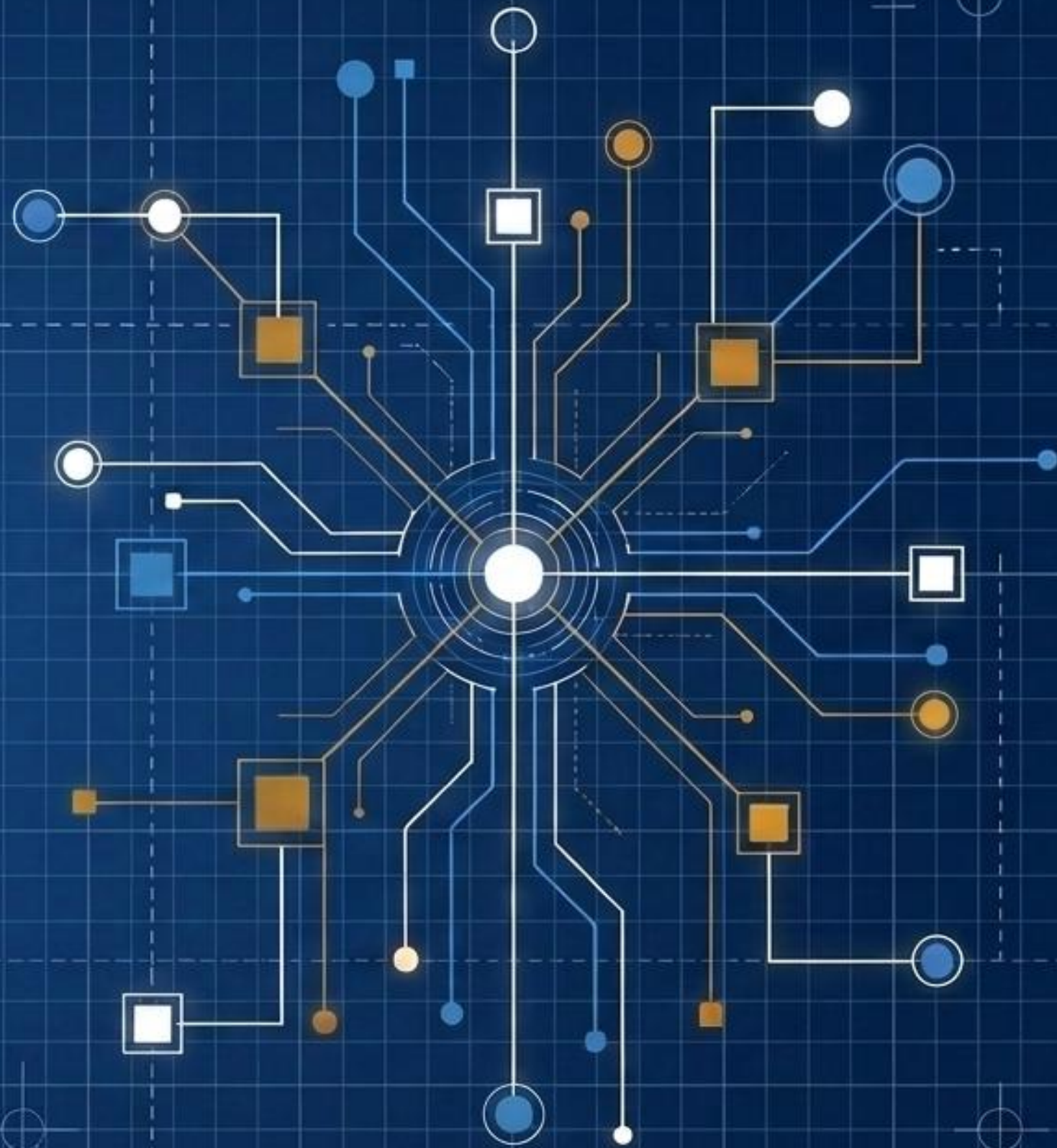


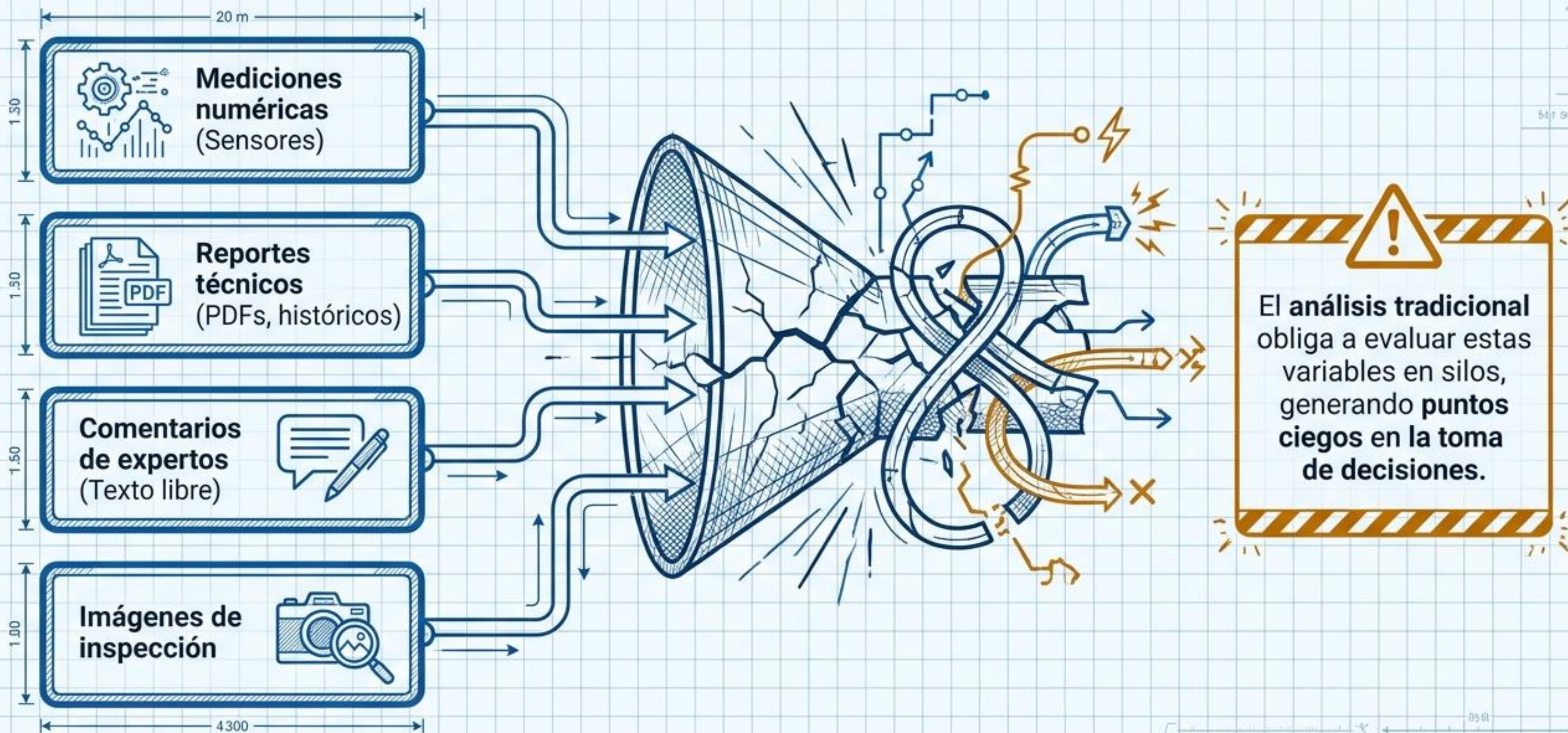
Jkey's Edu

De la Teoría a la **Producción:** LLMs y Agentes Autónomos en el Mundo Real

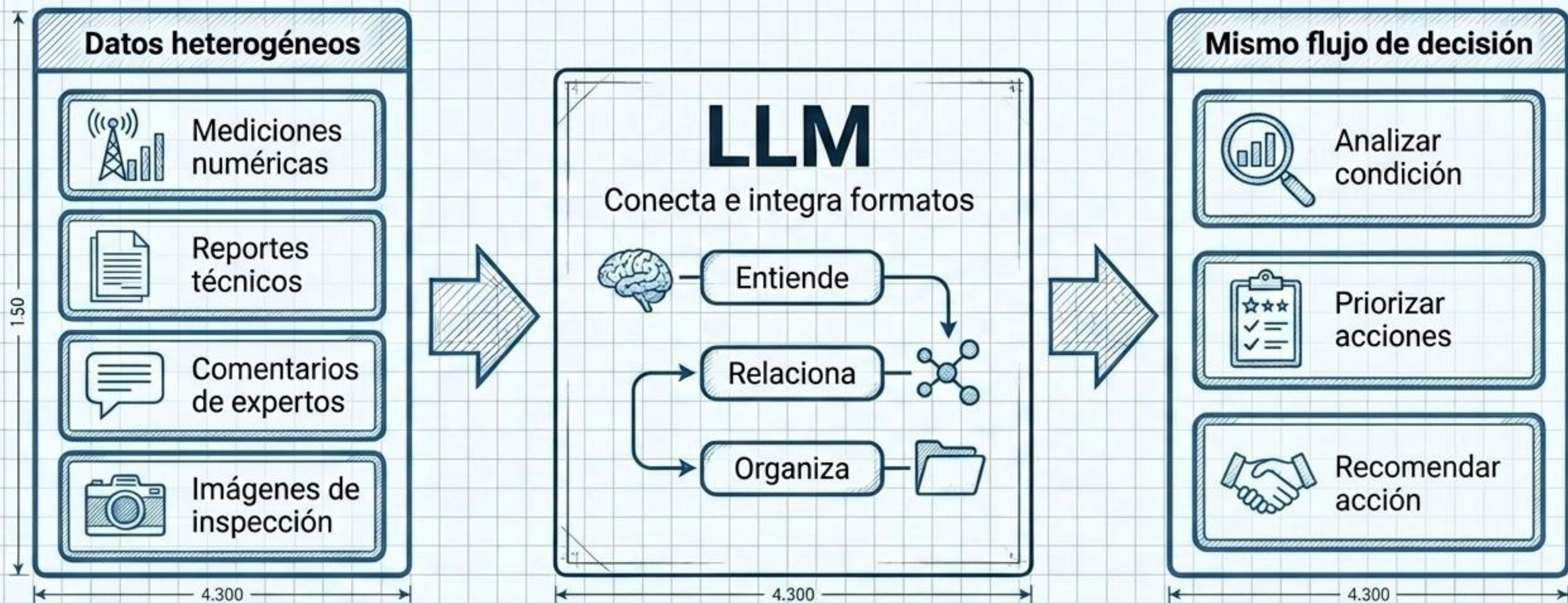
Fundamentos y Despliegue Seguro



En la industria, los datos hablan idiomas distintos



El modelo de lenguaje como motor de integración y decisión



El LLM permite unir múltiples tipos de datos en una decisión integrada.

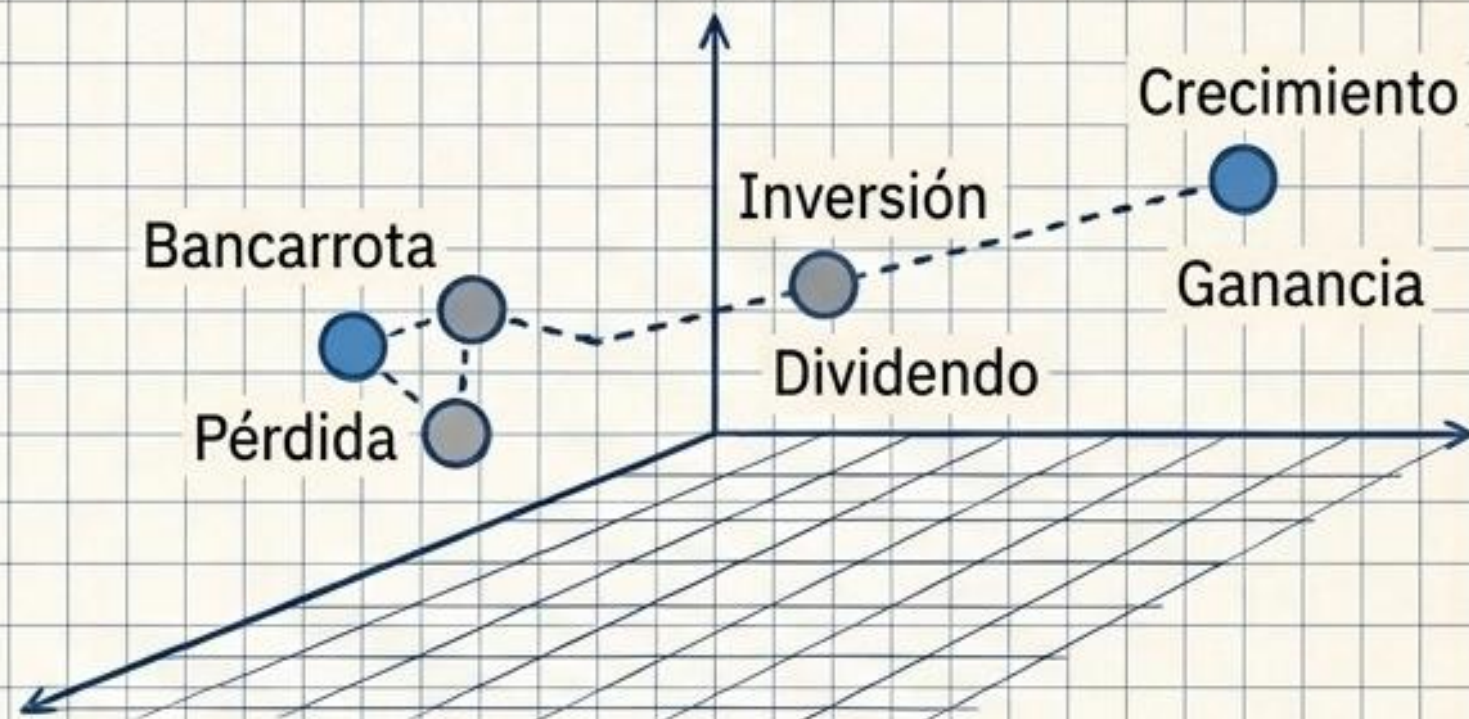
La IA no lee palabras, mapea relaciones financieras espaciales

TOKENS



El texto se divide en unidades procesables (tokens).

EMBEDDINGS



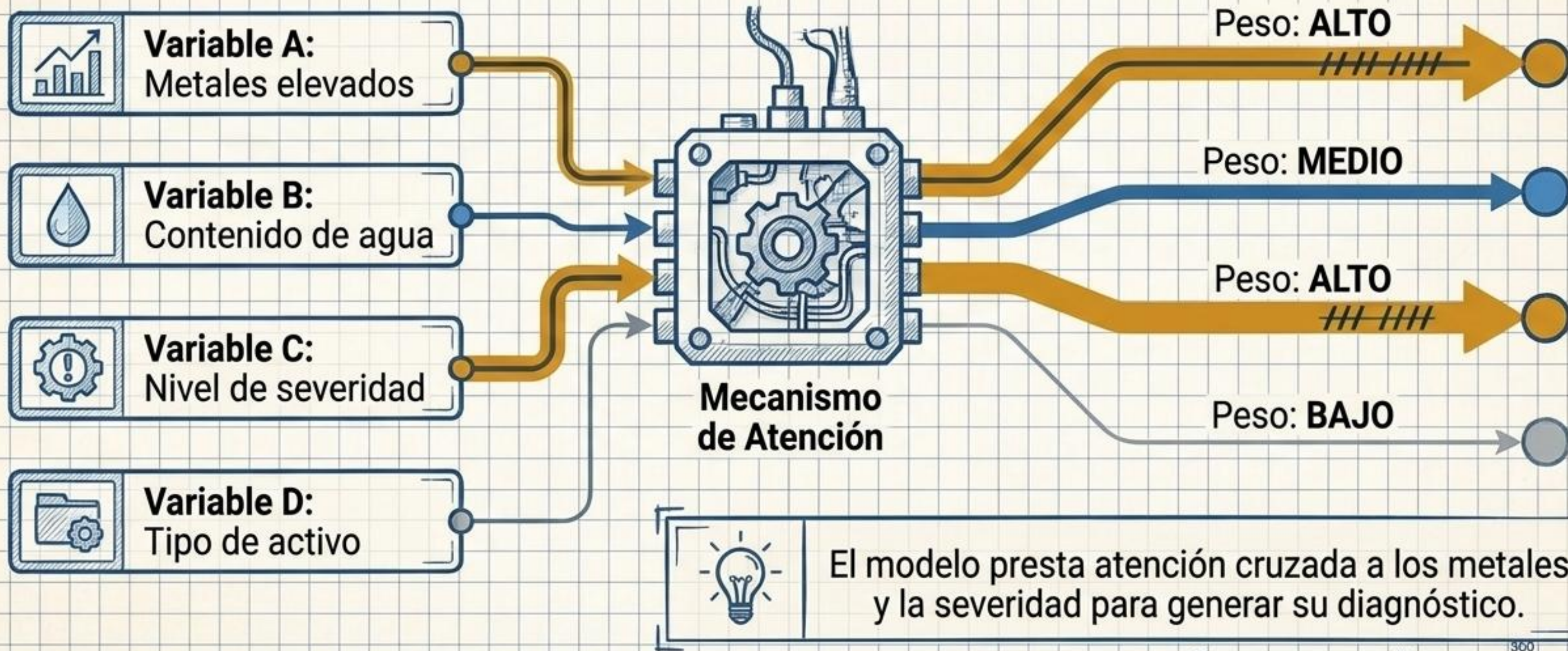
Mapeo espacial 3D de conceptos.



Cada token es un vector numérico. Conceptos relacionados financieramente se agrupan matemáticamente.

La Arquitectura Transformer: Asignando pesos de importancia

Permite que el modelo asigne distintos pesos a diferentes partes del contexto antes de responder.



La ventana de contexto es el límite estricto de memoria operativa



Regla de negocio: El LLM solo puede decidir sobre la información que cabe dentro de su ventana de atención inmediata. Cargar datos irrelevantes destruye la precisión.

💡 El LLM solo puede usar directamente la información que cabe dentro de su ventana de contexto.

El abismo de la implementación: Por qué los pilotos de IA mueren

De los proyectos de IA en
en organizaciones reales
no llegan a producción
(Fuentes: Gartner y McKinsey).

De los proyectos de IA en
en organizaciones reales
no llegan a producción
(Fuentes: Gartner y McKinsey).

70% - 85%

Diagnóstico

Quedan como prototipos,
demos o chats que nadie usa.

La causa raíz

Las empresas intentan usar la IA
como un simple oráculo de
preguntas (Chatbot) en lugar de
integrarla como un motor de
ejecución continua (Agente).

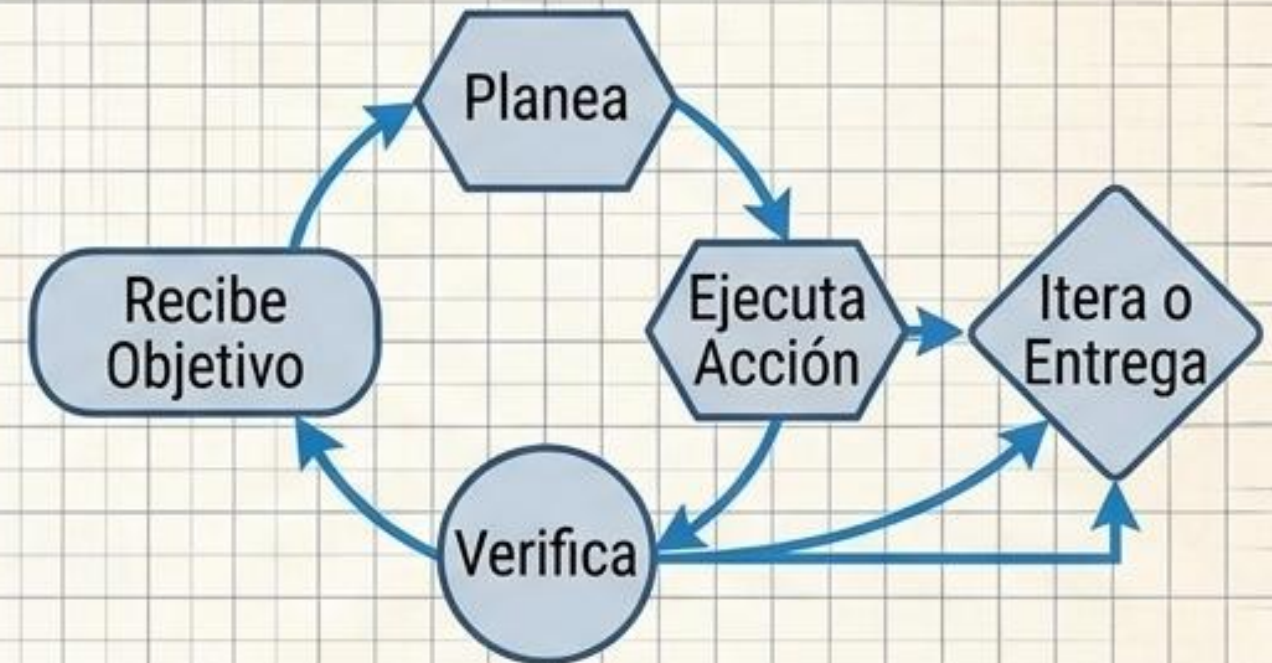
El salto arquitectónico: De conversar a operar

El Chatbot



- Depende del usuario para cada paso
- Sin estado
- Pasivo

El Agente de IA



- Ciclo de trabajo agéntico
- Capaz de usar herramientas
- Autónomo dentro de límites dados

El mensaje clave: La diferencia real no es que actúe solo, es la capacidad de iterar sobre errores y verificar sus propios resultados.

El estado base: El ADN de un agente en producción

Este archivo se carga antes de la primera interacción. Es lo que transforma un modelo genérico en un sistema experto.

AGENTS.md

Identidad

Quién es y cuál es su objetivo.

Contexto del Proyecto

Área, infraestructura disponible, qué base de datos usar.

Reglas de Operación

Restricciones críticas: Ej. Responder en español técnico, no inventar datos ausentes, señalar desviaciones con datos exactos.

MD (Markdown):

Formato de texto simple y ligero para escribir instrucciones estructuradas.

El 80% del valor de un agente proviene de este archivo
archivo de instrucciones, no del modelo base.

El commodity de los modelos vs. El valor de la orquestación

Arquitectura, Contexto, Hooks
y Reglas Organizacionales

El Diferenciador Competitivo
(El Activo Real)

Los modelos de lenguaje
son materia prima estadística.
No tienen contexto
organizacional ni
experiencia de planta.

Infraestructura Accesible
(Claude, GPT, Gemini)

La verdadera ventaja competitiva en 2026 no es tener acceso a una IA más potente, sino dominar la ingeniería de sistemas que convierte un modelo crudo en un agente estructurado y seguro.

El factor inquebrantable: El criterio humano

“ Lo que ningún algoritmo puede quitar: la intuición ética, la experiencia que se vive en el cuerpo y el juicio de saber qué automatizar. ”

Un modelo puede procesar millones de tokens y describir un fallo mecánico perfectamente, pero es el humano quien define las reglas, los límites éticos y asume la **responsabilidad** del impacto en el mundo físico.

Las personas construyen y dirigen a los agentes, no al revés.