



Rodrigo Liberoff
Arquitecto en ENCAMILA



Celeste Sánchez
Arquitecta en Infoavan

La liga de las IAs extraordinarias

Evento Virtual: 19 de Noviembre – 18:00 CET - www.azurecloudlab.net



Grupo Azure Cloud Lab



@azurecloudlab



THE LEAGUE OF EXTRAORDINARY ARTIFICIAL INTELLIGENCES



La liga de las IAs
extraordinarias

Nuestro objetivo



Seleccionar tres Inteligencias Artificiales que destaqueen y sean capaces de cumplir importantes misiones

Traerlas a nuestra "base" para estudiar sus habilidades

Probarlas y, si son aptas, incorporarlas como miembros oficiales de la liga de las IAs extraordinarias

Nuestros agentes



El señor Llama

Origen

EEUU

Afiliación

META AI

Alias

Llama 3.3, Llama 3.2,
Llama 3.1, Llama 3,
Llama Guard 2, Code
Llama FB, Llama
Guard 1, Code Llama,
Llama 2

Super poderes

LLM completo que destaca en tareas como matemáticas, traducción y razonamiento lógico.

Soporta 8 lenguajes, incluyendo español e inglés. Herramientas relacionadas con la seguridad de la respuesta (Llama Guard y Code Shield, CyberSecEval 2)

Optimizado para GPUs comunes, reduciendo necesidades de hardware

Nuestros agentes



Lady Mistral

Origen

EUROPA

Afiliación

MISTRAL AI

Alias

Mistral Large, Mistral Small, Mistral Nemo, Pixtral Large, Codestral, Minimistral, Mistral Embed, Mistral Moderation

Super poderes

LLM completo y competitivo (a nivel económico).

Soporte multilenguaje. Capacidades mejoradas en razonamiento matemático y resolución de problemas complejos.

Soporte nativo para llamadas de funciones y servicios externos.



Nuestros agentes



El increíble Phi-3

Origen

EEUU

Afiliación

MICROSOFT

Alias

Phi-3 mini, Phi-3
small, Phi-3 médium,
Phi-3.5 MOE, Phi-3.5
visión

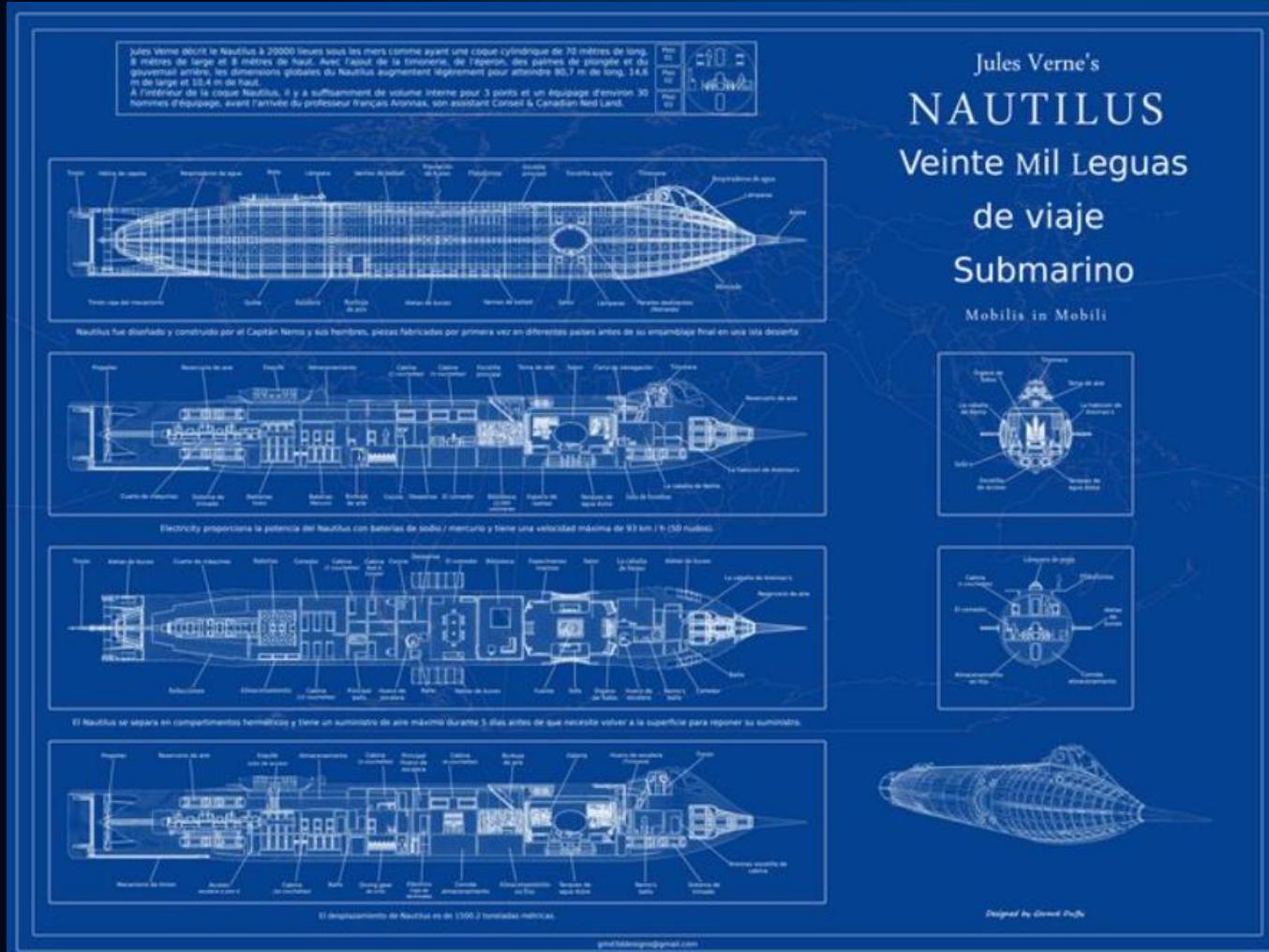
Super poderes

Es un SLM, es decir, un modelo pequeño y ligero capaz de operar en dispositivos con características limitadas.

Altamente escalable según demanda sin preocuparse de limitaciones de hardware.

Muy económico.

La Base



HARDWARE (sea en la nube o en local)
Puntos a tener en cuenta:

- CPUs / GPUs
- Memoria

Runtime para ejecutar los modelos
• Ollama



El plan

Crear una gran base (Nautilus)

- Desplegar Nautilus (utilizando AKS)
- Poner a punto los motores del nautilus (controlador de NVIDIA para gestionar la GPU)
- Preparar el soporte vital y los camarotes para los nuevos miembros de la tripulación (Desplegar Ollama)

Traer a los invitados

- Descarga de los modelos

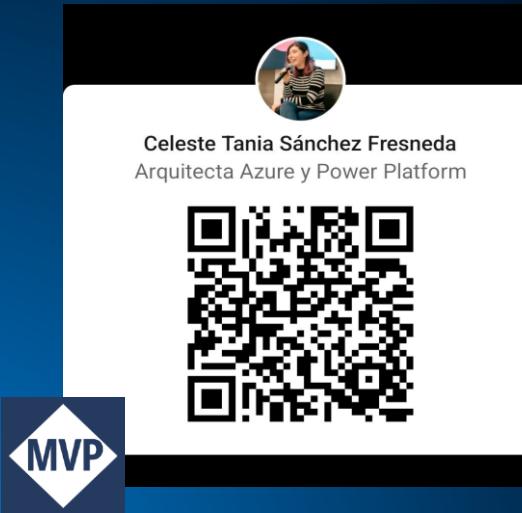
Es hora de actuar

- Ejecutar la aplicación para interactuar con cada modelo (creada con semantik kernel)

Manos a la obra



DEMO



<https://github.com/rliberoff/2025-azure-cloud-lab>

Muchas Gracias!!



Grupo Azure Cloud Lab



@azurecloudlab