

# Interacción Persona-Máquina

Agentes conversacionales basados en modelos de  
lenguaje

Profesor

Robiert Sepúlveda Torres

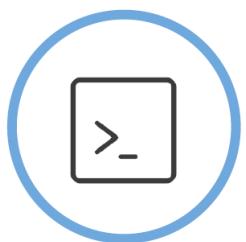


Universitat d'Alacant  
Universidad de Alicante

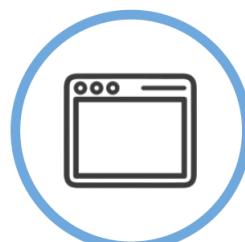


# Agentes conversacionales

## Motivación



Command Line



Graphical Interface



Natural Language Interface

# Agentes conversacionales

## Perspectiva del usuario

**Más intuitivo**, permite a las personas usar los sistemas computacionales de manera más natural

**Más atractivo**, la respuesta del sistema es más expresiva

**Más rápido**, permite a los usuarios completar ciertas tareas más rápido que usando una interfaz gráfica (GUI)

# Agentes conversacionales

## Perspectiva de la Industria

Experiencia de usuario mejorada

Permitir **nuevos** productos

Interacción más personalizada con la tecnología

Reducción de costos

# Agentes conversacionales

## Perspectiva de Investigadores

**Investigación básica y aplicada** sobre la interacción social y las necesidades de comunicación

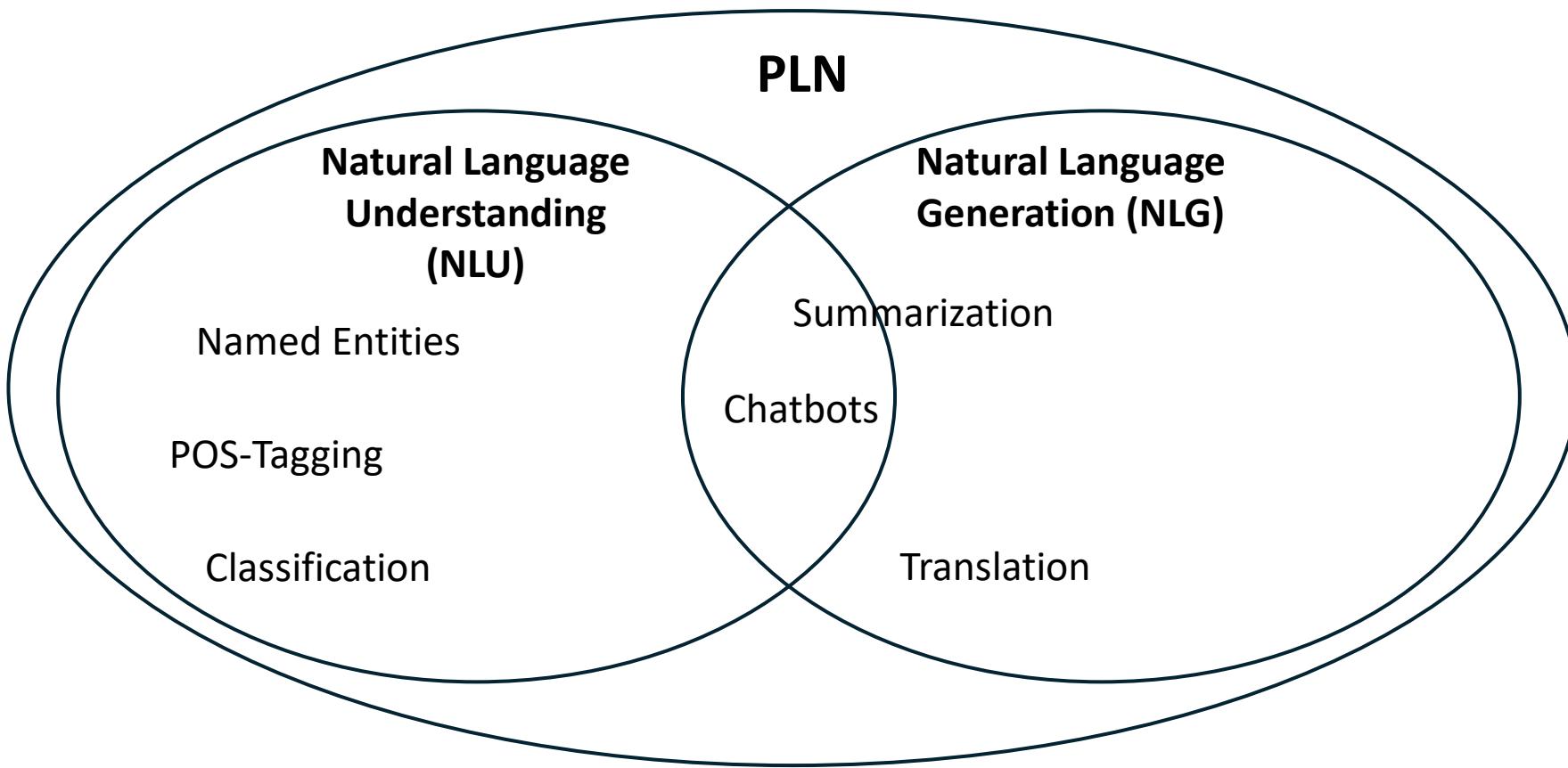
Se necesita **un nuevo lenguaje de diseño** para implementar este nuevo tipo de interfaz

Cómo construir **nuevas herramientas** para un procesamiento del habla más sólido y avanzado

Se requieren **nuevos marcos éticos y legales** para regular esta tecnología poderosa y potencialmente invasiva

# PLN-Agentes conversacionales

Ramas y tareas principales involucradas



# PLN-Agentes conversacionales

## NLU-NLG

### Natural Language Understanding (NLU)

La **comprensión del lenguaje natural (NLU)** utiliza análisis semántico y sintáctico para permitir que los ordenadores entiendan el lenguaje humano. El objetivo de NLU es comprender **la intención, el significado y el contexto**, en lugar de centrarse únicamente en el significado de palabras individuales.

### Natural Language Generation (NLG)

La **generación de lenguaje natural (NLG)** se encarga de generar lenguaje a partir de **datos estructurados y no estructurados**. Se encarga de estudiar como producir. Estudia como producir texto en lenguaje natural de **alta calidad, expresivo, claro y natural**.

# PLN-Agentes conversacionales

## Modelos de lenguaje

El lenguaje encierra en sí mismo una alta dependencia estadística entre sus componentes lo que se traduce en redundancia. (Dmitriev, 1991)



Ruleta de la suerte se sustenta en este principio

Ej: “Te dije que por aquí ...”

“Te dije que por aquí no era”

“Te dije que por aquí te ibas a caer”

# PLN-Agentes conversacionales

## Modelos de lenguaje

Tradicionalmente, el NLU y NLG se abordaban desde modelos de lenguaje (LM) muy específicos

Representación estadística que permite comprender y generar lenguaje natural

- Buenos en tareas muy concretas
- No generalizaban bien en contextos y tareas no vistas
- Comprensión limitada del texto

Actualmente:

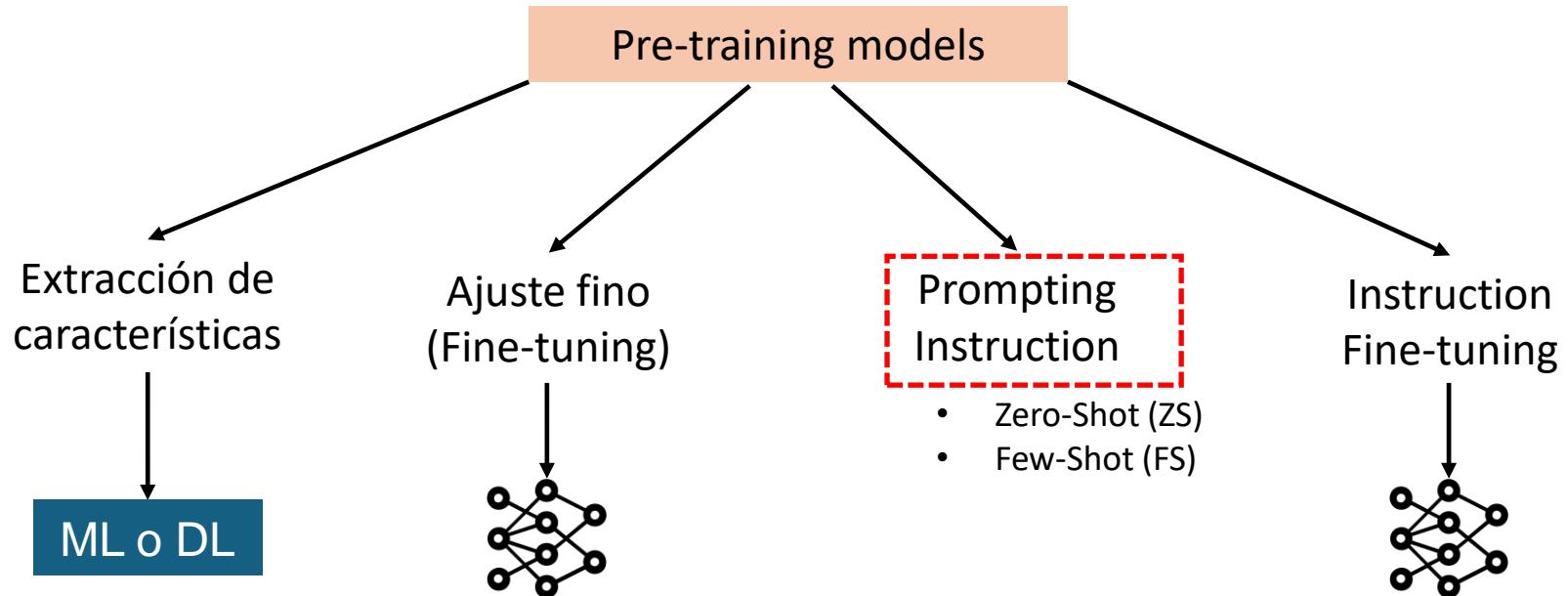
- Avances de hardware
- Arquitecturas modernas de redes neuronales
- Representaciones capaces de comprender el contexto del texto



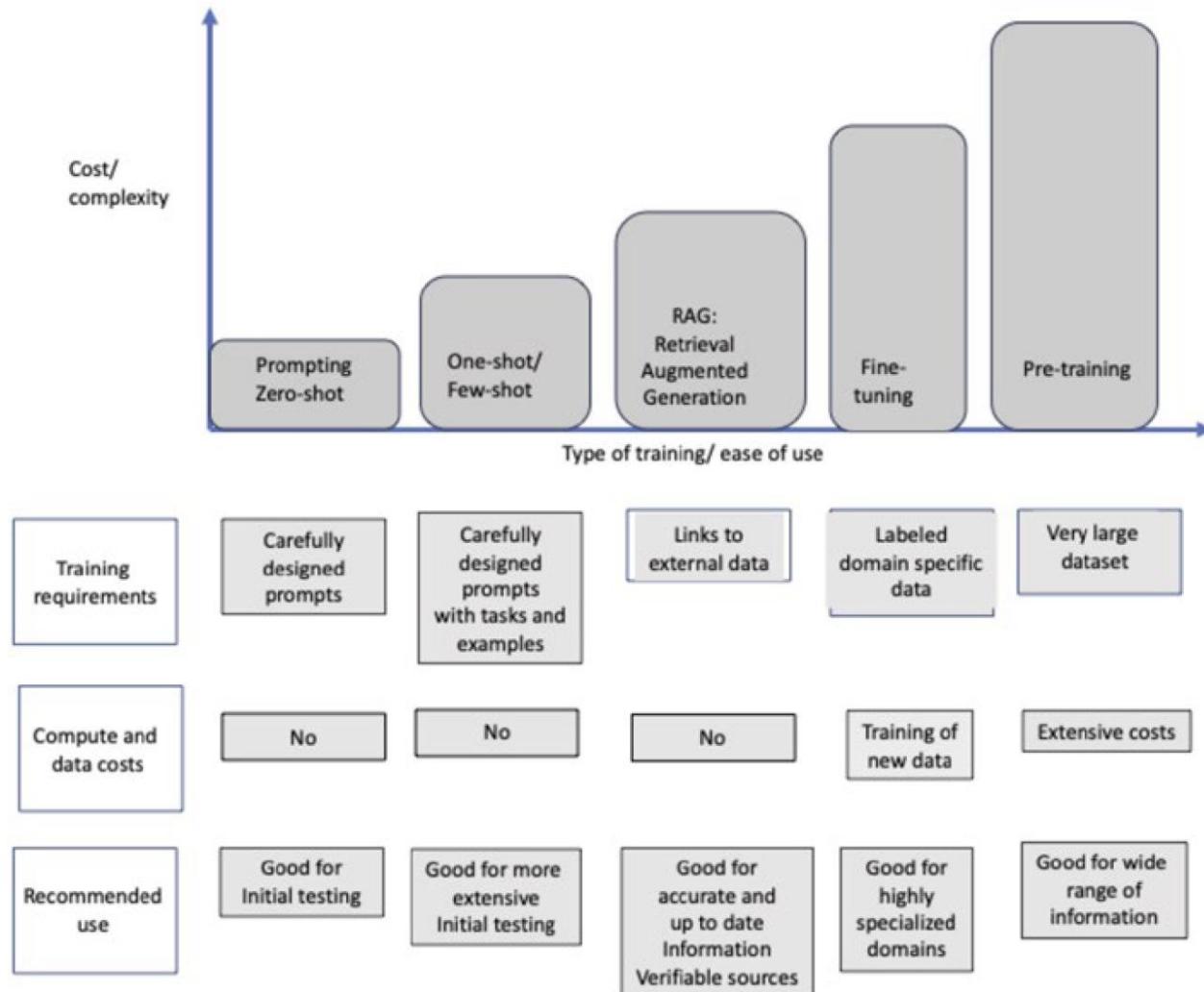
**LLM:** modelos de lenguaje más potentes que son capaces de exhibir grandes capacidades en la generación y comprensión de textos en un gran rango de tareas

Gran capacidad de transferencia de conocimiento entre dominio y tareas

# ¿Cómo podemos usar estos modelos de lenguaje?



# ¿Cómo podemos usar estos modelos de lenguaje?



Tomada de:  
McTear, M., &  
Ashurkina, M.  
Transforming  
Conversational AI.

# ¿Qué es la ingeniería de Prompts?

“It’s the ability to develop an intuition to get the most out of the model.”  Mira Murati (CTO de Open AI in 2024)

Técnica que utiliza una entrada inicial o ejemplo con el fin de guiar la predicción o generación de respuesta de un modelo en lenguaje natural

Explotar funcionalidades de los LLMs mediante lenguaje natural

Mejora la calidad y precisión de los resultados obtenidos

Permite la adaptación del modelo sin reentrenarlo

# ¿Qué es un Prompt?

Instrucciones o indicaciones que reciben los LLM para crear textos similares a los humanos mediante modelos generativos (IAG)

Medio de comunicación entre el usuario y el modelo

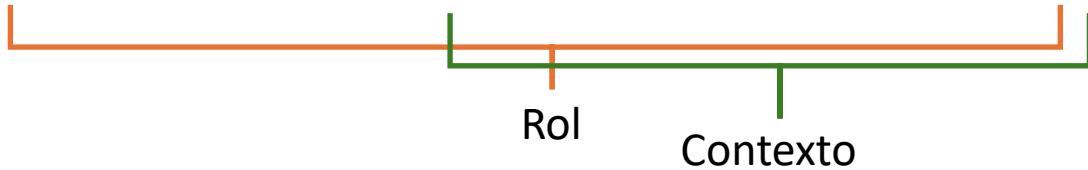
Diferencia entre que un modelo entienda o malinterprete tu consulta

Es importante considerar dos factores en los prompts para maximizar la eficacia de un modelo de lenguaje: **Contexto e Instrucción**

# ¿Qué es un Prompt?

**Contexto:** ayuda al modelo a comprender mejor el escenario

*“Como experto en planificación financiera”*



**Instrucción:** especifica la acción que se quiere llevar a cabo

*“Necesito que me des pautas para maximizar la ganancia en mis inversiones”*

# Ejemplos de Prompts

El chico se lo cree 

Como experto en planificación financiera.

Necesito que me des pautas para maximizar la ganancia en mis inversiones.



¡Perfecto! Como experto en planificación financiera, puedo darte un **marco estratégico claro** para maximizar la rentabilidad de tus inversiones **minimizando riesgos innecesarios**.

A continuación te presento las **pautas fundamentales**, estructuradas en 5 pilares clave:

## 1. Define objetivos claros y un horizonte temporal

Antes de invertir, responde con precisión:

-  **Plazo:** corto (1–3 años), medio (3–7 años) o largo (7+ años).
-  **Meta:** ¿Quieres ingresos pasivos, crecimiento patrimonial, retiro anticipado o preservar capital?
-  **Tolerancia al riesgo:** ¿Cuánto estás dispuesto a perder en el corto plazo sin entrar en pánico?

**⚠ Ejemplo:** una estrategia agresiva puede tener más volatilidad, pero rendimientos superiores a largo plazo; una conservadora prioriza estabilidad sobre ganancias.

# Ejemplos de Prompts

¿Por qué crees que se esconde el sol?

¡Qué bonita pregunta! ☀️✨

Depende de **cómo la mires**:

 **Desde una perspectiva científica**

El sol no se esconde en realidad —somos nosotros los que **giramos**.

La **rotación de la Tierra** hace que en ciertos momentos de la jornada un lugar quede de espaldas al sol, y eso es lo que llamamos **atardecer o noche**.

👉 Entonces, cuando “se esconde el sol”, en realidad estamos **viendo cómo la Tierra gira** y nos aleja de su luz directa.

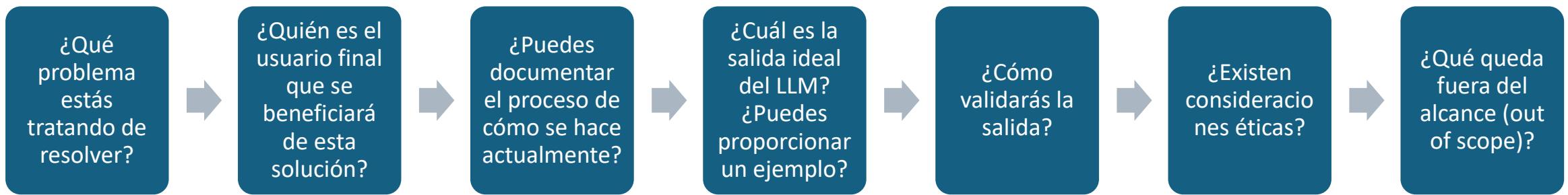
**Dar contexto adicional depende de la tarea, pero suele mejorar las respuestas del modelo**

# Ejemplos de Prompts

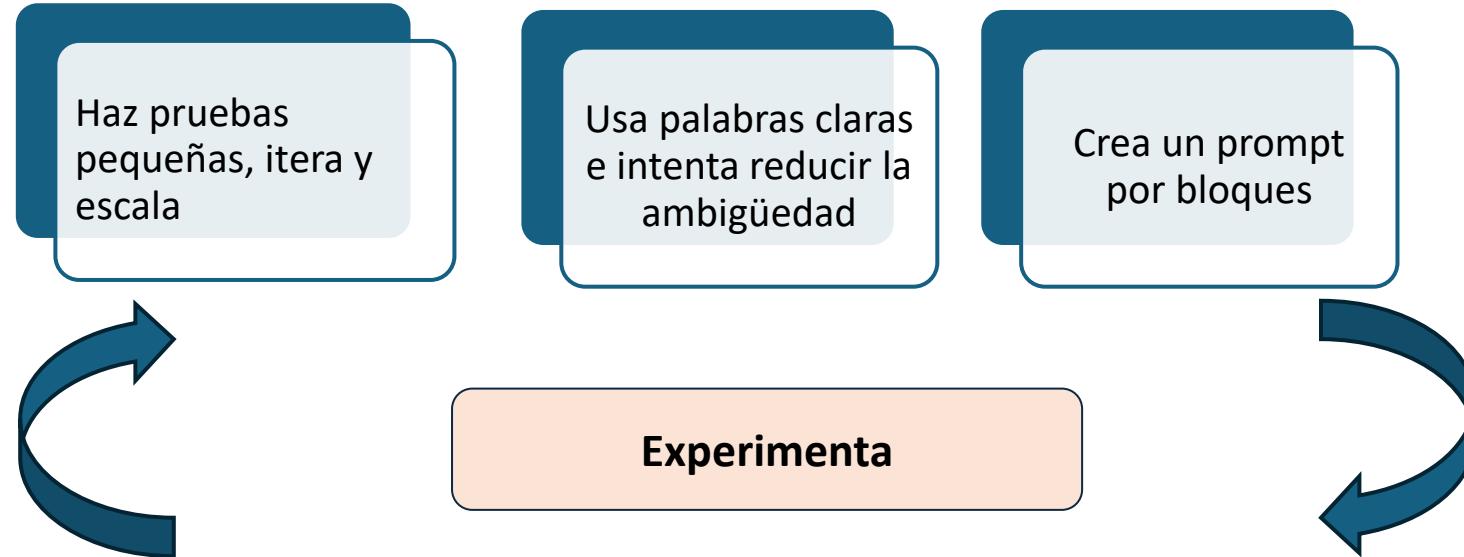
| Tareas                 | Prompts   |
|------------------------|---|
| Resumen de textos      | Realiza un resumen del siguiente documento:     |
| Traducción             | Traduce las siguientes oraciones a Español:     |
| Respuestas a preguntas | ¿Me puedes explicar por qué crecen las plantas? |

# ¿Cómo construir un buen prompt?

## Define el caso de uso

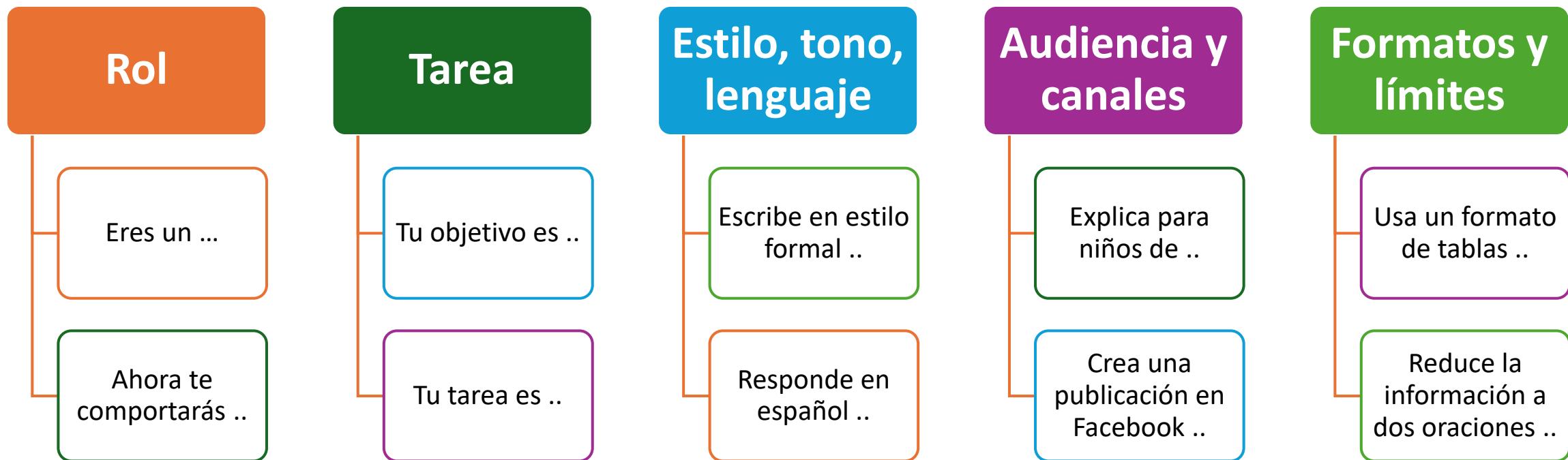


# ¿Cómo construir un buen prompt?



# ¿Cómo construir un buen prompt?

## Prompt por bloques



# Tipos de Prompts

**Zero-Shot Prompting:**  
no se da contexto ni  
ejemplos previos

**Few-Shot Learning:** se  
dan algunos ejemplos  
(1-5). Patrones y estilo

**Chain-of-Thought  
Prompting:** se pide que  
se detallen los pasos  
(razonar)

**Negative Prompting:** se  
especifica al modelo  
qué no tiene que hacer

**Hybrid Prompting:**  
combinación de  
diferentes métodos

**Prompt Chaining:** se  
divide una tarea  
compleja en órdenes  
más pequeñas

**ReAct (Reason-Action):**  
razonamiento y  
acciones intermedias  
para usar herramientas

**Self-Consistency:** se  
generan varios caminos  
de razonamiento para  
que el modelo evalúe  
cuál es mejor

**Self-Reflection  
Prompting:** se utiliza el  
propio modelo para  
evaluar su respuesta

# Tipos de Prompts

## Zero-Shot Prompting

Clasifica el siguiente mensaje en “tóxico” o “no tóxico”: [texto]

## Few-Shot Learning

Clasifica el siguiente comentario en “tóxico” o “no tóxico”.

Ejemplo: Eres un ladrón que se aprovecha de los demás: tóxico.

Comentario: [texto]

## Chain-of-Thought Prompting

Clasifica el siguiente comentario en “tóxico” o “no tóxico”.

Primero, razona paso a paso porqué el mensaje puede o no ser tóxico (por ejemplo: presencia de insultos, amenazas, lenguaje violento, etc.).

Luego, da tu respuesta final con las siguientes categorías: “tóxico” o “no tóxico”.

# Tipos de Prompts

## Negative Prompting

Analiza el siguiente comentario y responde si es “tóxico” o “no tóxico”.

- No incluyas términos ambiguos o respuestas que no sean “tóxico” o “no tóxico”.
- Evita cualquier tipo de análisis detallado o explicaciones adicionales

## Prompt Chaining

**Prompt\_1:** Clasifica el siguiente comentario en “tóxico” o “no tóxico”. [texto]

Respuesta: tóxico

**Prompt\_2:** ¿Puedes explicar por qué este mensaje es tóxico?

Respuesta: El mensaje es tóxico porque ...

**Prompt\_3:** ¿Puedes identificar los elementos en el mensaje que denotan toxicidad?

# Tipos de Prompts

## ReAct Prompting

Eres un asistente de domótica. Tu objetivo es ayudar a ejecutar tareas en el hogar. Debes seguir el patrón **Think–Action–Result**:

**Think** : razona paso a paso qué se debe hacer para cumplir la instrucción del usuario.

**Action**: describe claramente la acción que realizarías (por ejemplo, llamar a una API, activar un dispositivo, etc.).

**Result**: explica de forma breve y clara el resultado final esperado.

**Ejemplos:**

**Usuario**: “Enciende las luces del salón.”

**Think**: La instrucción implica encender la iluminación en la habitación “salón”. Debo identificar el dispositivo correcto.

**Action**: Activar luces → **encender\_luces(salón)**

**Result**: “Las luces del salón están encendidas.”

**Usuario**: “Pon la temperatura a 22 grados.”

**Think**: El usuario quiere ajustar el termostato principal a 22 °C.

**Action**: Ajustar termostato → **set\_termostato(22)**

**Result**: “La temperatura se ha ajustado a 22 grados.”

# Tipos de Prompts

## Self-Consistency Prompting

Clasifica el siguiente comentario como “tóxico” o “no tóxico” aplicando self-consistency: genera internamente varios caminos de razonamiento independientes, compara sus conclusiones y elige la respuesta más consistente. No reveles tu razonamiento interno. Devuelve solo:

1. etiqueta: “tóxico” | “no tóxico”
2. justificación breve (1 frase)
3. hasta 2 evidencias (citas textuales muy cortas del comentario)

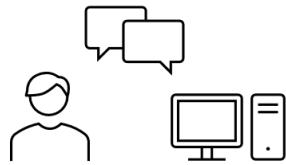
## Self-Reflection Prompting

Clasifica el siguiente comentario como “tóxico” o “no tóxico”. Genera una respuesta inicial con una justificación breve y evidencias.

1. Luego, reflexiona críticamente sobre esa respuesta: evalúa si es correcta, si falta información o si se cometió algún error.
2. Si es necesario, corrige la respuesta y entrega la versión final mejorada.

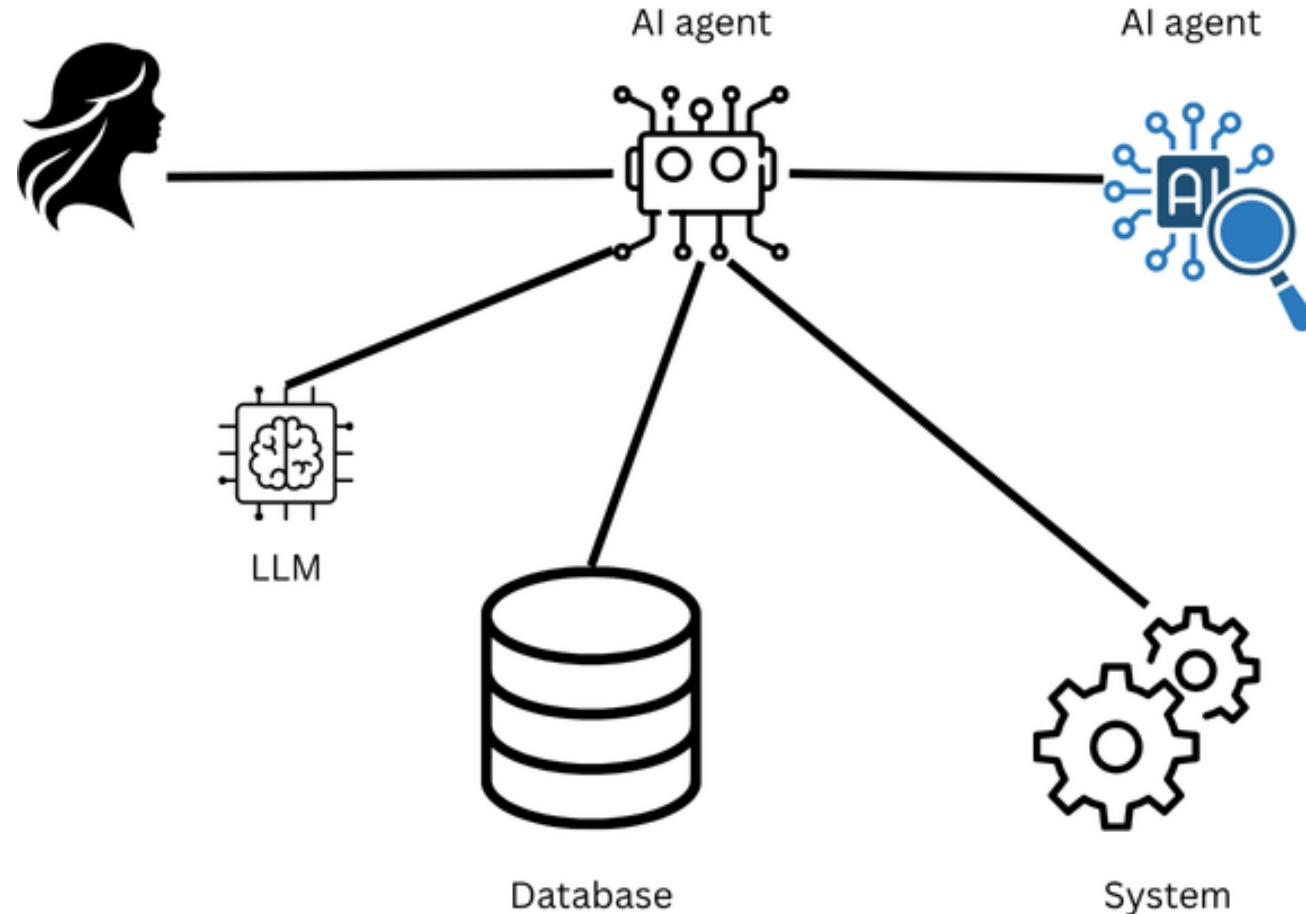
# ¿Qué es la ingeniería de Prompts?

Ingeniería de prompts o  
*Prompt engineering*



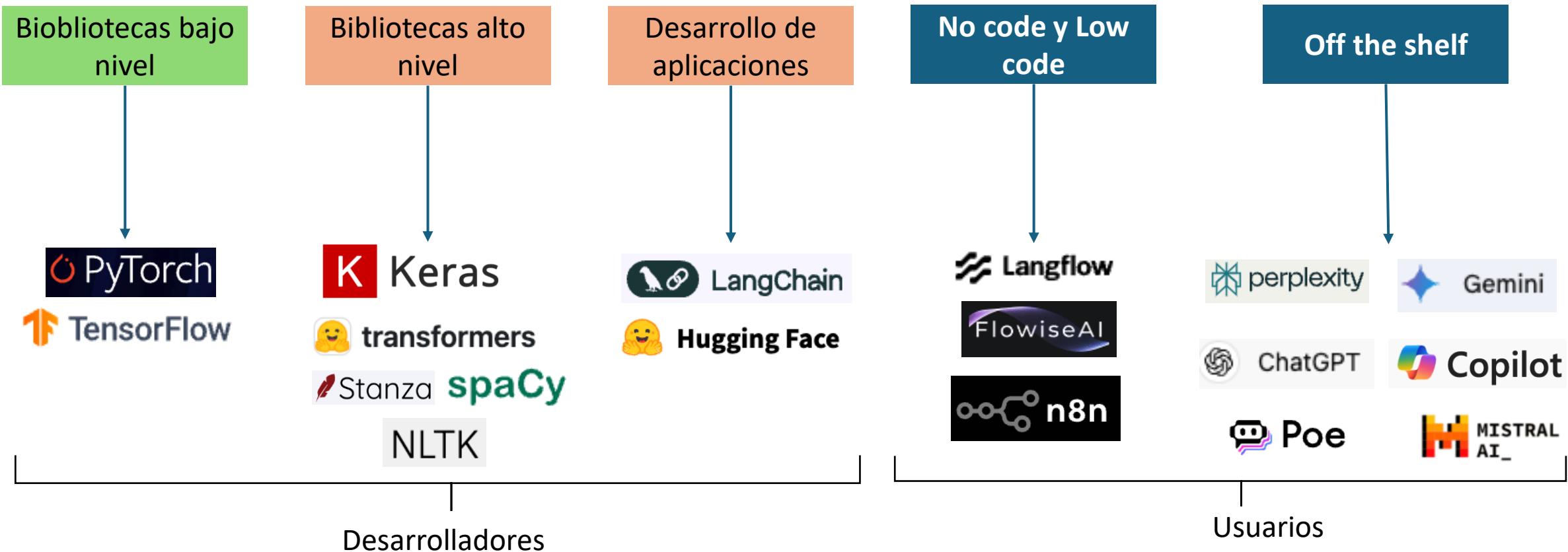
- Forma en la que nos comunicamos con un modelo generativo
- Desde lenguaje natural fluido hasta instrucciones explícitas
- + **pautado** = respuesta más precisa
- + **largo** = más probabilidad de perderse por el camino

# ¿Qué tipo de sistemas deberíamos construir en la actualidad?



# ¿Con qué herramientas contamos actualmente?

## Herramientas de PLN



# Conclusiones

- Podemos aprovechar el potencial de los LLMs para muchas tareas si diseñamos prompt adecuados.
- No existe una estrategia única que funcione en todos los contextos, por lo que tenemos que experimentar y ajustar.
- Existen otras estrategias más específicas que permiten adaptar aún más el modelo a una tarea (Fine-tuning o Instruction-tuning).
- Deberíamos usar prompts específicos para evaluar la calidad en la generación de estos modelos.