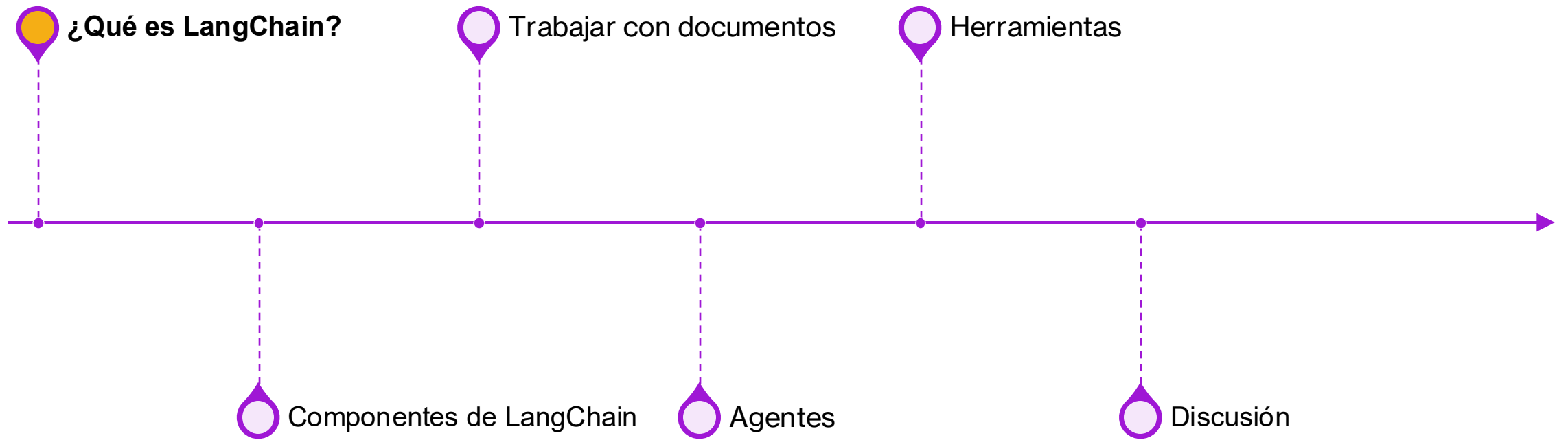




# DESCUBRIENDO EL PODER DE LANGCHAIN:

Creando aplicaciones inteligentes

# Indice





# ¿QUÉ ES LANGCHAIN?

Librería de python que permite hacer aplicaciones que interactúan con **Large Language Models (LLMs)**. Fue publicada el **17 de Octubre de 2022**



# ¿QUÉ ES UN LLM?

Un tipo de modelo de deep learning tipo **transformer** entrenado con **grandes cantidades** de datos capaz **entender y generar lenguaje humano** de manera natural.

# — ¿Que hace un LLM?

Encoder:

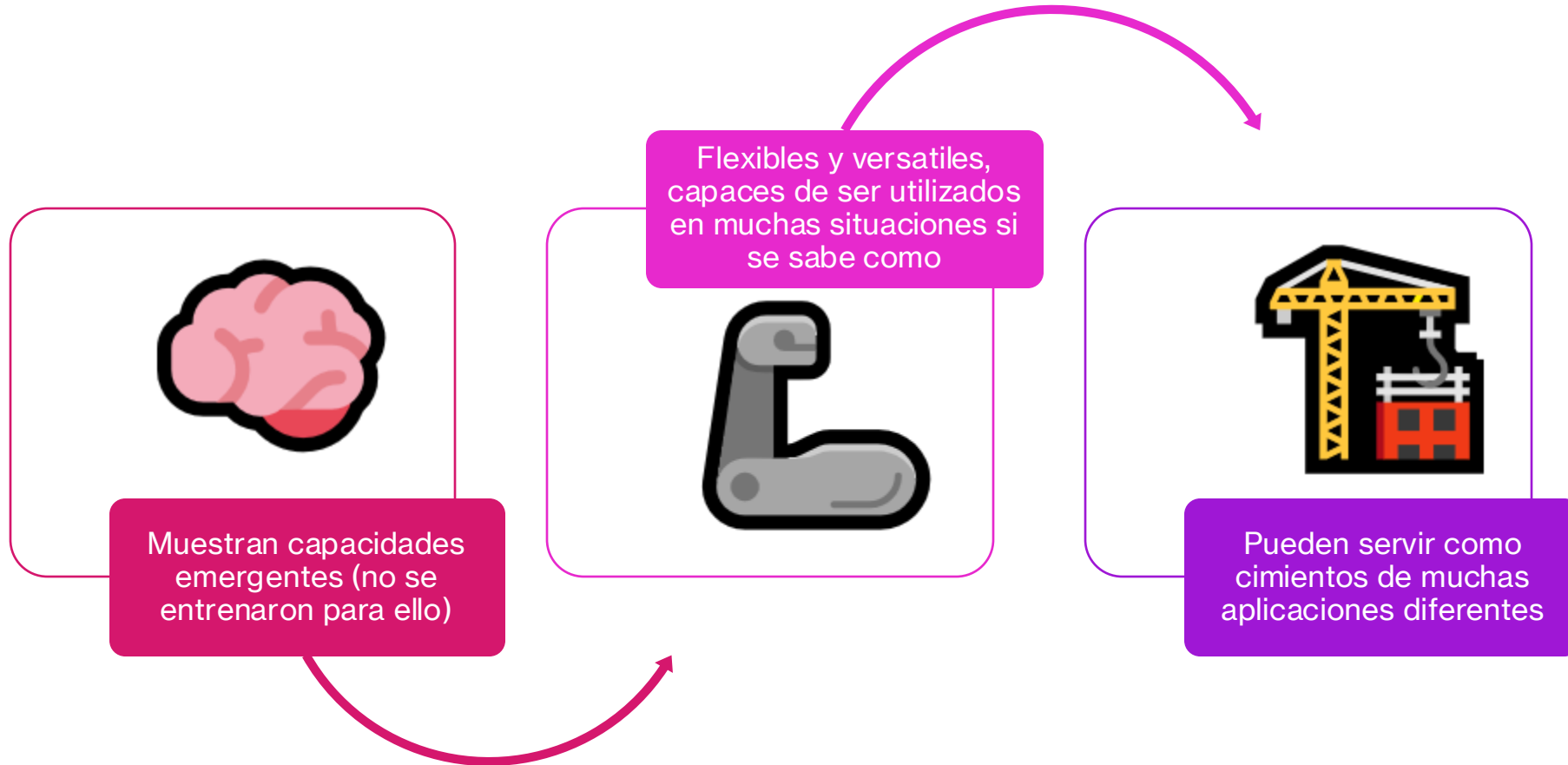
- vectores con significado semántico.

Decoder:

- continuar un texto dado



# De donde viene el *Hype* por los LLMs



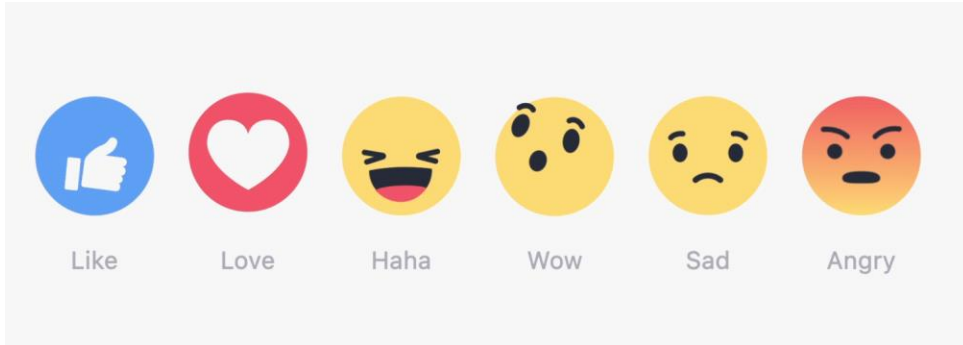


Los LLMs se pueden considerar  
**motores de razonamiento**  
y reducen el esfuerzo de  
desarrollar aplicaciones  
inteligentes de manera significativa



por los  
LLMs

# Ejemplo real: Reacción o tópico



- Analizar temáticas con mensajes de redes sociales es complicado.
- La mayoría de mensajes son reacciones a otros mensajes y por lo tanto no tienen temática definida.
- Un LLM puede discernir sobre que mensajes son reacciones y cuales tratan un tema concreto.
- Para ello solo necesito escribir un **prompt**.



# Ejemplo

You are a text classification assistant trained to analyze messages from social networks. Your task is to categorize each message into one of three categories: 'Topical', 'Reactive'.

- A Topical message discusses a clear, specific topic or issue.
- A Reactive message is a response or reaction to another message and doesn't focus on a specific topic.

Classify each message accurately based on these criteria and following this instructions:

Instructions: {format\_instructions}.

Here are some examples to guide you:

Example 1:

Message: 'The new environmental policies will impact carbon emissions significantly.'

Classification: topical

Example 2:

Message: "I completely agree with what you said!"

Classification: reactive

Example 3:

Message: "Artificial intelligence will change the future of healthcare."

Classification: topical

Example 4:

Message: "Interesting point, but have you considered other perspectives?"

Classification: reactive

TASK:

Message: "{message}"

Classification:



Like



Love



Haha



Wow

- Analizar temáticas

- La mayoría de me

tanto no tienen te

- Un LLM puede dis

tratan un tema co

- Para ello solo nec

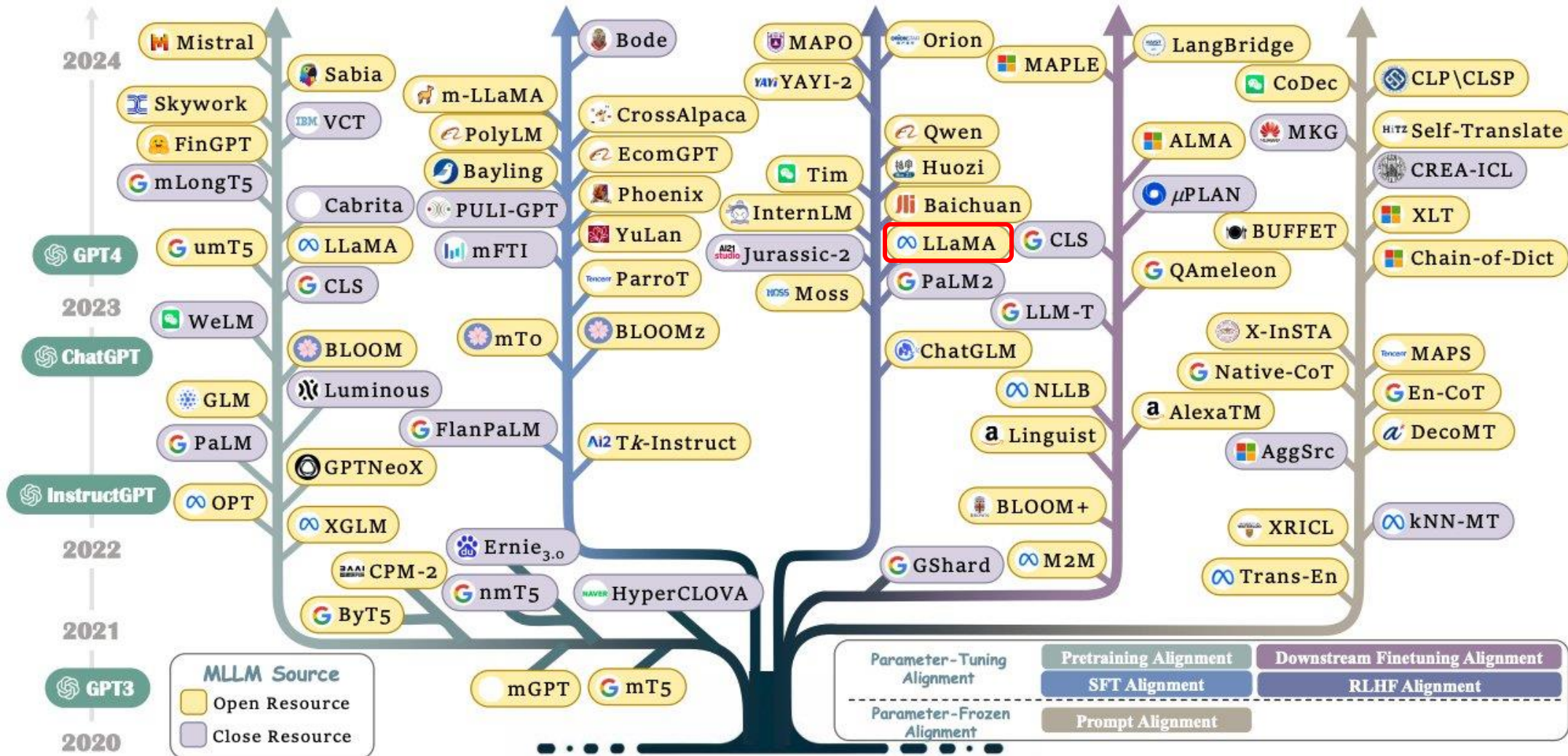


# LANGCHAIN

⚡ Permite aprovechar las capacidades de los LLM para desarrollar aplicaciones inteligentes



# ¿ CUALES HAY?



Overview of Multilingual

# LLama 3.1

---

- Publicado el **23 de julio de 2024** con licencia **open-source**.
- Entrenado en la generación de texto y en modo chat.
- Tres tamaños
  - Pequeño (8b): 8 Billones de parámetros
  - Mediano (70b): 70 Billones de parámetros
  - Grande (405b): 405 Billones de parámetros
- Disponible en version *raw* e *instruct*.
- Código en Ollama  
<https://github.com/ollama/ollama>





# Consumo de memoria

Model Size	FP16	FP8	INT4
8B	16 GB	8 GB	4 GB
70B	140 GB	70 GB	35 GB
405B	810 GB	405 GB	203 GB

Model Size	1k tokens	16k tokens	128k tokens
8B	0.125 GB	1.95 GB	15.62 GB
70B	0.313 GB	4.88 GB	39.06 GB
405B	0.984 GB	15.38	123.05 GB



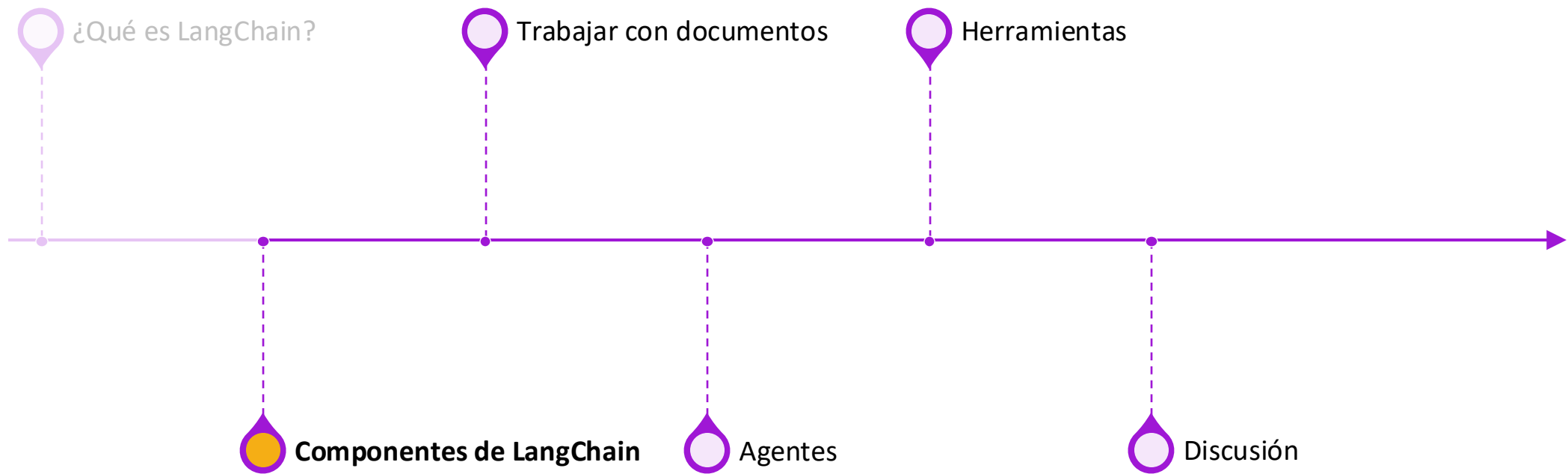
# Caso de uso

---

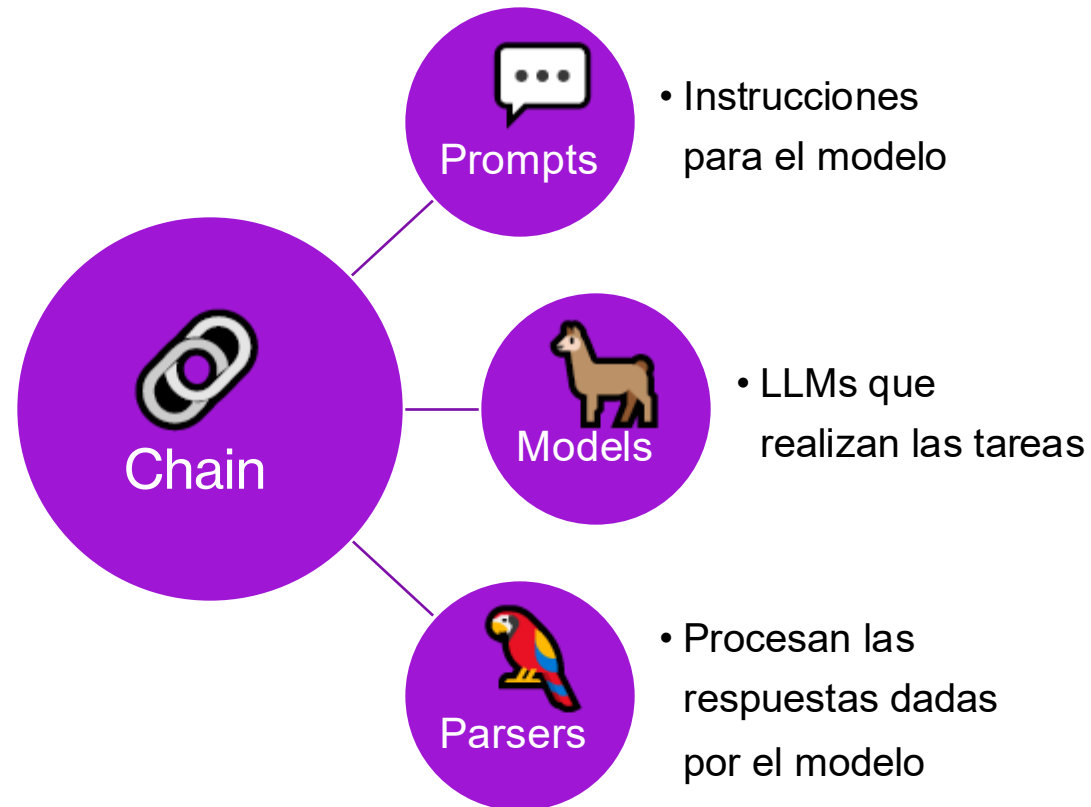
- Aplicación que simula ser el sombrero seleccionador
- El sombrero es el encargado de elegir en qué casa de Hogwarts debe ir un alumno según su personalidad.
  - 🦁 **Gryffindor**: Valientes, atrevidos, caballerosos y audaces.
  - 🦫 **Hufflepuff**: Leales, trabajadores, justos y pacientes.
  - 🦅 **Ravenclaw**: Inteligentes, curiosos, creativos y sabios.
  - 🐍 **Slytherin**: Ambiciosos, astutos, ingeniosos y decididos.



# Indice



# Componentes básicos de LangChain





# Consejos prompt engineering

- **Es un proceso muy artesanal**
- Depende del modelo, conoce sus limitaciones
- Incluye ejemplos
- Divide las tareas en etapas
- Da hueco al modelo para que razone
- Se especifico con la salida



# Ejemplo

---

Act as the sorting hat and tell to what Hogwarts house will {name} go. Hogwarts houses have the next traits:

Gryffindor: Brave, daring, chivalrous, and bold. They value courage and are often seen as leaders willing to stand up for what's right.

Hufflepuff: Loyal, hardworking, fair, and patient. Hufflepuffs prize dedication and kindness, with a strong sense of justice and inclusivity.

Ravenclaw: Intelligent, curious, creative, and wise. Ravenclaws value learning, wit, and wisdom, always seeking knowledge and understanding.

Slytherin: Ambitious, cunning, resourceful, and determined. Slytherins are driven, often setting high goals and valuing loyalty within their group.

Always follow this instructions when answering:

Select one of the following options: gryffindor, hufflepuff, ravenclaw, slytherin only return one option with no explanation, always reply in lower case.

# Ejemplo: Pon ejemplos

---

Act as the sorting hat and tell to what Hogwarts house will {name} go. Hogwarts houses have the next traits:

**Gryffindor:** Brave, daring, chivalrous, and bold. They value courage and are often seen as leaders willing to stand up for what's right.

**Hufflepuff:** Loyal, hardworking, fair, and patient. Hufflepuffs prize dedication and kindness, with a strong sense of justice and inclusivity.

**Ravenclaw:** Intelligent, curious, creative, and wise. Ravenclaws value learning, wit, and wisdom, always seeking knowledge and understanding.

**Slytherin:** Ambitious, cunning, resourceful, and determined. Slytherins are driven, often setting high goals and valuing loyalty within their group.

Always follow this instructions when answering:

Select one of the following options: gryffindor, hufflepuff, ravenclaw, slytherin only return one option with no explanation, always reply in lower case.

# Ejemplo: Especifica formato salida

---

Act as the sorting hat and tell to what Hogwarts house will {name} go. Hogwarts houses have the next traits:

Gryffindor: Brave, daring, chivalrous, and bold. They value courage and are often seen as leaders willing to stand up for what's right.

Hufflepuff: Loyal, hardworking, fair, and patient. Hufflepuffs prize dedication and kindness, with a strong sense of justice and inclusivity.

Ravenclaw: Intelligent, curious, creative, and wise. Ravenclaws value learning, wit, and wisdom, always seeking knowledge and understanding.

Slytherin: Ambitious, cunning, resourceful, and determined. Slytherins are driven, often setting high goals and valuing loyalty within their group.

Always follow this instructions when answering:

Select one of the following options: gryffindor, hufflepuff, ravenclaw, slytherin only return one option with no explanation, always reply in lower case.

# Ejemplo: Dejele espacio para pensar



```
{  
  name: str,  
  house: int  
}
```



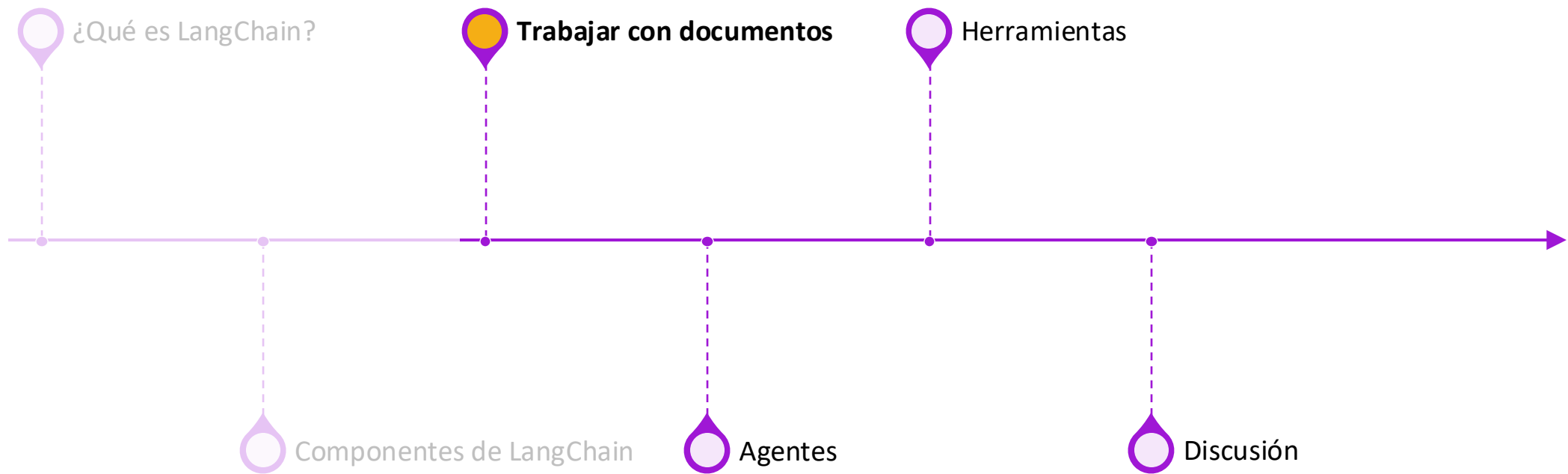
```
{  
  name: str,  
  house: int,  
  reason: str  
}
```

# Limitaciones: Ventana de contexto

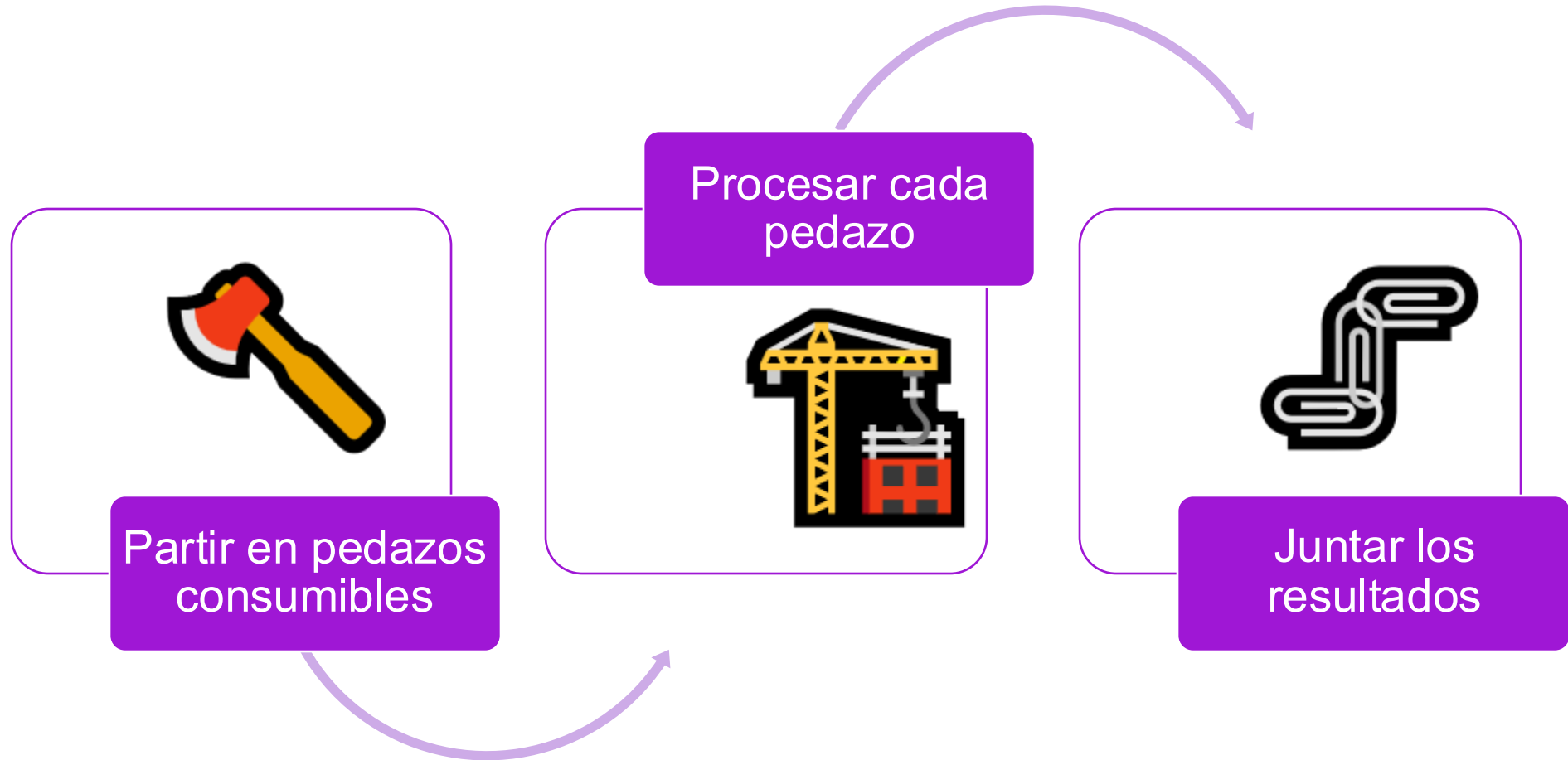
- Los LLM tienen un número máximo de texto que pueden procesar.
- Ventana de contexto = N°
- LLama 3.1 ventana 128k tokens



# Indice

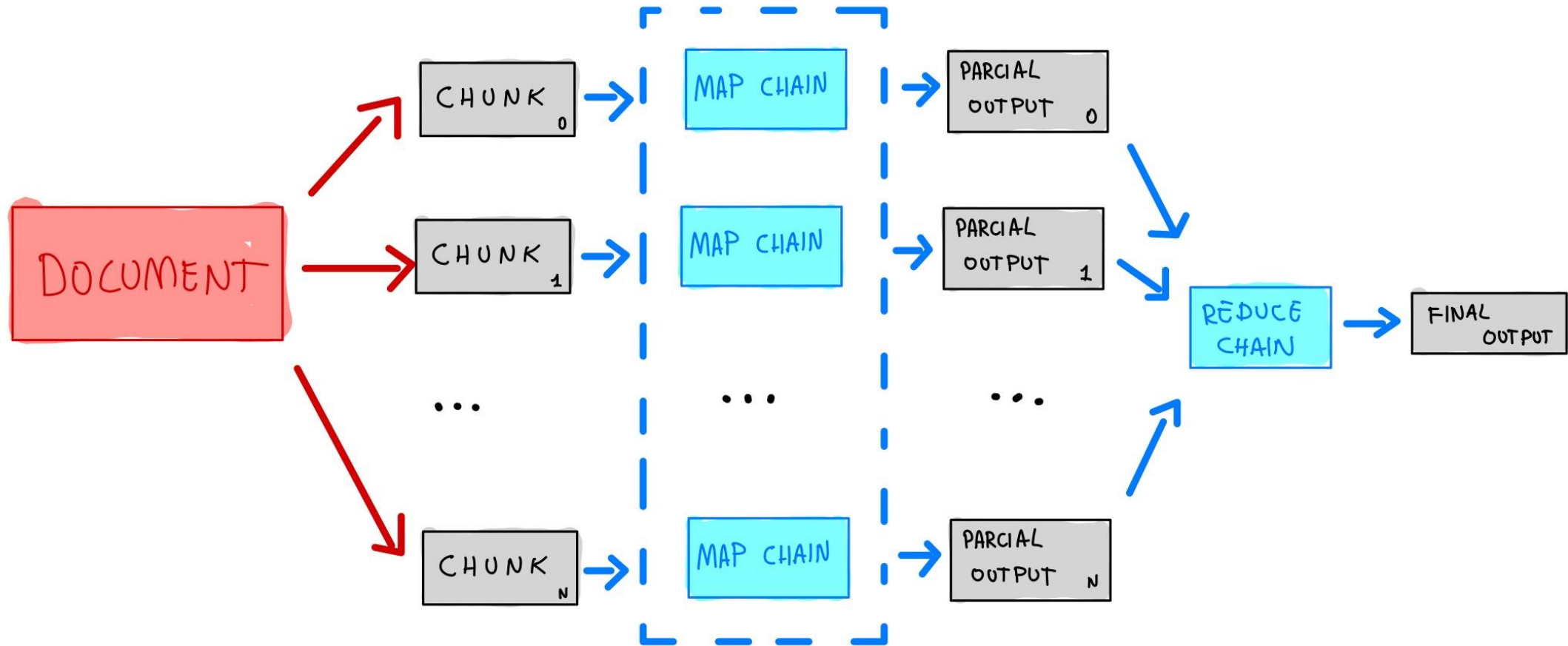


# Solventar limitacion

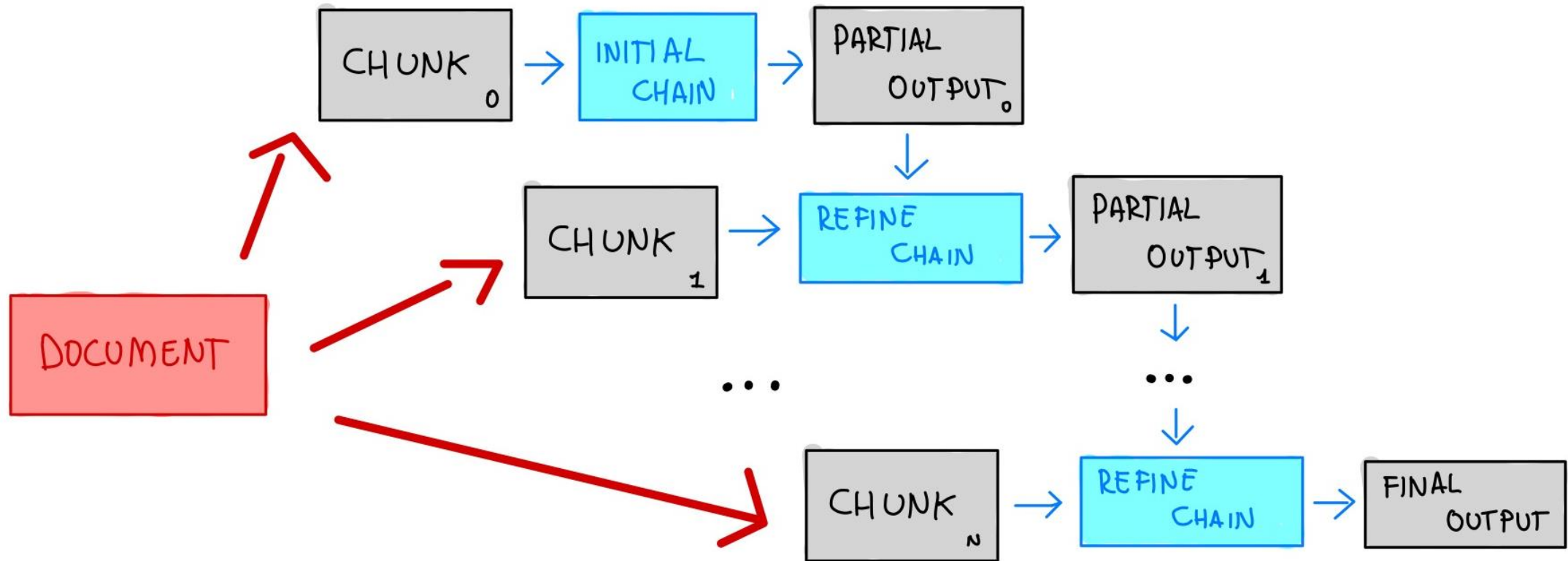




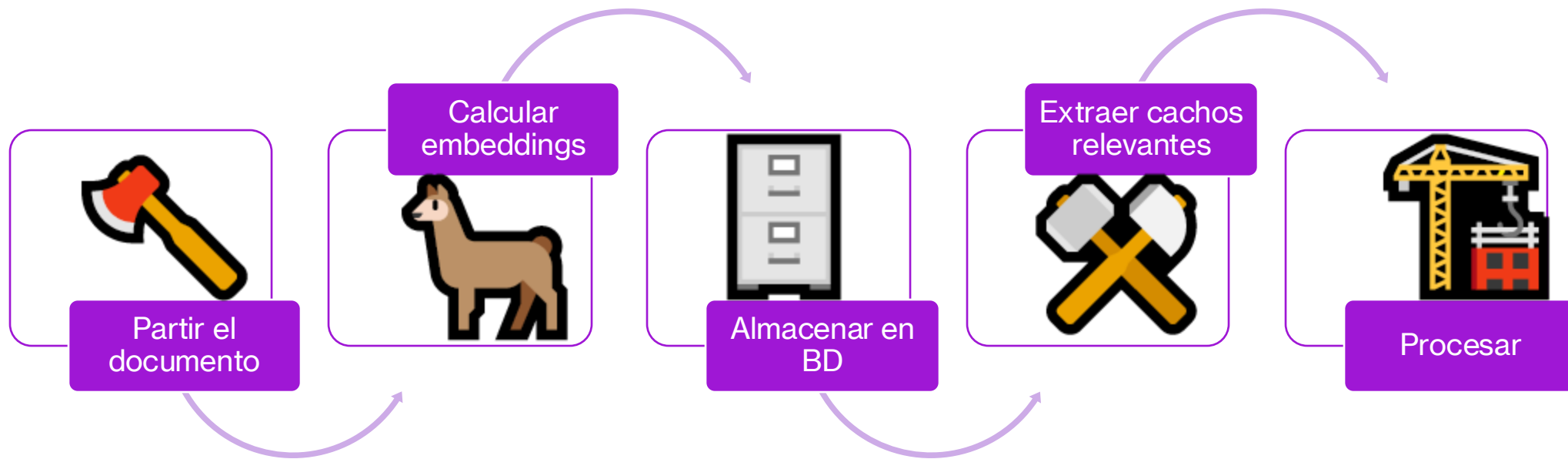
# Juntar los resultados: Map - Reduce



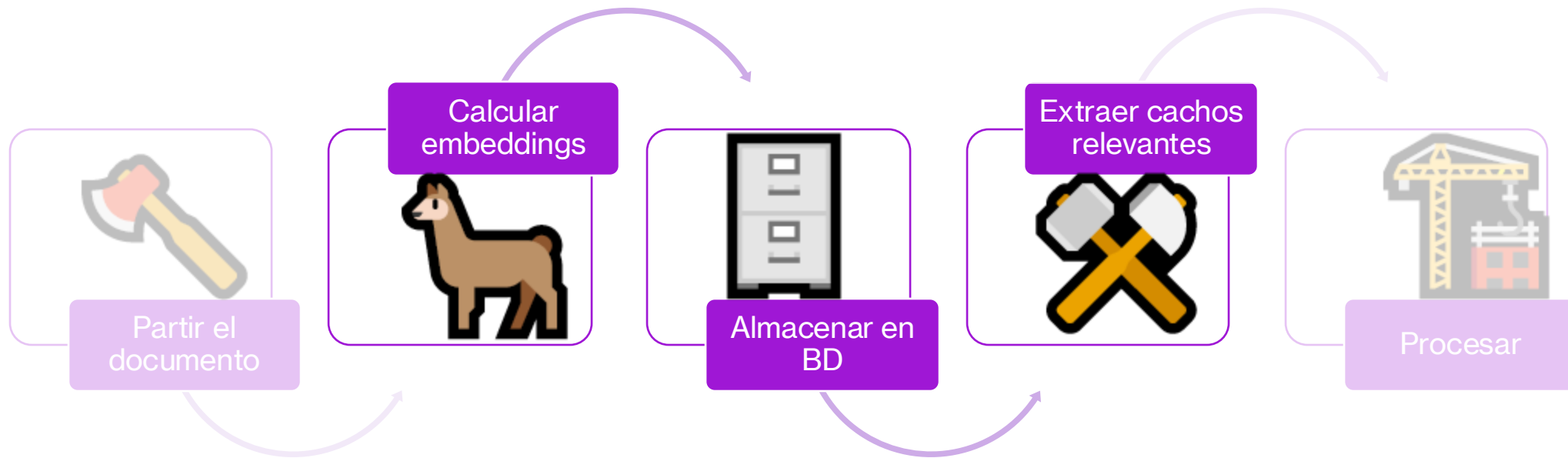
# Juntar los resultados: refine



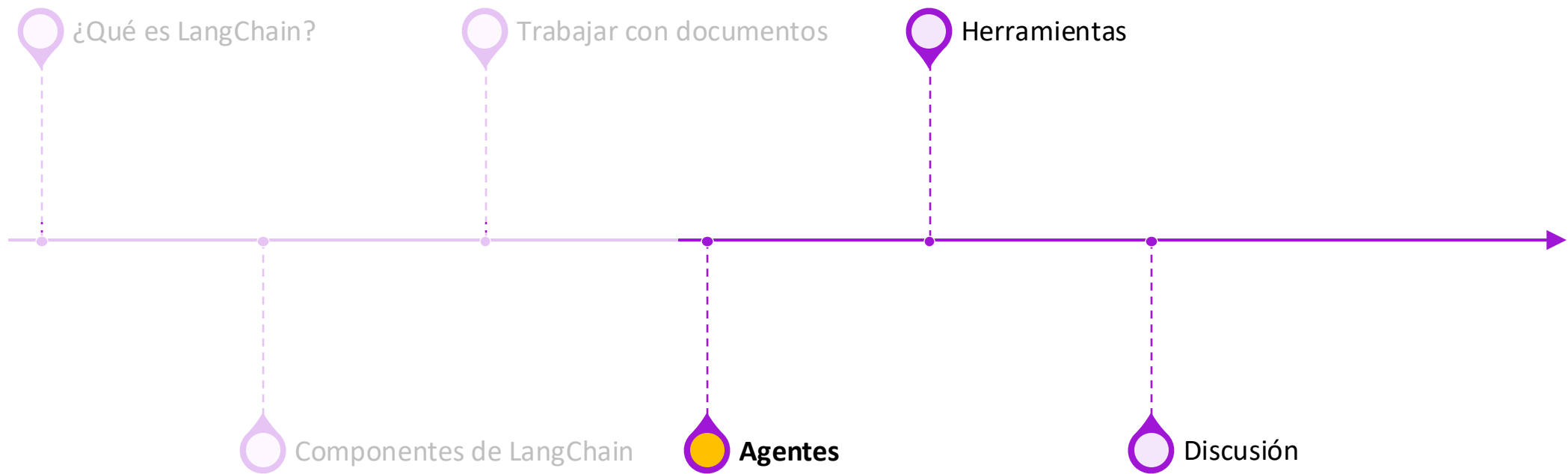
# Interactuar con un documento



# Procesar partes de un documento



# Indice





# LANGGRAPH

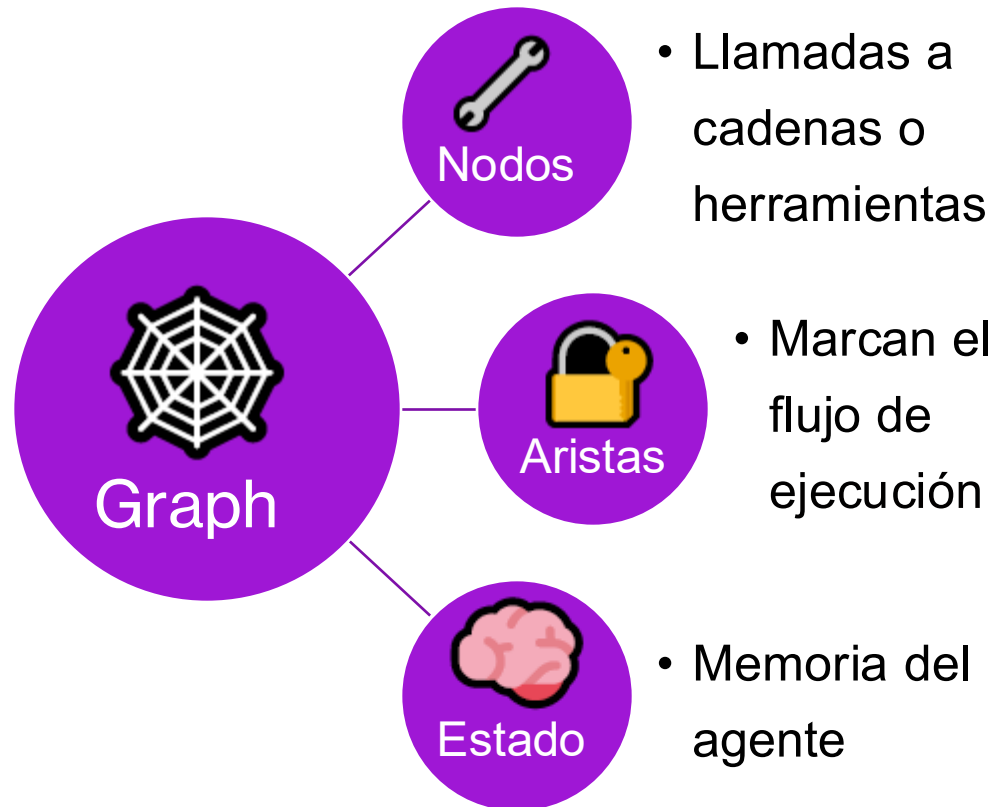
⚡ Construida encima de LangChain y permite crear **agentes inteligentes** facilmente.



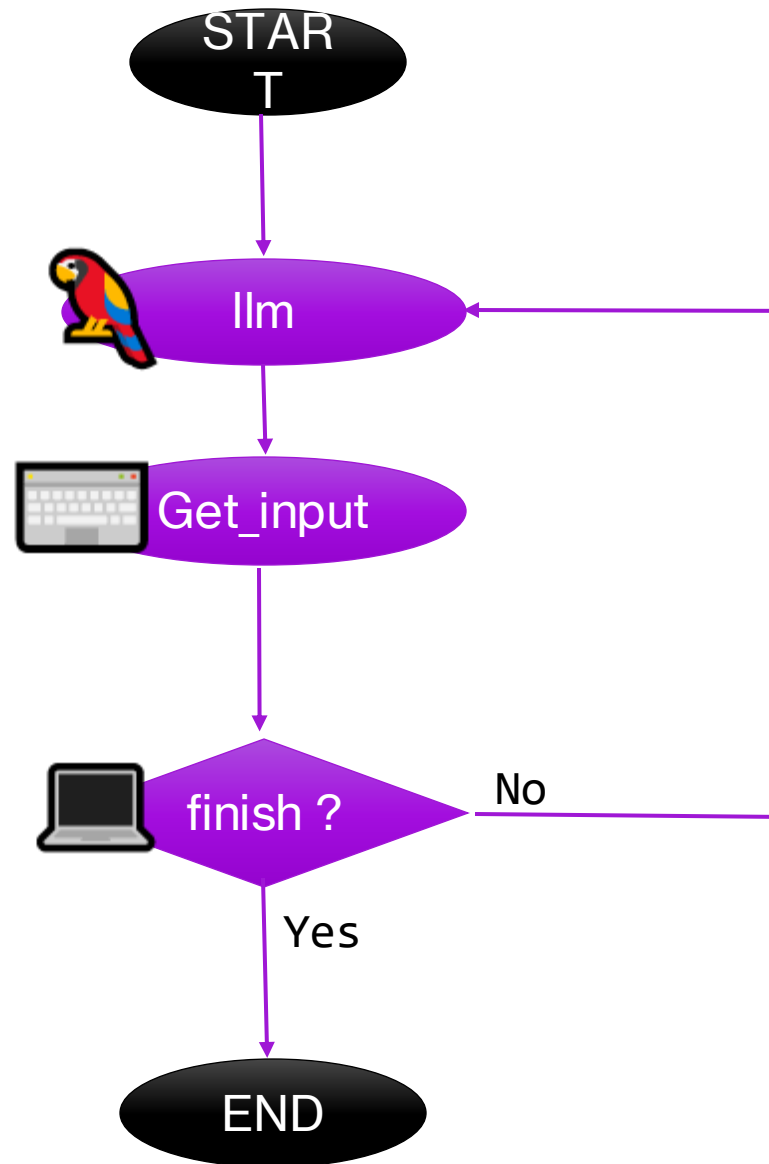
# ¿QUÉ ES UN AGENTE?

Los agentes utilizan un LLM para elegir la secuencia de acciones a realizar. Hasta el momento, las cadenas consisten en una secuencia de acciones programada “a fuego” en el código. En los agentes, se utiliza un modelo de lenguaje como un motor de razonamiento para determinar qué acciones tomar y en qué orden.

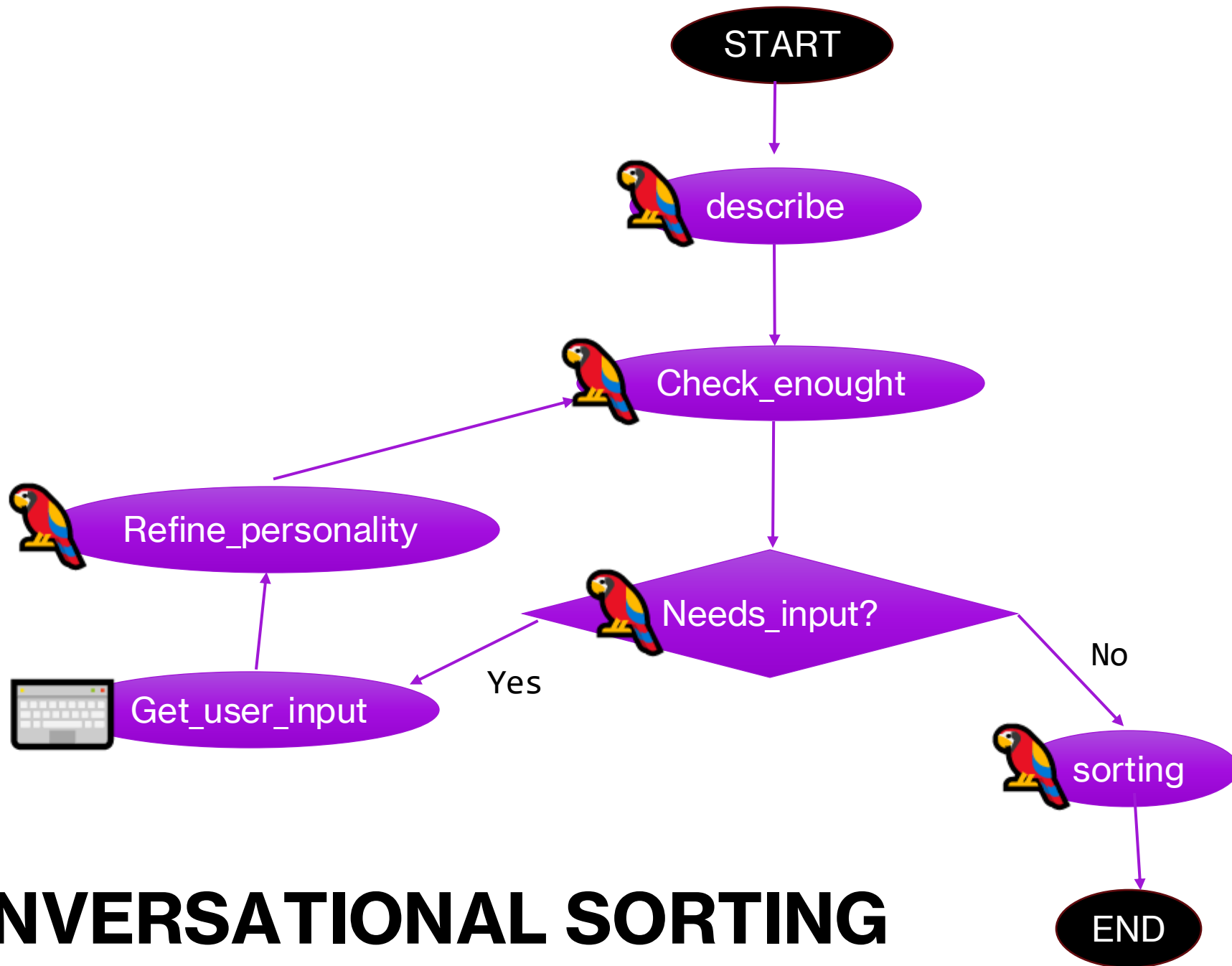
# Componentes básicos de LangGraph





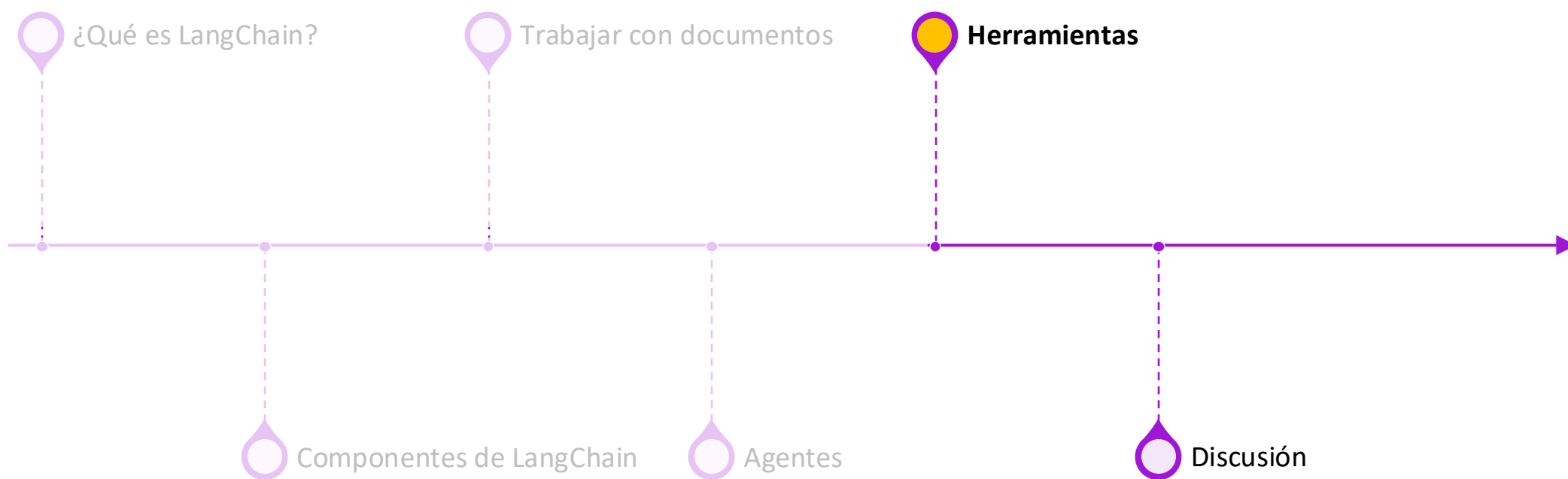


# Chat bot conversational



# CONVERSATIONAL SORTING

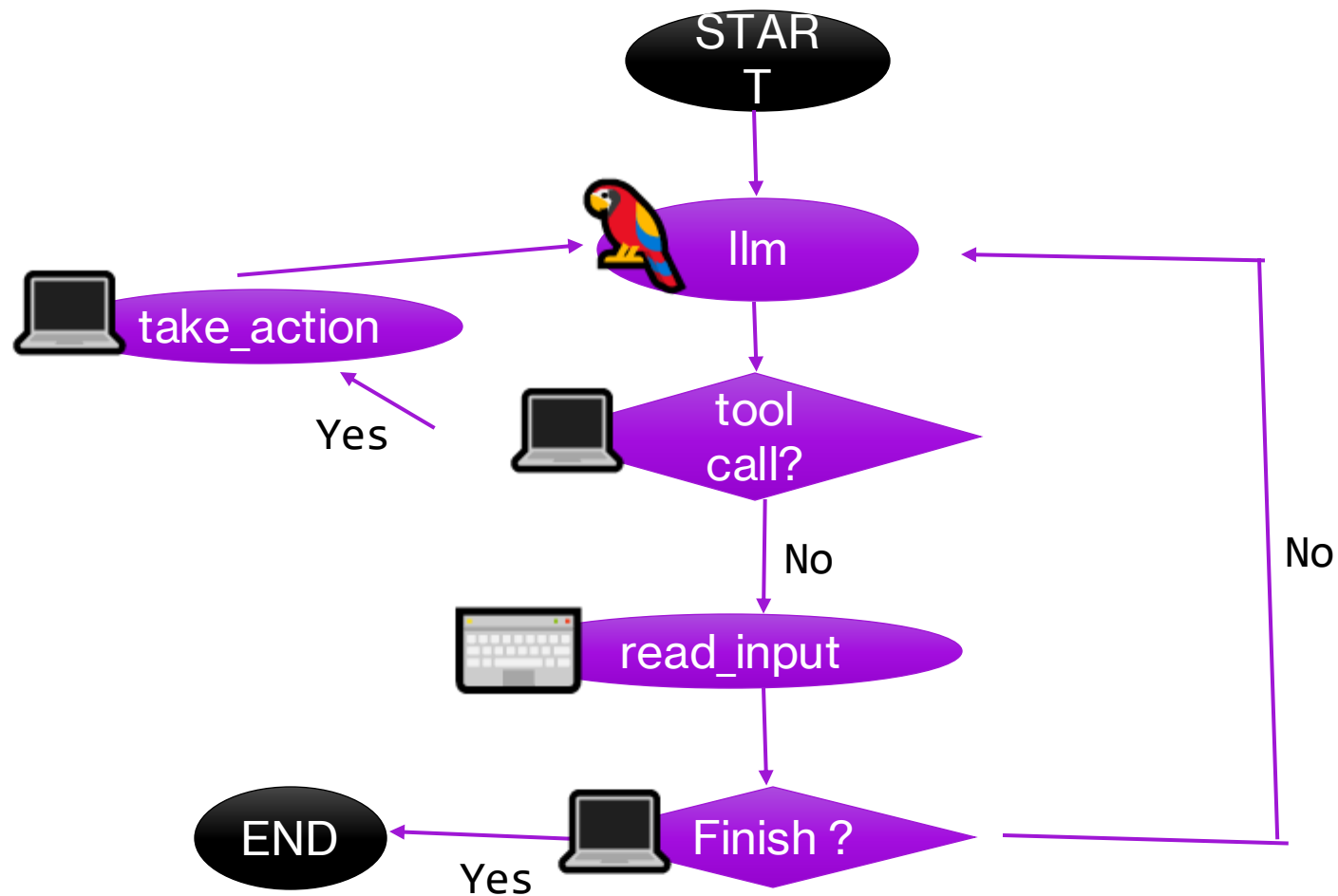
# Indice



# Herramientas

- Las herramientas son funciones que el LLM puede llamar.
- Permiten aumentar las capacidades de un LLM (buscar en internet, hacer cálculos ... etc).
- Es el LLM el encargado de decidir cuando tiene que usar una herramienta y cómo usarla.

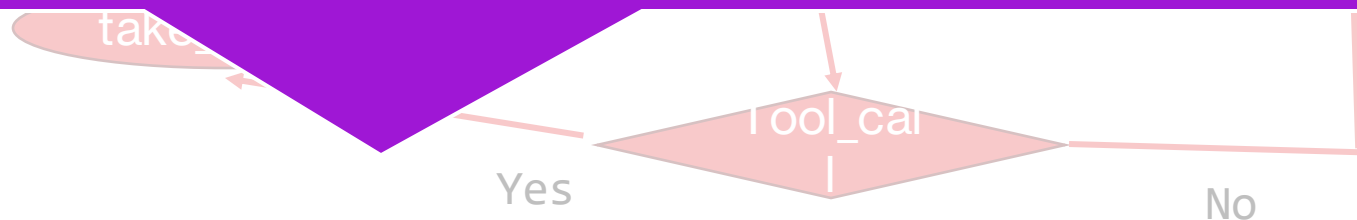




# Chat bot with tools

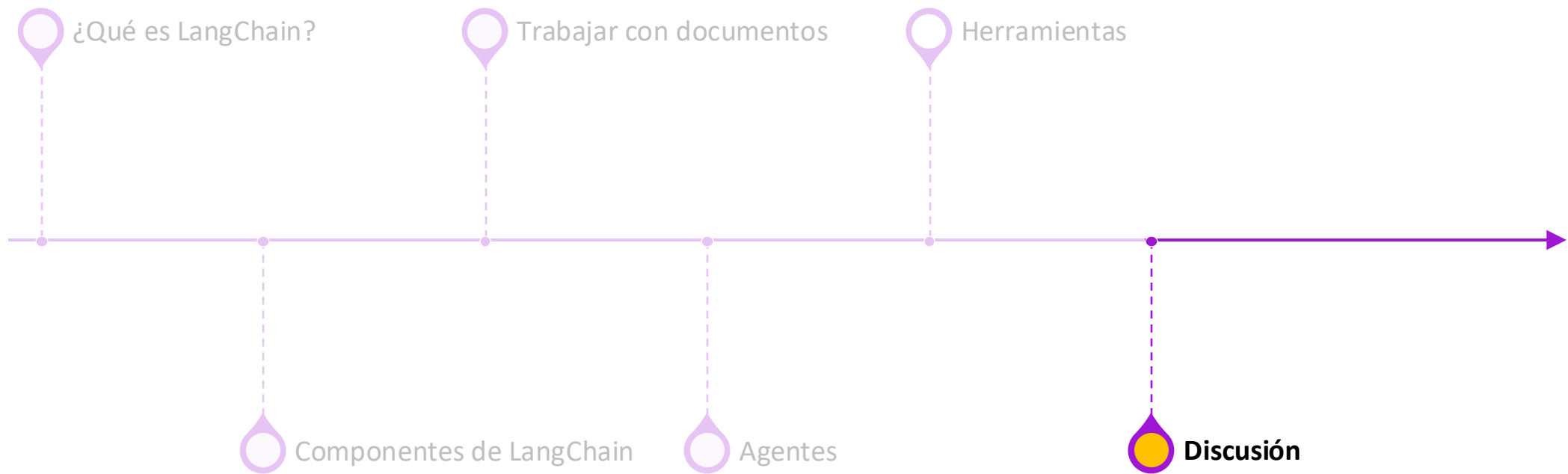


Solo los modelos más grandes 70b  
y 405b son capaces de usar  
herramientas.



# Chat bot with tools

# Indice





# Discusión

- El desarrollo de los LLMs avanza muy deprisa.
- Cada semana se presentan modelos nuevos más capaces y eficientes.
- Las librerías están en constante cambio.
- Los modelos han mejorado mucho en un año, pero aun siguen siendo difíciles de controlar y manejar.
- El área está creciendo tanto que esta charla ya se podría convertir en una asignatura.







# CODIGO DISPONIBLE

[https://github.com/panizolledotangel/LangChain\\_llama\\_talk](https://github.com/panizolledotangel/LangChain_llama_talk)

 angel.panizo@upm.es