

Cursos Extraordinarios

Verano 2024

**“Inteligencia Artificial y Medios:
estrategias innovadoras frente a la
desinformación”**



Las tecnologías de Inteligencia Artificial en el ámbito periodístico

TODOS LO QUE NECESITAS ES UN MODELO DE LENGUAJE



Inteligencia Artificial

Capacidad de las **máquinas** para realizar **tareas** que normalmente requieren **destrezas humanas**.

Esto incluye:

Aprendizaje,

percepción,

razonamiento,

toma de decisiones,...

reconocimiento de voz,

síntesis de voz,

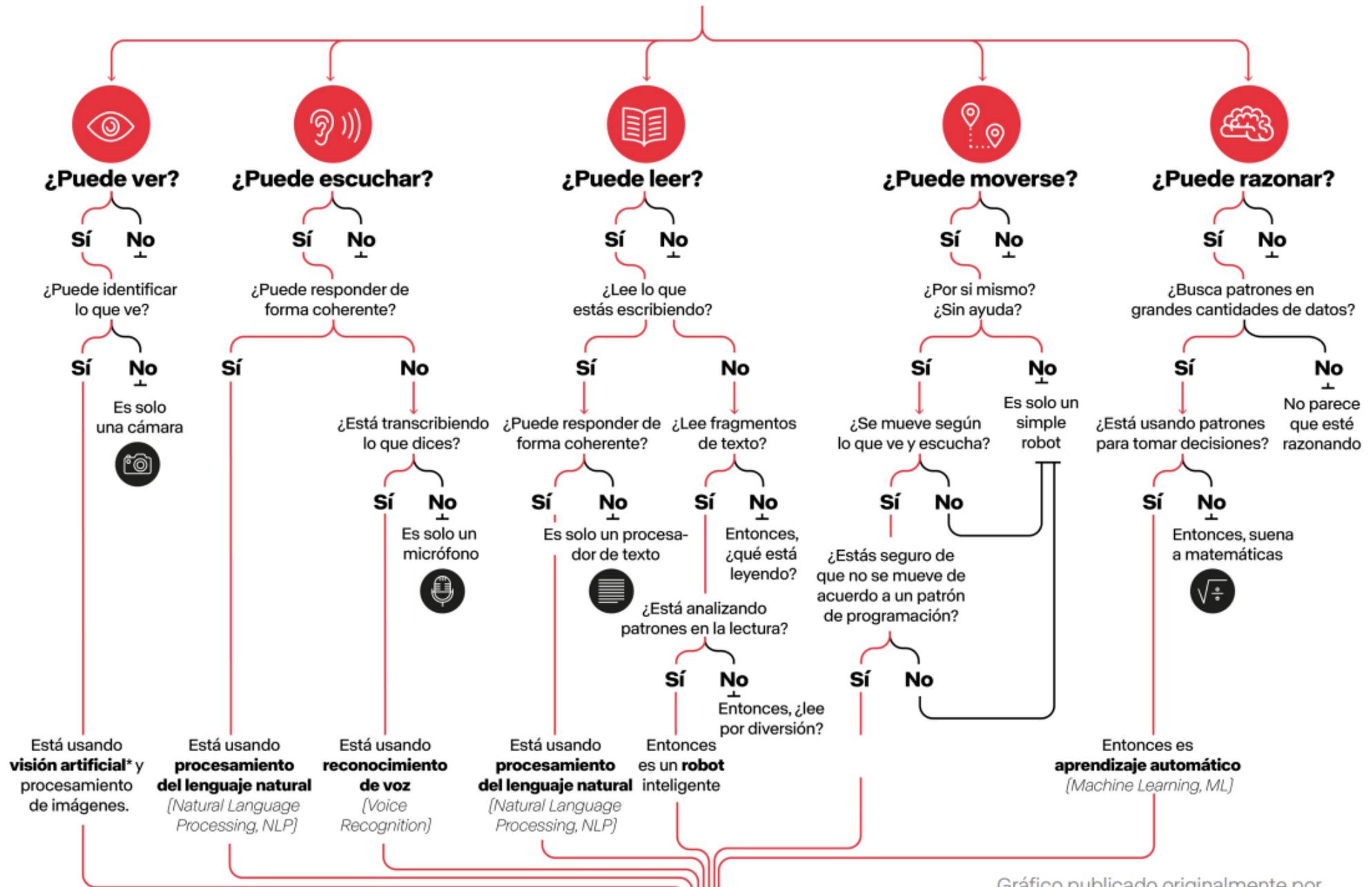
visión,

comprensión del lenguaje natural,





¿Está utilizando IA?





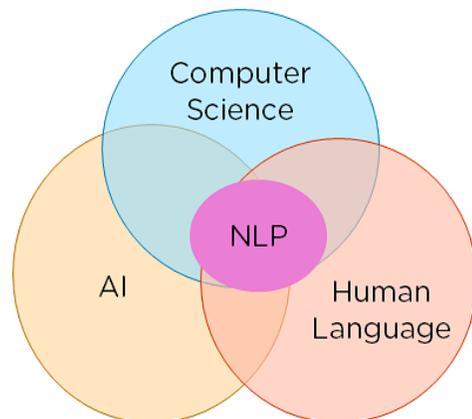
Inteligencia Artificial y Procesamiento del Lenguaje Natural

¿Qué es el procesamiento del lenguaje natural?

- ✓ Término genérico que abarca todo aquello que permite a las máquinas *procesar* el *lenguaje humano* tanto en forma *escrita*, *verbal*, o *visual*.

¿Porqué es importante el procesamiento del lenguaje natural?

- ✓ Componente/Capacidad fundamental de los sistemas de IA.



Capacidades de un sistema de IA

- Percepción
- Aprendizaje
- Representación del conocimiento
- Razonamiento



Inteligencia Artificial y Procesamiento del Lenguaje Natural



Inteligencia Artificial y aplicaciones

- **Generación de Contenidos:**
 - Narrativas Automáticas: Plataformas que generan noticias a partir de datos.
 - Asistentes de Redacción: Herramientas que ayudan a redactar y editar artículos.
- **Curación de Contenidos:**
 - Agregadores de Noticias: Algoritmos que seleccionan y presentan noticias relevantes para audiencias específicas.
 - Personalización: IA que adapta el contenido según las preferencias del usuario.
- **Moderación de Contenidos:**
 - Clasificación y etiquetado: Sistemas que clasifican y etiquetan automáticamente el contenido según su adecuación y relevancia.
 - Detección de comportamiento tóxico: Algoritmos entrenados para identificar lenguaje abusivo, amenazas, imágenes, spam, y otros tipos de contenido inapropiado.

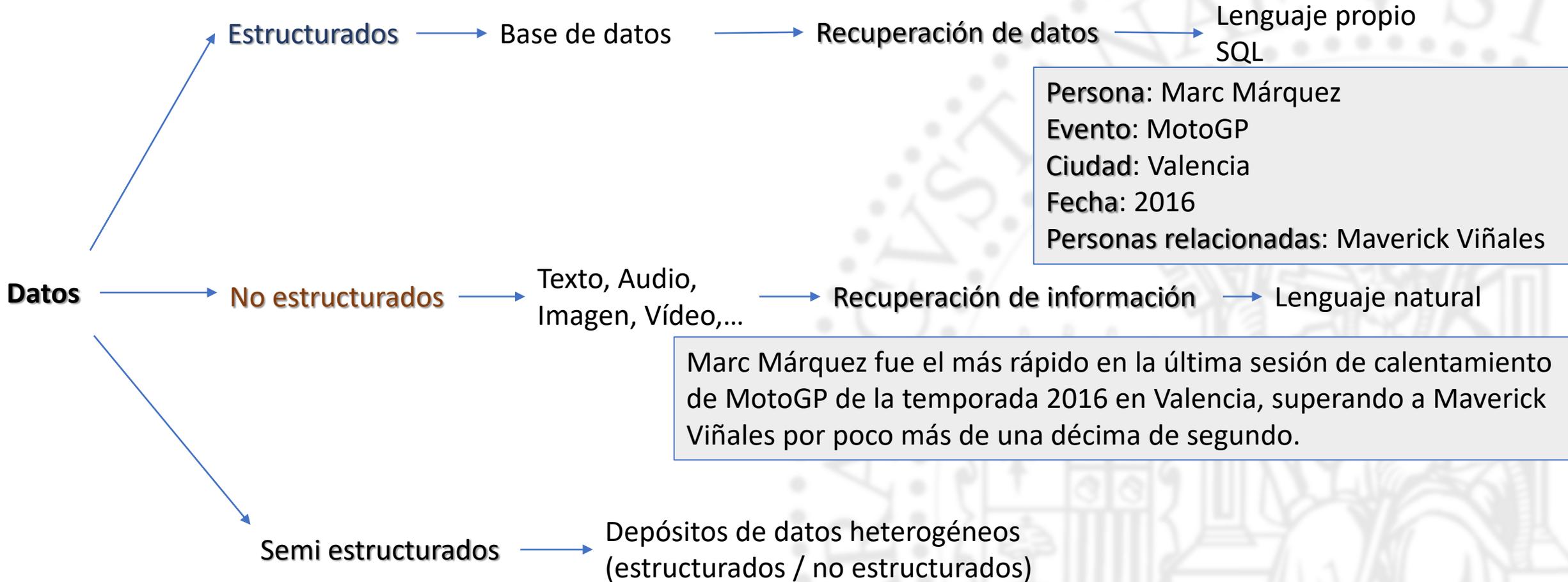


Inteligencia Artificial y aplicaciones

- **Verificación de Datos:**
 - **Fact-checking Automatizado:** Herramientas que verifican la veracidad de los hechos.
 - **Detección de Noticias Falsas:** Algoritmos que identifican y filtran noticias falsas y desinformación.
- **Análisis de Audiencia:**
 - **Análisis de Sentimientos:** Herramientas que analizan cómo reacciona la audiencia a diferentes temas.
 - **Segmentación de Audiencias:** IA que segmenta a los lectores en grupos específicos para personalizar el contenido.
-

Inteligencia Artificial e información

El periodista trabaja con datos y los datos ...



Inteligencia Artificial e información

¿Cómo representamos la información?

Tesauros:

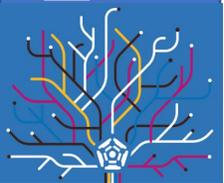
- un tesoro es un instrumento de control terminológico utilizado para traducir a un lenguaje más estricto el idioma natural empleado en los documentos y por los indizadores. (UNESCO)
- vocabulario controlado y dinámico de términos relacionados semántica

Ontologías:

- las ontologías son teorías que especifican un vocabulario relativo a un cierto dominio. Este vocabulario define entidades, clases, propiedades, predicados y funciones y, las relaciones entre estos componentes.

La ontologías y los tesauros han adquirido una importancia capital debido a que “existe la necesidad de **describir la semántica de un dominio** de forma que el humano lo entienda y que pueda ser procesado por el ordenador”

¿Embeddings y modelos de lenguaje?



Modelo de lenguaje

- Representación matemática de la utilización de la **lengua**
- **Comprender**, generar y predecir el texto en función del **contexto**

Hasta 2013

- Bolsa de palabras (Bag-of-Words): contar frecuencia de aparición de las palabras en un texto, no se tiene en cuenta el orden o la estructura gramatical. Clasificación de textos y análisis de sentimiento.
- Modelos estadísticos: modelar la probabilidad de ocurrencia de palabras o secuencias de palabras en un texto. Uso en reconocimiento automático del habla, traducción automática y generación de texto.





Inteligencia Artificial e información

¿Qué ocurrió en 2013?

“Efficient Estimation of Word Representations in Vector Space”, Thomas Mikolov et al.

<https://doi.org/10.48550/arXiv.1301.3781>

- Aparición del Word2vec: representación vectorial semántica (embedding) de palabras usando redes neuronales

¿Embedding? (incrustación)

- Representación de una palabra/frase/documento en un espacio vectorial de alta dimensión.
- Un embedding transforma datos categóricos en una forma numérica que conserva la información y las relaciones contextuales de los datos originales, permitiendo una manipulación y análisis más eficientes. Los embeddings capturan las relaciones semánticas y contextuales entre las palabras.



Inteligencia Artificial y modelos de lenguaje

Pasamos de

Frases generadas de forma automática con el modelo de trigramas

sólo quiero mencionar es que el horario de mañana comienzan las fiestas del proyecto que hay que cumplimos con el resto de la cuenta atrás para Málaga y mañana el presidente del consejo de ministros que ha evolucionado puedo permitir el tratado de Amsterdam

señor Fischler presentará ahora es una de ellas estaba ingresado por una posible sanción de nueve grados más esperados como la afirmación de que dice debe cesar económica que el tribunal así lo haya clarificado todavía no han logrado calar una cuenta en matar uno de los reclusos etarras pero por encima de todo lo que sucede hoy

qué duda cabe juzgado el juez los dos islamistas radicales han pedido más ayuda según el fiscal le acusó de haber abonado los tres últimos meses su cargo



Inteligencia Artificial y modelos de lenguaje

A

Generación de texto red neuronal recurrente con vocabulario de 40 tokens

Las temperaturas bajarán en el norte y en el sur, lloviznas en el este y en Galicia

y en Asturias, con algunas nubes más por la tarde, mientras que en las Islas Baleares se mantienen los cielos básicamente despejados. Aquí tenemos las nubes que se extienden por muy poco espacios de las próximas horas. De hecho, esta es la situación de cara a la jornada del domingo, continuaremos hablando de una situación marcada por la estabilidad. Por tanto también con intervalos de nubes altas y medias en la costa mediterránea. En general tiempo estable en el centro y en el oeste del país. Esto se va a notar más porque el viento es de componente norte que va a traer un tiempo muy parecido al de hoy. En el resto del país con cielos despejados, pero también con nieblas en el norte de las islas de mayor relieve.

Inteligencia Artificial y modelos de lenguaje

Pero seguía sin comprender lo que genera...

y en 2017

“Attention Is All You Need”, Ashish Vaswani et al. <https://doi.org/10.48550/arXiv.1706.03762>

Nueva arquitectura de redes neuronales: Transformer

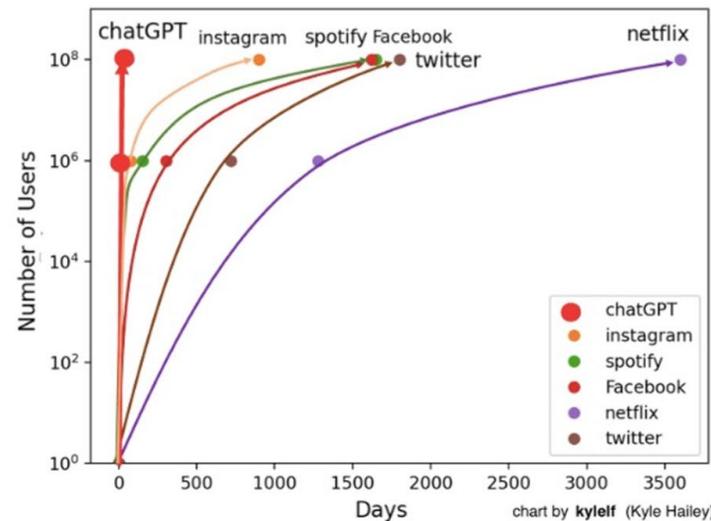
... y aparecen los grandes modelos de lenguaje (GPT, Llama, Phi, Gemini, ...)

... y aparecen los grandes modelos de lenguaje multimodales (Texto + imagen + audio + ...)



... y se abre un mundo de oportunidades no previstos ni esperados.

Noviembre 2022 ...



Evolución del número de usuarios en función de los días transcurridos desde su lanzamiento para diversos productos tecnológicos

Inteligencia Artificial y modelos de lenguaje

... y aparecen los problemas

Las alucinaciones en los LLM se refieren a la generación de contenido que es irrelevante, inventado o inconsistente con los datos de entrada.

Taxonomía de las alucinaciones

- **Alucinaciones de hechos**
Ocurre cuando el modelo genera contenido incorrecto sobre hechos.
- **Alucinaciones de fidelidad**
Ocurren cuando el modelo produce contenido que no es fiel o es inconsistente con el contenido proporcionado.

Pregunta a GPT3.5

¿qué relación hay entre el pueblo de Calasanz y el conde de la Viña?

Respuesta:

La relación entre el pueblo de Calasanz y el conde de la Viña se remonta a la época medieval. El conde de la Viña era un noble que poseía tierras y propiedades en la zona de Calasanz. Como señor feudal, tenía autoridad sobre el pueblo y sus habitantes.

El conde de la Viña ejercía su poder y controlaba los recursos del pueblo, como la tierra y los impuestos....



Inteligencia Artificial y modelos de lenguaje

Estrategias de Mitigación de Alucinaciones

1. **Mejora de la Calidad de los Datos:** Asegurar la precisión y la completitud de los datos de entrenamiento para minimizar la introducción de desinformación y sesgos.

Soluciones en nuestras manos:

Prompt Engineering

“Retrieval Augmentation Generation” - RAG

2. **Mejoras en el Entrenamiento:** Desarrollar mejores arquitecturas y estrategias de entrenamiento, como el modelado de contexto bidireccional y técnicas para mitigar el sesgo de exposición.

Solución en nuestras manos:

Fine-tuning

3. **Técnicas Avanzadas de Decodificación:** Emplear métodos de decodificación más sofisticados que equilibren la aleatoriedad y la precisión para reducir la aparición de alucinaciones.

Solución en nuestras manos: **Temperatura, top_p, top_k**



Inteligencia Artificial y modelos de lenguaje

Mejora de la calidad de los textos permitiendo ir a buscar a fuentes confiables

Enter query:

busca la relación que hay entre el pueblo de Calasanz y el conde de la Viña

> Entering new AgentExecutor chain...

Invoking: `Wikipedia` with `Pueblo de Calasanz`

Page: Aragon

Summary: Aragon (ARR-ə-gən, US also -gon, -gohn; Spanish and Aragonese: Aragón [araˈɣon] ; Catalan: Aragó [əɾəˈɣo]) is an autonomous community in Spain, coextensive with the medieval Kingdom of Aragon. In northeastern Spain, the Aragonese autonomous community comprises three provinces (from north to south): Huesca, Zaragoza, and Teruel. Its capital is Zaragoza. The current Statute of Autonomy declares Aragon a historic nationality of Spain.

Covering an area of 47720 km² (18420 sq mi), the region's terrain ranges diversely from permanent glaciers to verdant valleys, rich pasture lands and orchards, through to the arid steppes of the central lowlands. Aragon is home to many rivers—most notably, the river Ebro, Spain's largest river in volume, which runs west–east across the entire region through the province of Zaragoza. It is also home to the highest mountains of the Pyrenees.

As of January 2021, the population of Aragon was 1326261, with slightly over half living in ...



Inteligencia Artificial y modelos de lenguaje

Mejora de la calidad de los textos permitiendo ir a buscar a fuentes confiables

Invoking: `Wikipedia` with `Conde de la Viña`

Page: Concha y Toro

Summary: Concha y Toro is the largest producer and exporter of wines from Latin America and one of the ten largest wine companies in the world, with more than 33 million cases sold per year in 2014, distributed in 135 countries. Concha y Toro Winery is located in Santiago de Chile, but the company has vineyards in Chile, Argentina and the United States, with more than 10,000 hectares under cultivation in 2016. Their Casillero del Diablo and Don Melchor wines have international recognition. As of 2010 the company is one of the official sponsors of the English football club Manchester United F. C.

Invoking: `user_input` with `{'query':

'No se encontró información sobre la relación entre el pueblo de Calasanz y el conde de la Viña. ¿Hay algo más en lo que pueda ayudarte?}'`

Enter your answer: El conde de la Viña tenía una casa en Calasanz

Final answer:

Según la información disponible, no se encontró una relación específica entre el pueblo de Calasanz y el conde de la Viña. Sin embargo, mencionas que el conde de la Viña tenía una casa en Calasanz. Lamentablemente, no tengo acceso a información detallada sobre propiedades o residencias específicas de personas. Te recomendaría buscar información adicional en fuentes locales o históricas para obtener más detalles sobre la relación entre el conde de la Viña y su casa en Calasanz.



Inteligencia Artificial y modelos de lenguaje

Habilidades emergentes: habilidades no vistas en modelos más pequeños

Tres habilidades emergentes típicas que se han observado en los LLMs:

1. Aprendizaje en contexto (“In-context learning”)

Se proporcionan algunos ejemplos de entrada-salida (conocidos como "contexto")

Ejemplo 1:

Pregunta: ¿En qué año comenzó la Segunda Guerra Mundial?

Respuesta: La Segunda Guerra Mundial comenzó en 1939.

Ejemplo 2:

Pregunta: ¿Quién fue el primer presidente de los Estados Unidos?

Respuesta: El primer presidente de los Estados Unidos fue George Washington.

Ejemplo 3:

Pregunta: ¿Cuál es la capital de Francia?

Respuesta: La capital de Francia es París.

Nueva pregunta:

Pregunta: ¿Quién escribió "Don Quijote de la Mancha"?

Respuesta generada:

Respuesta: "Don Quijote de la Mancha" fue escrito por Miguel de Cervantes.



Inteligencia Artificial y modelos de lenguaje

2. Seguimiento de instrucciones (Instruction following)

El modelo entiende y responde de manera adecuada y precisa a las instrucciones, preguntas o tareas específicas que le sean proporcionadas.

Instrucción:

Instrucción: Describe brevemente un gadget futurista y su función principal en una sola frase.

Respuesta generada:

Respuesta: El "HoloPad" es una tableta holográfica portátil que proyecta pantallas 3D interactivas en el aire, permitiendo a los usuarios trabajar y jugar sin necesidad de una pantalla física.

3. Razonamiento paso a paso (step-by-step reasoning)

Los LLMs pueden realizar razonamientos complejos y resolver problemas paso a paso. Esto incluye habilidades matemáticas, lógicas y analíticas que no se enseñaron explícitamente pero que surgen de la vasta cantidad de datos y ejemplos en los que fueron entrenados.



Inteligencia Artificial y modelos de lenguaje

¿Cómo nos relacionamos con el modelo de lenguaje?

¿Qué es un “prompt”?

Instrucciones y contexto que pasamos a un modelo de lenguaje para lograr una tarea deseada.

Ingeniería de “Prompt”

Optimizar “prompts” para el uso eficiente de los modelos de lenguaje.

Un “prompt” está compuesto por

1. Instrucciones

Instrucción: Clasifica el texto en negativo, neutral o positivo

2. Contexto

Contexto: estamos en un restaurante de comida rápida

3. Datos de entrada

Texto: La hamburguesa me ha gustado aunque picaba un poco.

4. Salida

Sentimiento:
(Neutral/Positivo)

Inteligencia Artificial y modelos de lenguaje

Ejercicio:

Abrimos chatGPT (<https://chat.openai.com/>)

o Ollama (<http://signal4.cps.unizar.es:8505>)

Cambiar “La hamburguesa me ha gustado aunque picaba un poco.” por:

- “La hamburguesa me ha gustado muchísimo, se la recomiendo a mis amigos.”
- “La hamburguesa me ha sentado mal.”
- “La hamburguesa de este restaurante, ni fu ni fa”

Con este último texto, cambiar el contexto a:

- “estamos en un restaurante de comida con varias estrellas michelín”
¿qué ocurre con la valoración?



Inteligencia Artificial y modelos de lenguaje

Instrucción: Extraiga el nombre de las comunidades autónomas y ciudades en el siguiente texto.

Texto: La pipirrana tiene múltiples variantes en Andalucía, Castilla-La Mancha, Extremadura o Murcia, y hace siglos la hicimos en versión para estómagos sensibles. Hoy nos vamos al mismísimo Jaén a aprender a preparar la típica de allí en un escenario más propio imposible: la Casería de Jesús, una casa del siglo XVI en mitad de un olivar. A los mandos del dornillo -la fuente en la que se prepara en Jaén-, Ana María Gutiérrez, creadora del blog Cocinando entre olivos. Todo, en el vídeo de arriba.

Formato deseado:

Lugar: <lista separada por comas de nombres de comunidades autónomas y ciudades>

Instrucción: Extraiga los nombres de las personas que hayan escrito un blog y el nombre del blog en el siguiente texto.

Texto: La pipirrana tiene múltiples variantes en Andalucía, Castilla-La Mancha, Extremadura o Murcia, y hace siglos la hicimos en versión para estómagos sensibles. Hoy nos vamos al mismísimo Jaén a aprender a preparar la típica de allí en un escenario más propio imposible: la Casería de Jesús, una casa del siglo XVI en mitad de un olivar. A los mandos del dornillo -la fuente en la que se prepara en Jaén-, Ana María Gutiérrez, creadora del blog Cocinando entre olivos. Todo, en el vídeo de arriba.

Formato deseado:

Formato json con el nombre de la persona y el nombre del blog

Inteligencia Artificial y modelos de lenguaje

Ejemplo Comunicación Aumentativa y Alternativa

<https://arasaac.org/>



Problema de los tableros de comunicación:

Poca flexibilidad a la hora de crear frases

Verbos sin conjugar

Concordancias de género, número y tiempo

Ejemplos

Ainara ser científico, Elena estar contento y Nuria querer ser bombero. → Ainara es científica, Elena está contenta y Nuria quiere ser bombera.

Ayer olvidar llaves coche → Ayer olvidé las llaves del coche

<http://signal4.cps.unizar.es:8501/>

"Flexionar una frase" modificar la estructura gramatical o las palabras de una frase para adaptarla a diferentes contextos, necesidades comunicativas o estilos. Esto puede incluir cambios en el tiempo verbal, el número, el género, el grado de formalidad, entre otros aspectos gramaticales y estilísticos.

Habilidad emergente de los nuevos modelos de lenguaje de gran escala (GPT4, LLAMA3, ...)

Inteligencia Artificial y modelos de lenguaje

Técnicas de “prompting”

1. Zero-shot

no proporcionamos al modelo ningún ejemplo de texto.

Clasifica el texto en neutro, negativo o positivo.

Texto: Creo que las vacaciones están bien.

Sentimiento:

2. Few-shot

Proporcionamos al modelo ejemplos

Positivo ¡Esto es increíble!

¡Esto es malo! Negativo

¡Wow, esa película fue genial! Positivo

¡Qué horrible espectáculo! --



Inteligencia Artificial y modelos de lenguaje

Ejemplo de 1-shot con palabras que no existen

Prompt

Un "Tankuko" es un animal pequeño y peludo originario de Tanzania. Un ejemplo de una oración que usa la palabra tankuko es: Estábamos viajando por África y vimos estos lindos tankukos.

Hacer un "santule" significa saltar muy rápido. Un ejemplo de una oración que usa la palabra santule es:

GPT3.5

Estábamos corriendo una carrera y él santuleó sobre los obstáculos con increíble agilidad.

LLAMA3

El niño se emocionó y comenzó a santulear de alegría.

Phi3

Durante el festival local, los jóvenes bailaban al ritmo pulsante mientras realizaban sensacionales santules en el suelo del atardecer, capturando momentos inolvidables para todos los asistentes.



Inteligencia Artificial y modelos de lenguaje

3. Chain-of-thought (cadenas de pensamiento)
'Cadena de Pensamiento' (CoT) permite a un Modelo de Lenguaje Grande (LLM) abordar tareas complejas descomponiéndolas en los pasos constituyentes necesarios.
CoT permite capacidades de razonamiento complejo a través de pasos de razonamiento intermedios.

Prácticamente incorporado en la última generación de modelos de lenguaje



Inteligencia Artificial y modelos de lenguaje



Fui al mercado y compré 10 manzanas. Le di 2 manzanas a mi vecino y otras 2 al mecánico. Luego fui y compré 5 manzanas más y comí una. ¿Cuántas manzanas me quedaban?

Llama3:8b



Un problema de matemáticas delicioso!

Vamos a resolverlo paso a paso:

1. Compraste 10 manzanas al principio.
2. Le di 2 manzanas a tu vecino y otras 2 al mecánico, lo que significa que te quedaron:
 $10 - 2 \text{ (vecino)} - 2 \text{ (mecánico)} = 6 \text{ manzanas}$
3. Luego compraste 5 más, por lo que ahora tienes: $6 + 5 = 11 \text{ manzanas}$
4. Comiste una, así que te quedaban: $11 - 1 = 10 \text{ manzanas}$

Así que, después de todo eso, te quedaron 10 manzanas.



Inteligencia Artificial y modelos de lenguaje

Llama3:8b



traduce al español y resuelve el problema. it takes Amy 4 minutes to climb to the top of the slide. It takes her 1 minute to slide down. The water slide closes in 15 minutes. How many times can she slide before it closes?



¡Claro! La traducción al español es:

"Le lleva a Amy 4 minutos subir hasta el final de la piscina. Le lleva un minuto bajar por ella. La piscina cierra en 15 minutos. ¿Cuántas veces puede deslizarse antes de que cierre?"

Ahora, resolvamos el problema:

Amy necesita 4 minutos para subir y 1 minuto para bajar, lo que significa que cada ciclo (subir-bajar) le toma 5 minutos.

Si la piscina cierra en 15 minutos, podemos dividir ese tiempo entre el tiempo que Amy necesita para un ciclo:

$15 \text{ minutos} \div 5 \text{ minutos por ciclo} = 3 \text{ ciclos}$

Por lo tanto, Amy puede deslizarse 3 veces antes de que la piscina cierre.

Inteligencia Artificial y modelos de lenguaje

... y cuando pasamos a la multimodalidad (GPT4-visión)

PROMPT = ("The uploaded series of images is from a single video. "

"The frames were sampled every {FRAME_EXTRACTION_FREQUENCY_SECONDS} seconds. " "Make sure it takes about {FRAME_EXTRACTION_FREQUENCY_SECONDS // 2} seconds to voice the description of each frame. "

"Use exclamation points and capital letters to express excitement if necessary. " "Describe the video using Marv Albert style.")



"Alright, HERE WE GO! The game is TIED at 115! The crowd is on their FEET! The clock's ticking down! Dribbling at the top... He's waiting for the FINAL SHOT! JUST UNDER 6 SECONDS LEFT! He steps back... TAKES THE THREE... \n\nYES! HE NAILS THE THREE-POINTER AT THE BUZZER! OH, BABY, WHAT A SHOT! The crowd is going NUTS! Teammates are ECSTATIC! Celebration on the court as CONFETTI flies! \n\nThe arena is ELECTRIC, FANS are hugging, the team is swarmed with joy! What a CLUTCH PERFORMANCE to win the GAME and the SERIES! Unbelievable FINISH! Absolute MADNESS in the arena! And the hugs keep coming, WHAT A NIGHT!"



Inteligencia Artificial y modelos de lenguaje

¿Qué le falta?

Interaccionar con el “exterior”

Hasta ahora solo utilizamos el conocimiento del “modelo”

¿Es posible incluir información nueva o interaccionar con el “exterior”?

Colaboración de LLMs y herramientas externas

RAG – Retrieval Augmented Generation

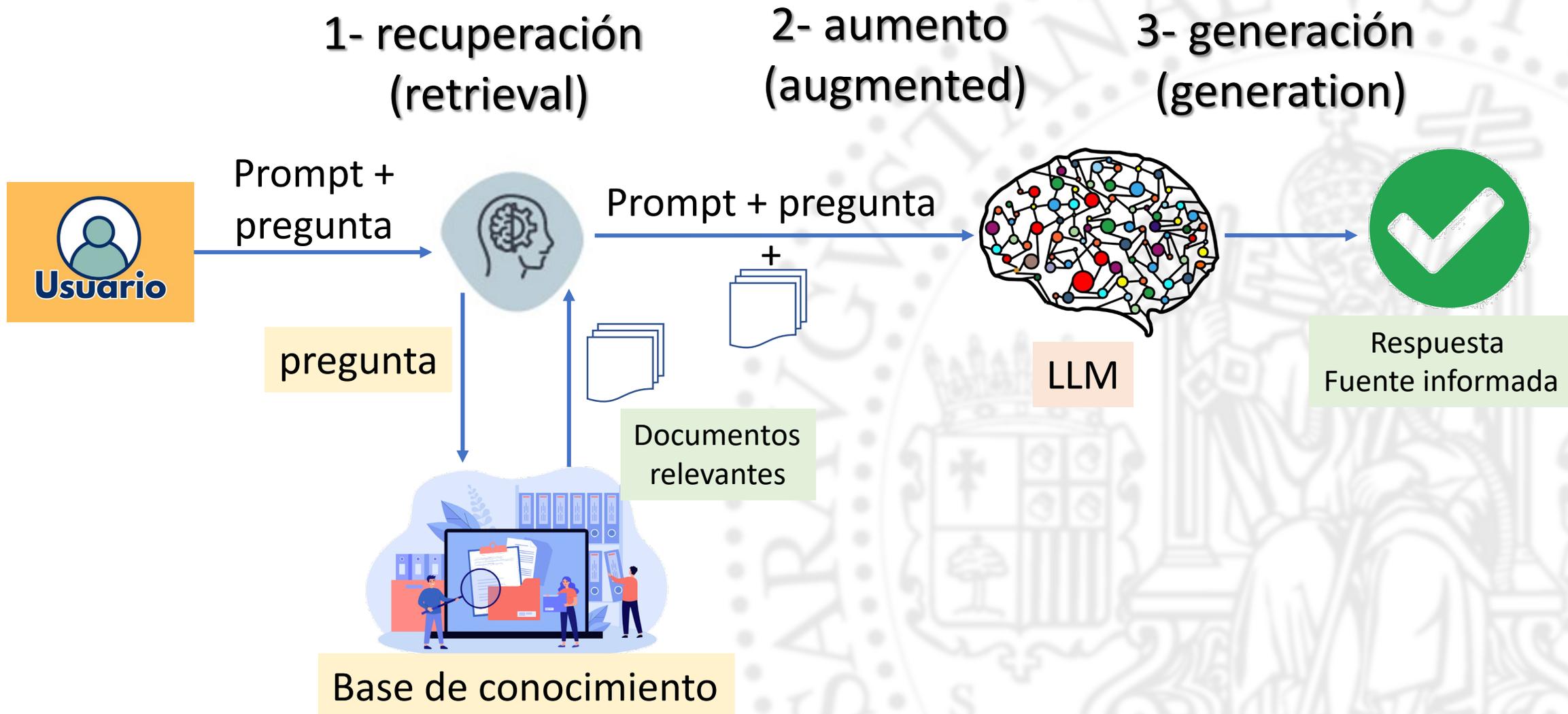
ReAct – Reasoning and Acting

Agentes autónomos

Inteligencia Artificial y modelos de lenguaje

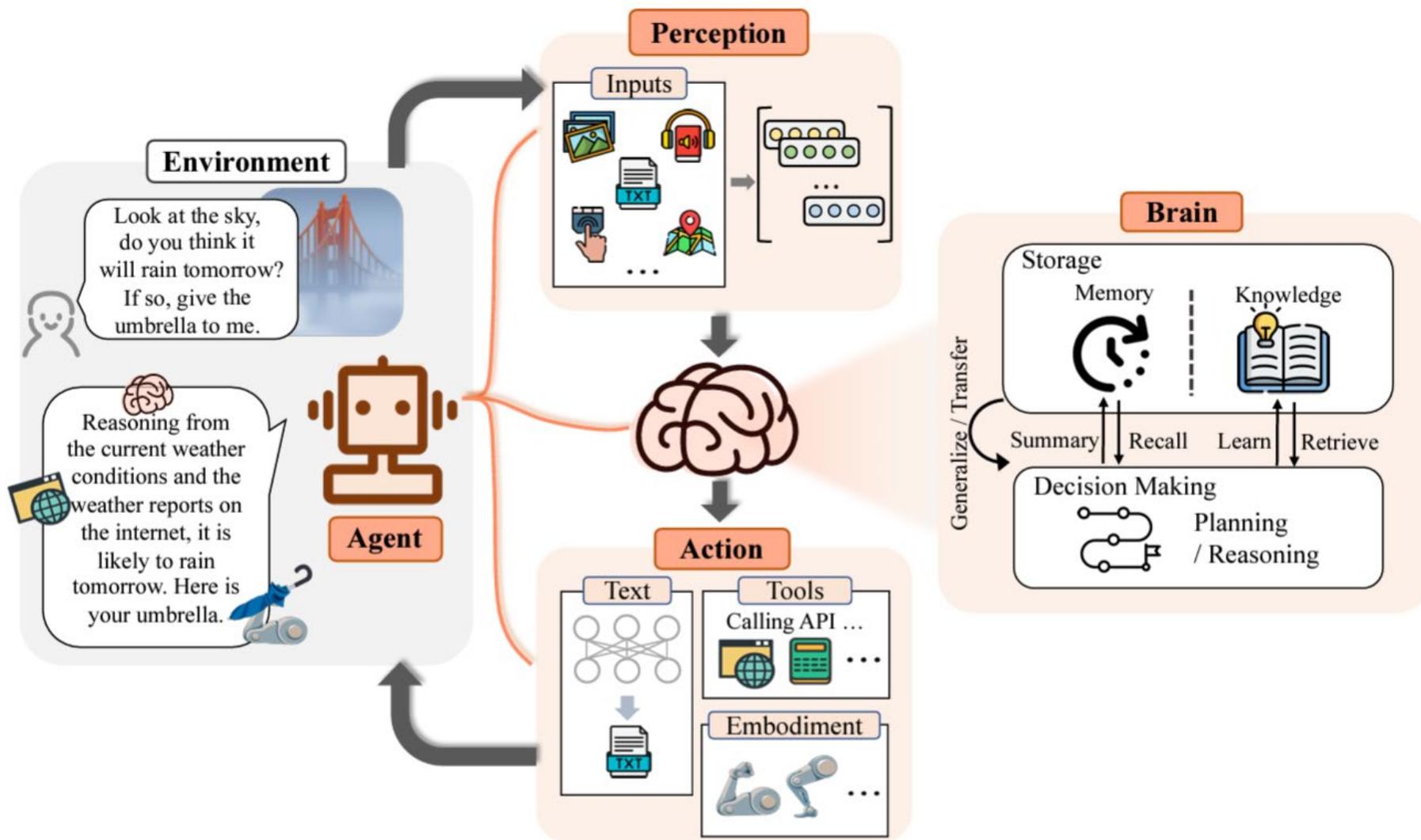
RAG - Retrieval Augmented Generation

3 fases



Inteligencia Artificial y modelos de lenguaje

Marco conceptual de Agentes autónomos



Inteligencia Artificial y modelos de lenguaje

Demos

Recuperación de información

Texto

<http://signal4.cps.unizar.es:8504/>

Multimodal

<http://signal4.cps.unizar.es:8509/>

RAG – Agente

<http://signal4.cps.unizar.es:8510/>

