ventaja:** Pueden basar la acción en la parte **no observable** del estado (lo que recuerdan o infieren).

Ejemplo: Un sistema de seguridad que recuerda que una puerta fue abierta hace 5 minutos, aunque ahora parezca cerrada.



4. 🌐 Arquitecturas Deliberativas (Basadas en Objetivos)

- **Mecanismo:** Percepción -> Planificación hacia el Objetivo -> Acción.
- **Características:**
 - Tienen **metas u objetivos** definidos.
 - Utilizan un modelo del mundo para **planificar** secuencias de acciones (planes) para alcanzar la meta.
 - La deliberación (pensar) es más costosa en tiempo.
- **Uso:** Tareas complejas que requieren previsión (e.g., planificación logística, sistemas de navegación GPS).

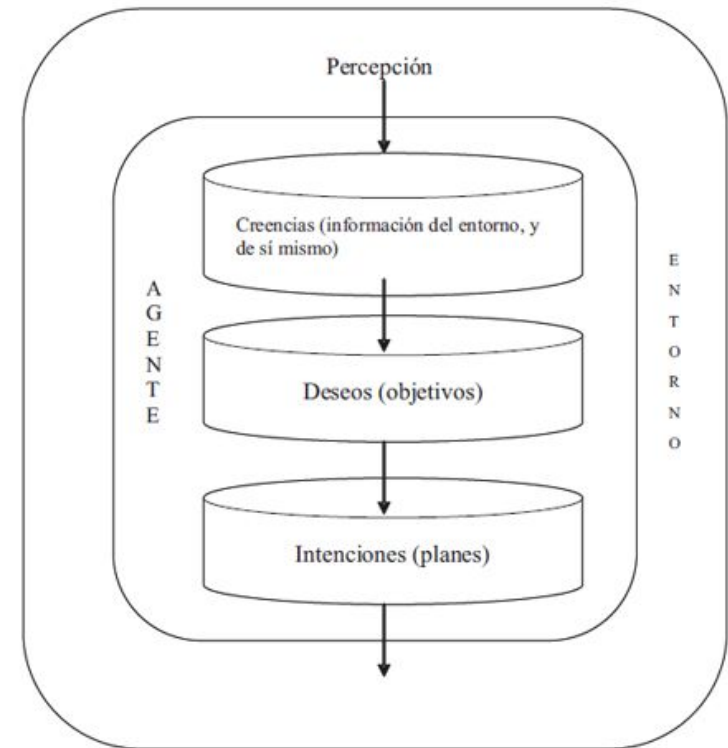


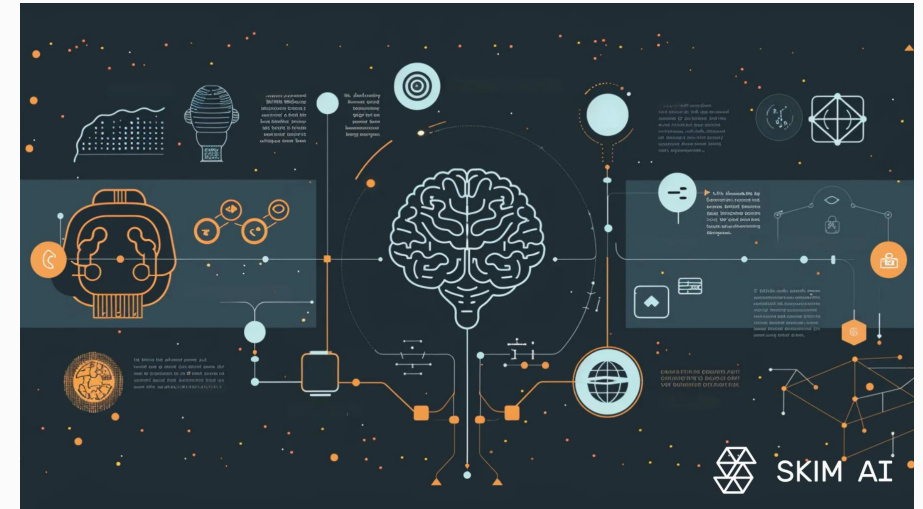
Figura 2. Agentes deliberativos

5. Deliberativas: El Proceso de Planificación

- 1. Definición del Objetivo:** El estado deseado (e.g., llegar a Madrid).
- 2. Búsqueda de Solución:** Explorar el espacio de estados y acciones posibles (e.g., diferentes rutas).
- 3. Generación del Plan:** Seleccionar la mejor secuencia de acciones encontradas (e.g., la ruta más corta o rápida).
- 4. Ejecución del Plan:** Llevar a cabo las acciones paso a paso, volviendo a planificar si hay desviaciones.

6. Arquitecturas Basadas en Utilidad

- **Mecanismo:** Percepción->Maximización de Utilidad
->Acción.
- **Función de Utilidad:** Mide el grado de "felicidad" o **preferencia** del agente por un estado determinado o resultado.
- **Toma de Decisiones:** Eligen la acción que maximiza la **Utilidad Esperada**, considerando la probabilidad de los resultados.
- **Nota:** Son los agentes más racionales, ideales para entornos con incertidumbre.



7. Utilidad: Manejo de la Complejidad

- **1. Incertidumbre:** Pueden evaluar la probabilidad de que una acción tenga éxito. (Ej. Si el 90% de probabilidad de éxito me da una utilidad de 10, y el 10% de fracaso me da una utilidad de -100).
- **2. Objetivos Conflictivos:** Permiten ponderar la importancia de diferentes metas. (Ej. Asignar 80% de utilidad a la **seguridad** y 20% a la **velocidad**).
- **Fórmula Conceptual:** Acción Óptima = $\max_A \sum \text{Probabilidad}(\text{Resultado}) \times \text{Utilidad}(\text{Resultado})$

8. Arquitecturas Híbridas (Combinación)

- **El Reto:** Los agentes reactivos son rápidos pero miopes; los deliberativos son inteligentes pero lentos.
- **Solución Híbrida:** Combina un componente reactivo (para reacciones urgentes) con un componente deliberativo (para planificación a largo plazo).

Ejemplo: Un coche autónomo. Utiliza el módulo reactivo para frenar de golpe ante una emergencia (rápido) y el módulo deliberativo para planificar la ruta y la estrategia de adelantamiento (lento).



Bloque 3: Ciclo de Vida de un Agente

1. El Ciclo de Vida del Agente (Bucle P-P-A)

- El funcionamiento dinámico de cualquier agente autónomo se basa en un ciclo iterativo y continuo.
- Este ciclo se denomina: **Percepción, Planificación** (o Razonamiento) y **Acción**.
- Cada ciclo es una interacción con el entorno que busca acercar al agente a sus objetivos.

Nota: Muestra un gráfico de ciclo con flechas.



2. Fase 1: Percepción (Sense)

- **Entrada de Datos:** El agente utiliza sus **sensores** (cámaras, micrófonos, APIs, bases de datos) para recoger información del entorno.
- **Procesamiento:** Los datos brutos son procesados e interpretados para formar una **percepción** coherente y significativa del estado actual del mundo.
- **Resultado:** Actualización del **Estado Interno** del agente (su "memoria" o modelo del mundo).

3. Fase 2: Planificación (Think/Reason)

- **Propósito:** Determinar la mejor acción a tomar en el momento actual.
- **Proceso Clave:**
 - **Agentes Reactivos:** Mapeo directo (si la percepción es X, la acción es Y).
 - **Agentes Deliberativos:** Implica un **razonamiento** complejo: evaluar posibles futuros, generar o seleccionar un plan que maximice la probabilidad de éxito de la meta.
- **Salida:** La **decisión** (la próxima acción o secuencia de acciones a ejecutar).

4. Fase 3: Acción (Act)

- **Ejecución:** La decisión tomada en la fase de Planificación se traduce en comandos para los actuadores.
- **Actuadores:** Pueden ser físicos (motores de un robot, ruedas de un vehículo) o de software (enviar un correo electrónico, modificar un archivo, ejecutar un contrato inteligente).
- **Efecto:** La acción cambia el estado del entorno. Esto completa el ciclo.
- **Nota:** Una acción exitosa acerca al agente a su meta.

5. El Bucle P-P-A es Iterativo y Dinámico

- La acción produce un cambio en el entorno, que es inmediatamente captado en la **siguiente Percepción**.
- Este bucle continuo permite al agente **adaptarse** y **corregir** errores o desviaciones del plan en tiempo real.
- **Aprendizaje en el Ciclo:** La retroalimentación de las acciones anteriores (éxito o fracaso) se utiliza para mejorar los procesos de Planificación y el Modelo Interno en ciclos futuros.

6. Conclusión y futuro

- Los agentes autónomos representan un avance crucial en la IA al permitir que los sistemas actúen de forma independiente y racional.
- El diseño de su inteligencia está determinado por la **arquitectura** elegida, desde la simple reactividad hasta la compleja deliberación por utilidad.
- El ciclo **Percepción-Planificación-Acción** es el motor operativo que impulsa su interacción dinámica con el mundo real.
- **El Futuro:** La tendencia es hacia sistemas híbridos y multiagente, capaces de manejar objetivos más abstractos y colaborar a gran escala.

