

# Desmitificando a Maxia: Como Funciona Realmente um Modelo de Linguaxe Grande

Unha viaxe visual dende os datos brutos ata o diálogo intelixente.

# Que é un LLM? Un Motor de Patróns Lingüísticos.

Un Modelo de Linguaxe Grande (LLM) é unha intelixencia artificial deseñada para **entender, xerar e razoar con texto**. Non "pensa" no sentido humano; funciona aprendendo os patróns estatísticos da linguaxe a partir dunha cantidade masiva de textos para predicir a palabra máis probable que vén a continuación.

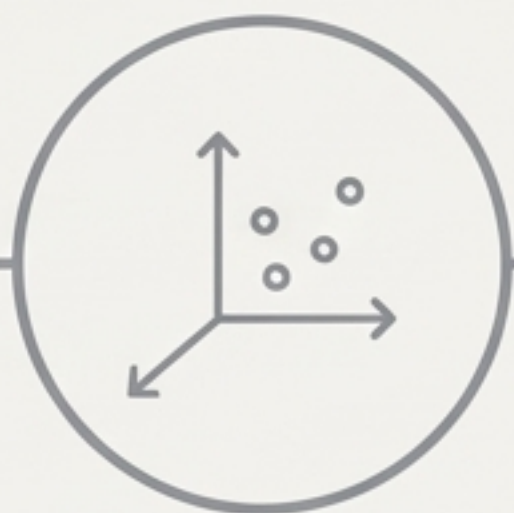


# A Anatomía dun LLM en Seis Pasos

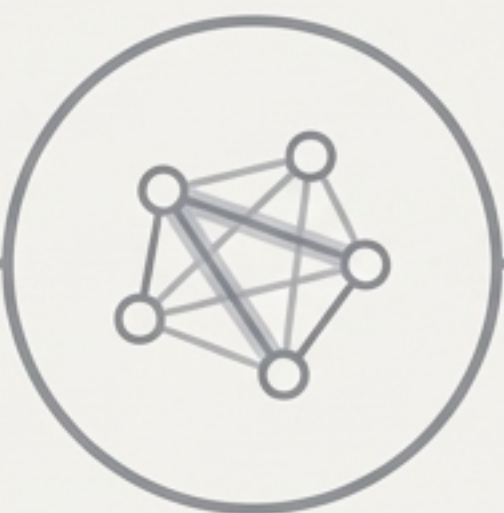
Para pasar de datos a diálogo, un LLM segue un proceso de construción e funcionamento ben definido. Exploraremos cada unha destas etapas.



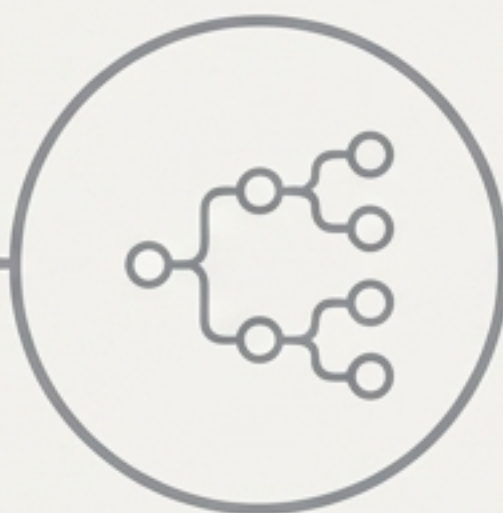
**1. Adestramento**



2. Vectores



3. Transformer



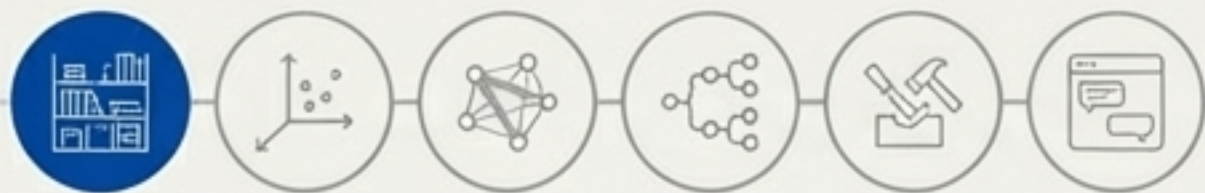
4. Predición



5. Axuste Fino



6. Uso Práctico



# 1. Adestramento: Aprender a Linguaxe do Mundo

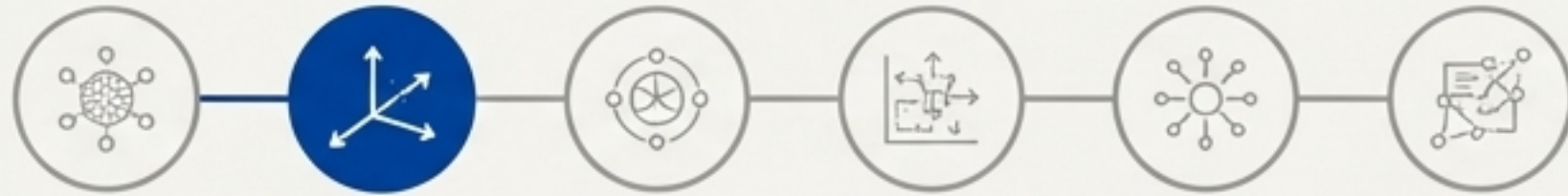
O modelo aliméntase con millóns de textos (libros, artigos, webs). O obxectivo non é memorizar, senón **aprender as relacións estatísticas e semánticas entre as palabras.**

O modelo aprende que a palabra "París" adoita aparecer preto de "Francia", ou que "comer" está intimamente relacionado con "alimentos".

*"É como se o modelo lese toda a biblioteca da humanidade e comezase a notar os patróns subxacentes no uso da linguaxe."*



## 2. Vectores: Traducindo Palabras a Coordenadas de Significado



### 2. Vectores

Cada palabra convértese nun conxunto de números chamado **vector**. Estes números non representan letras, senón a posición da palabra nun 'espazo de significado' multidimensional.

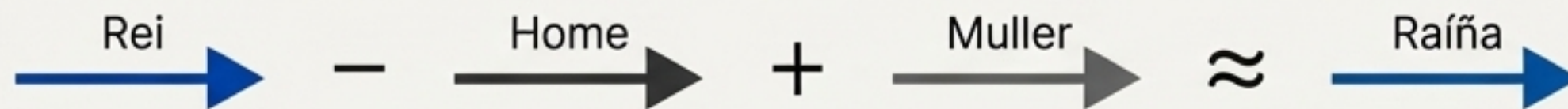
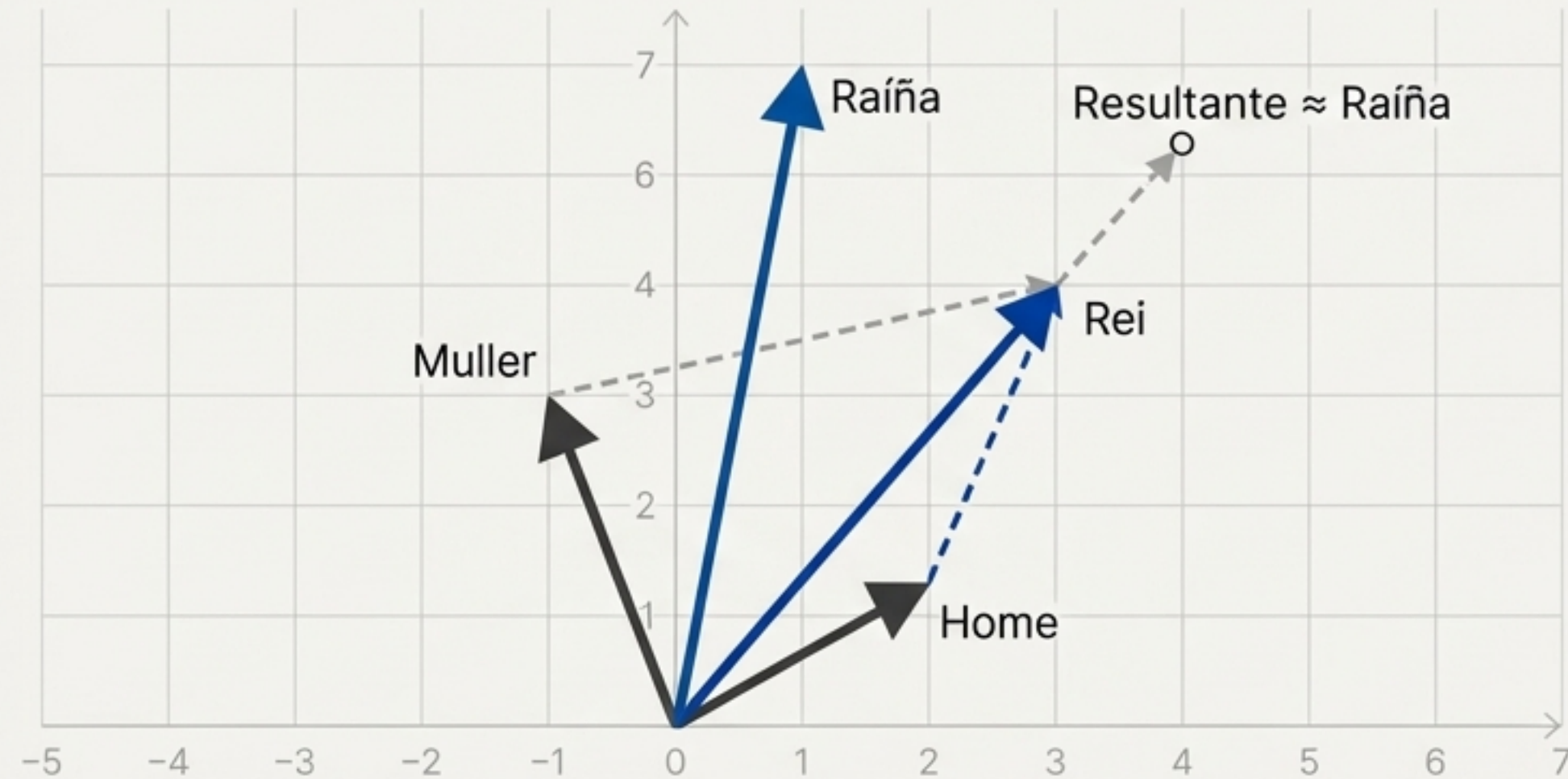
**Insight Clave:** A proximidade neste espazo matemático reflicte a proximidade semántica. O modelo pode calcular que os vectores de 'rei' e 'raíña' están moito máis preto entre si que os de 'rei' e 'avión'.



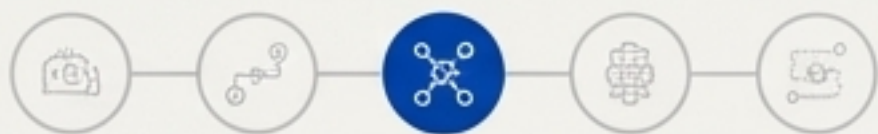
Design conceptual y da una beautiful ea intuitivo visualizado "star map of meaning".

# A Aritmética do Significado

A representación vectorial permite realizar operacións matemáticas que reflicten relacións do mundo real. Este é un dos fundamentos do razoamento analóxico nos LLM.



[Vector de "Rei"] - [Vector de "Home"] + [Vector de "Muller"] ≈ [Vector de "Raíña"]

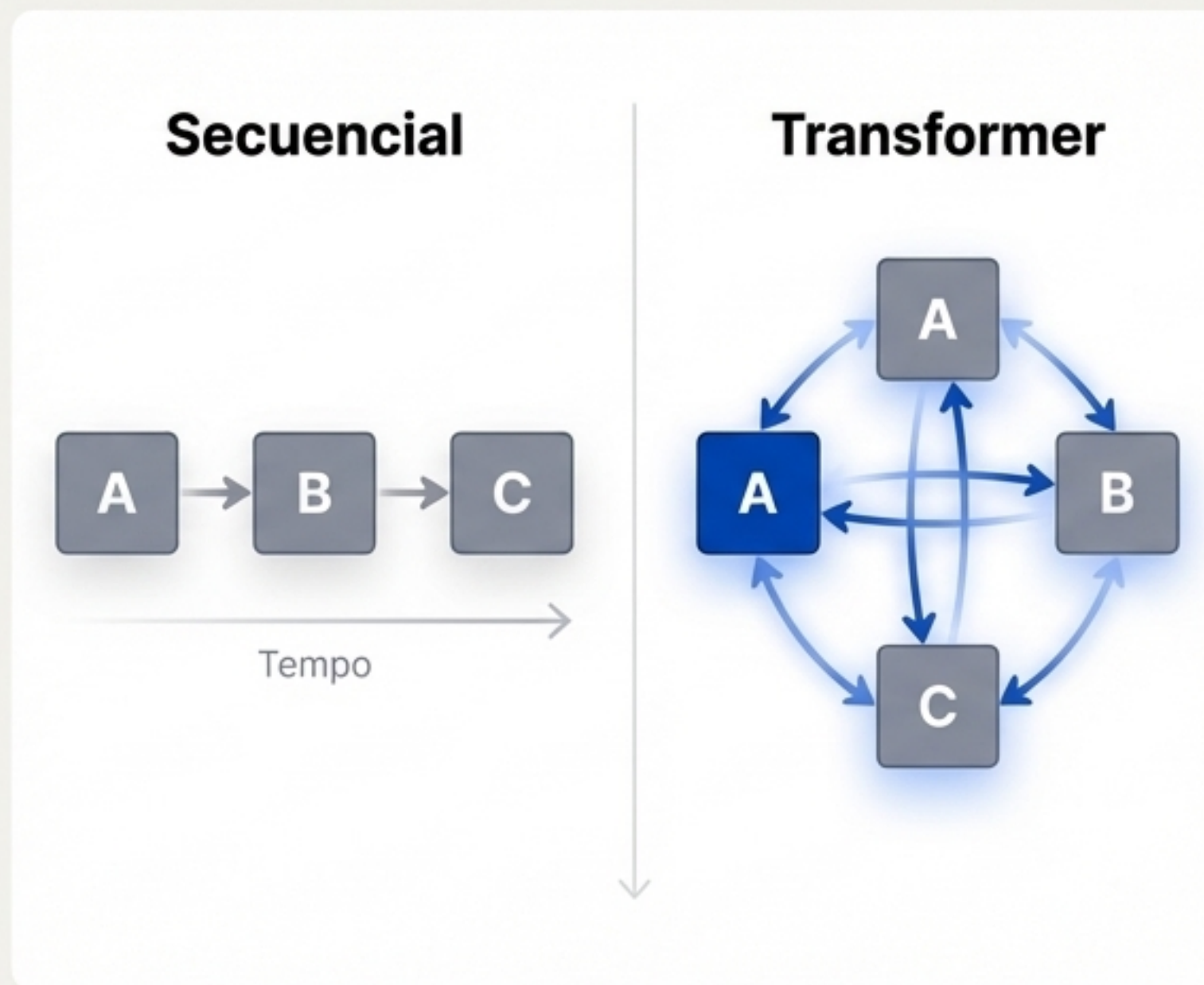


### 3. Transformer

## 3. A Revolución de 2017: A Arquitectura Transformer

Os LLM modernos (GPT, Gemini, Claude) usan unha arquitectura chamada **Transformer**. A súa innovación clave é a capacidade de procesar todas as palabras dun texto simultaneamente, non de forma secuencial.

- **Visión Global:** Entende o contexto completo da frase ou parágrafo.
- **Relacións a Distancia:** Conecta palabras relevantes aínda que estean afastadas.
- **Mecanismo de Atención:** Permítelle 'enfocarse' nas partes máis importantes do texto para unha tarefa específica.



# O Superpoder do Transformer: O Mecanismo de Atención

A "atención" permite ao modelo ponderar a importancia de diferentes palabras ao interpretar o significado doutra. Deste xeito, resolve ambigüidades e comprende estruturas gramaticais complexas.

O **can** que perseguía o gato **caeu** ao río



Atención: 'caeu' está  
fortemente conectado con 'can'



## 4. O Motor Principal: Predición da Seguinte Palabra

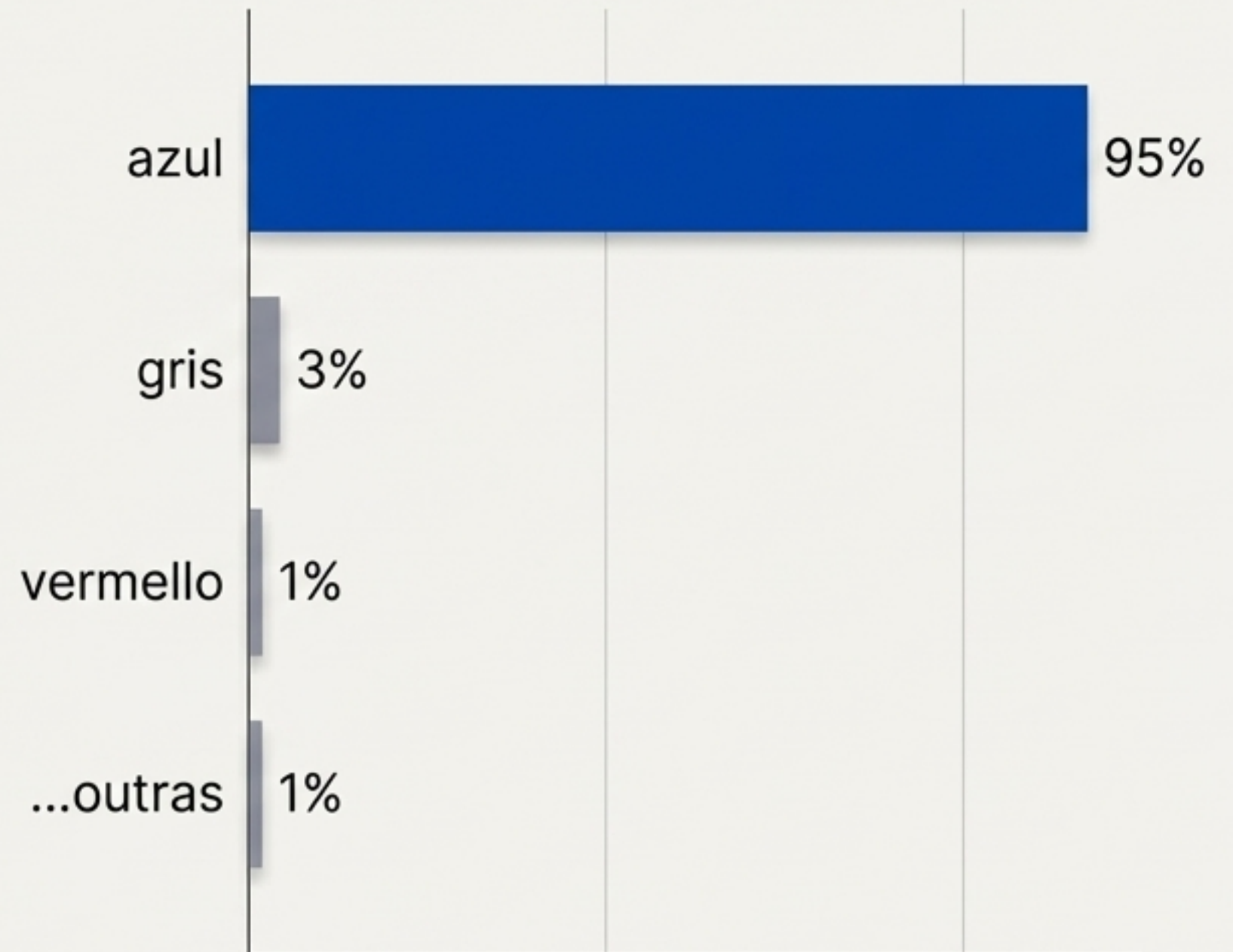
Unha vez adestrado, a tarefa fundamental dun LLM é sinxela: dado un anaco de texto, calcular as probabilidades de cal será a seguinte palabra.

### Exemplo Visualizado:

Input: O ceo é de cor...

Cálculo de Probabilidades do Modelo:

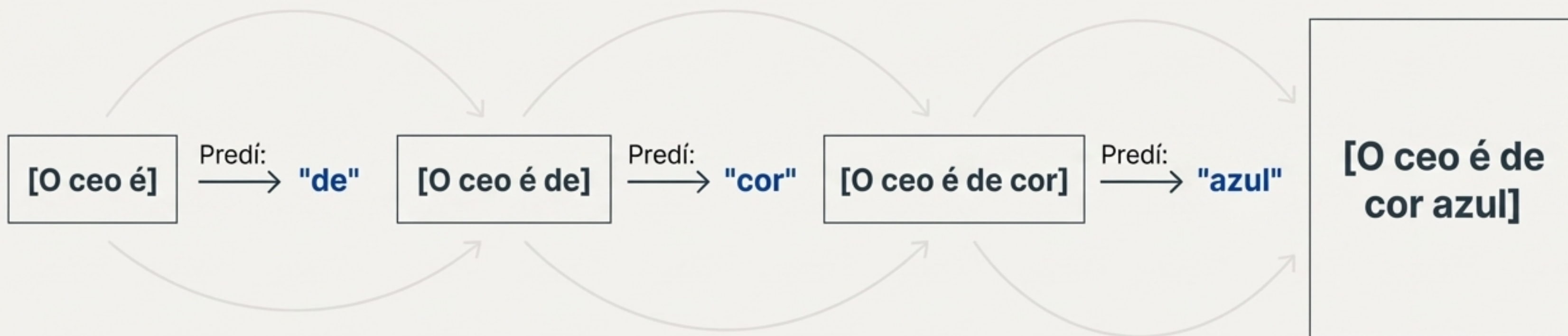
azul  
gris  
vermello  
...outras



# Construindo Respostas, Unha Peza de Cada Vez

O modelo non xera parágrafos enteiros dunha vez. Funciona nun bucle:

1. Predí a palabra máis probable.
2. Engade esa palabra ao texto de entrada.
3. Usa o novo texto como **contexto** para predicir a seguinte palabra.
4. **Repite** o proceso ata completar a resposta.

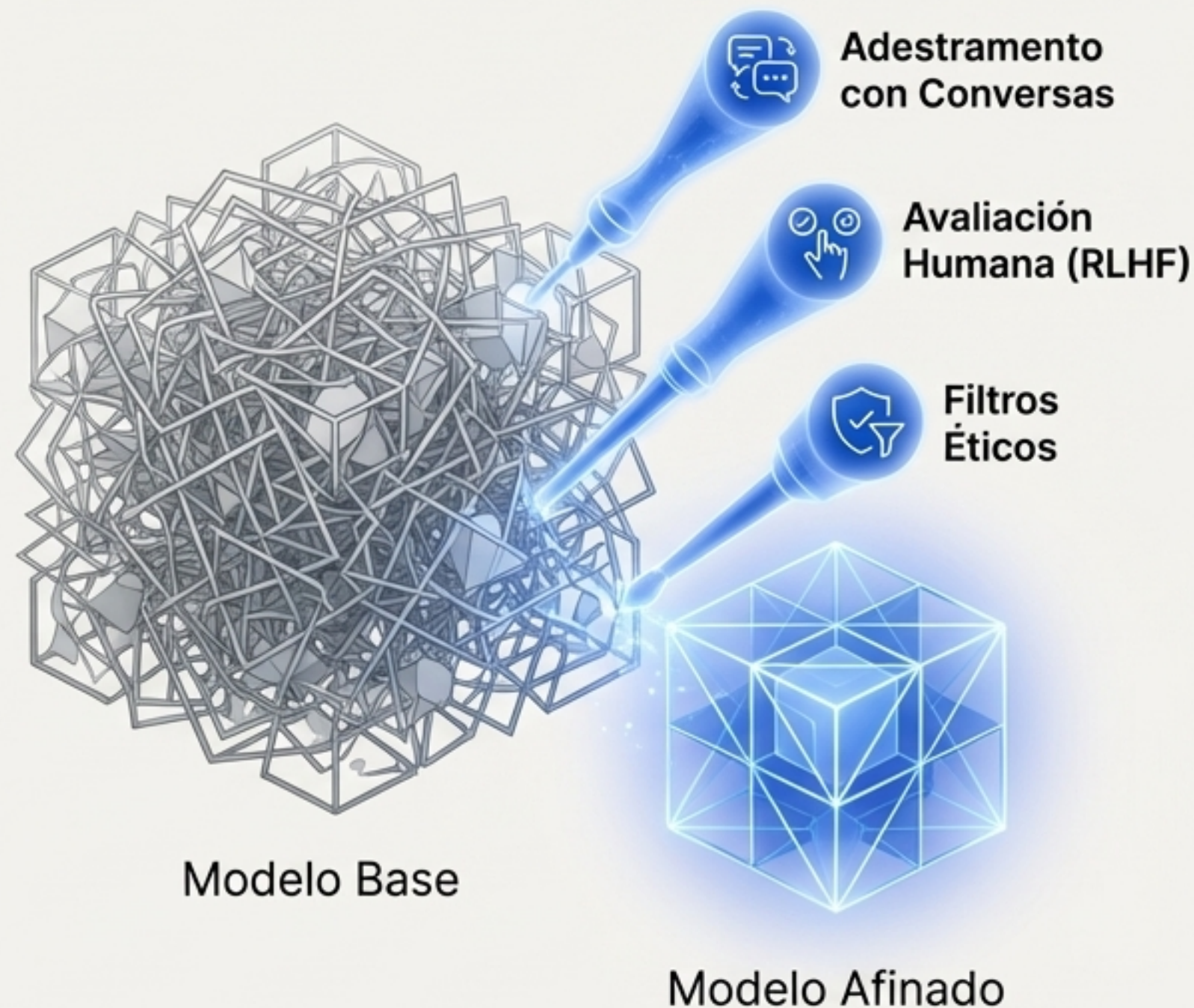




## 5. Axuste Fino: Do Coñecemento Bruto á Personalidade Útil

O adestramento inicial crea un motor de coñecemento xeral. O axuste fino é o proceso de esculpilo para que se comporte de maneira útil, segura e conversacional.

- **Adestramento con Conversas:** Ensínaselle a dialogar usando exemplos de conversas reais.
- **Avaliación Humana (RLHF):** Os humanos puntúan e corrixen as respostas para mellorar o ton, a precisión e evitar prexuizos.
- **Filtros Éticos:** Impléméntanse regras para evitar a xeración de contido perigoso ou inadecuado.

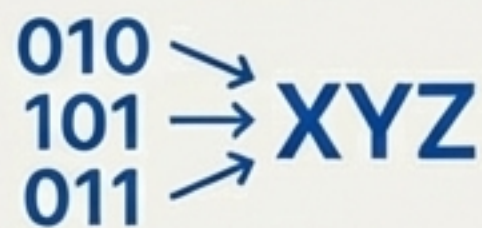




# 6. Na Práctica: O Ciclo Completo da Interacción

Cando lle escribes a un LLM, desencadeas todo este proceso en milisegundos.

**4. Saída de Texto:** Os números da resposta convértense de novo en palabras que podes ler.

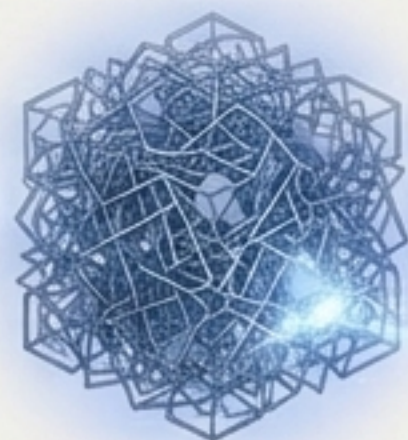


**1. Entrada do Usuario:** Escribes a túa pregunta.



**2. Vectorización:** O teu texto convértese en números (vectores).

**3. Procesamento e Xeración:** O modelo busca os patróns máis probables e comeza a xerar unha resposta palabra por palabra.

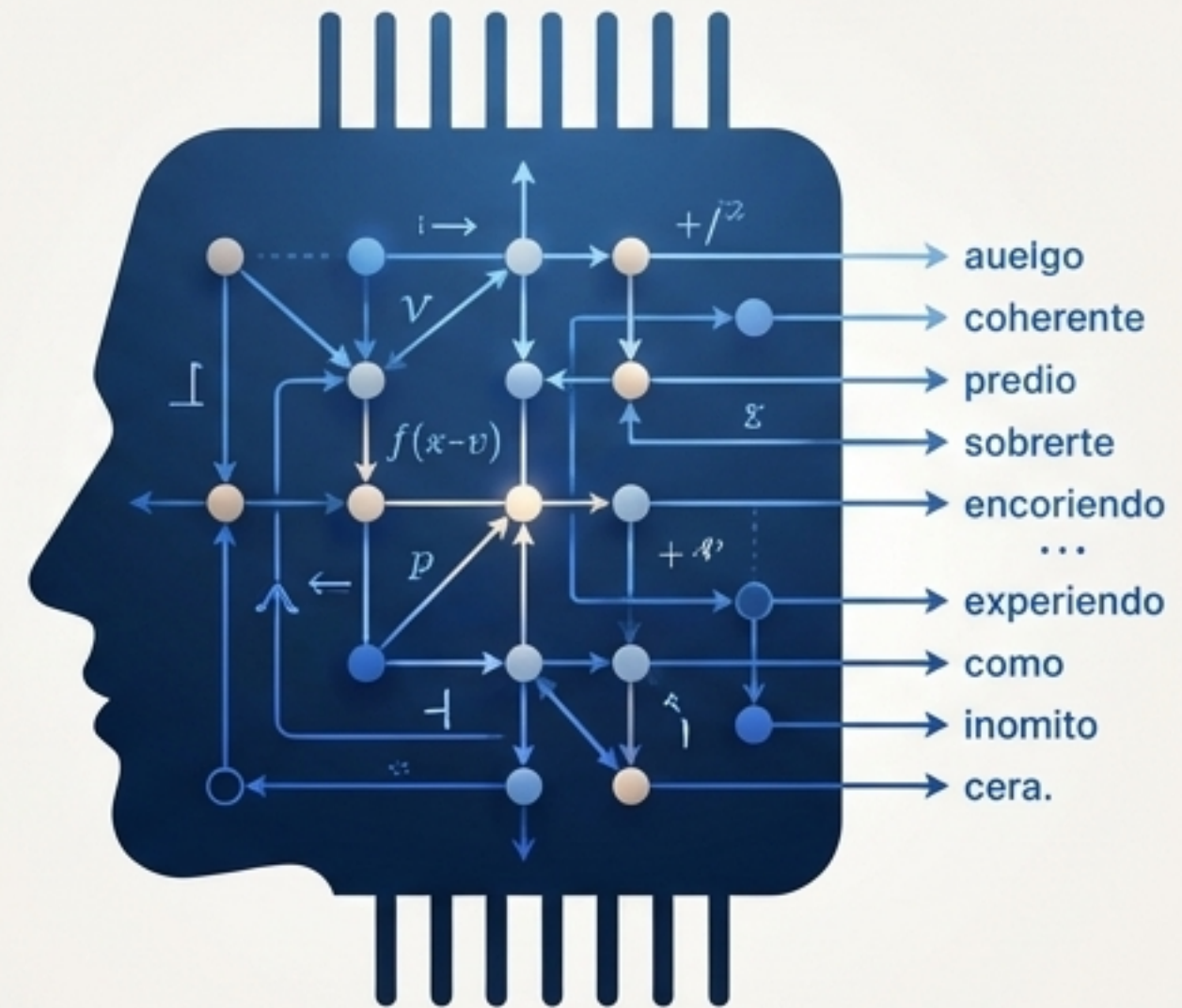


Non é Comprensión, é unha Predición Sofisticada

**Un LLM non pensa nin entende como unha persoa, pero é moi bo predicindo texto coherente e con sentido, porque aprendeu como usamos a lingua a unha escala sobrehumana.**



**PENSAMENTO HUMANO:**  
Complexidade biolóxica, emoción, contexto.



**PREDICCIÓN COMPUTACIONAL:**  
Probabilidade matemática, estrutura, escala.

# Da Probabilidade á Proeza: Capacidades Emerxentes

Esta complexa arquitectura de predición de patróns dá lugar a unha sorprendente gama de habilidades útiles:



Responder Preguntas



Escribir Historias



Resumir Textos



Traducir Idiomas



Programar Código



Manter Conversas

