



28 de junio de 2025
Bogotá, Colombia



“La automatización está impulsando el declive de las tareas banales y repetitivas

Amber Rudd

Secretario de Estado para el Trabajo y las Pensiones en 2019

Lo que pasa en algún lugar del mundo...

I'm excited to announce that the first draft of Data Engineering AI Agents is nearly complete and will be released soon!

★ This marks **HOTTECHSTACK** first major open-source project.

- Data Conversation Agents: Develop a suite of data-focused agents designed to collaboratively address specific challenges.
- AI Data Pipelines: Automatically generate code to build efficient Data Engineering Pipelines.
- AI Data Infrastructure: Create an end-to-end system for AI-driven code generation and execution within the infrastructure, delivering concrete results.

Reinvent Your Data Engineering Workflow
Automate complex data engineering tasks with GenAI agents that require minimal human assistance*

Data Conversation Agents
Collaborate with multiple specialized AI agents to solve complex data engineering tasks through natural conversation
[Get Started →](#)

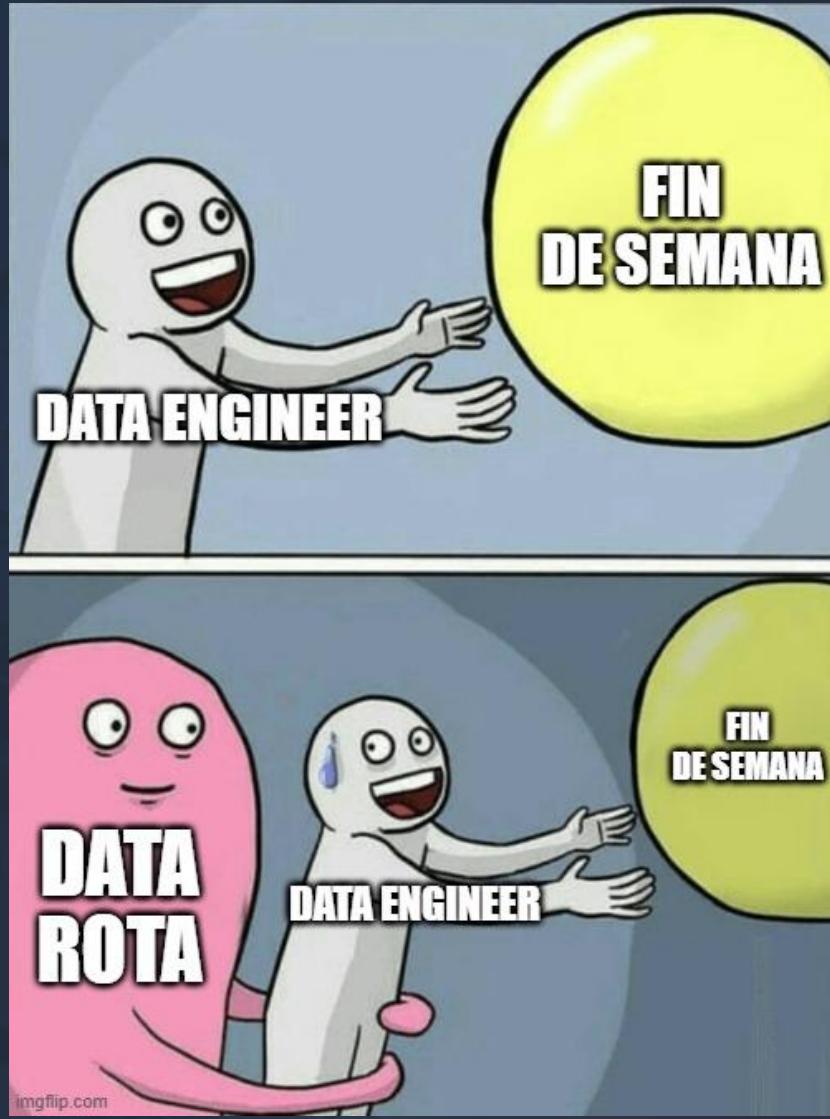
AI Data Pipelines
Generate optimized code for ETL/ELT pipelines with intelligent agents that understand your data architecture
[Get Started →](#)

AI Data Infrastructure
Build and deploy scalable data infrastructure with AI-generated configurations and implementation code
[Get Started →](#)

Specialized Data Engineering Strategies
Choose from various approaches tailored to your specific data challenges



... y una anécdota profesional (recurrente)



Agregando agentes a tu data pipeline: robusteciendo el proceso

Carlos Fernando Chicata

Data & Cloud solution engineer
TRD



Agenda

- 1) Contexto de un data pipeline
- 1) Entendamos los incidentes de un data pipeline
- 1) AI y Agentes: construcción
- 1) Lecciones de todo esto



Entendamos los data pipeline



AWS S3
Almacenamiento



AWS Glue
Procesamiento



AWS Step Function
Orquestamiento



AWS Lambda
Procesamiento



AWS Event Bridge
Gatillador y comunicación



Entendamos los data pipeline

Mapeamiento de datos Normalización
Transformación de datos
Gestión de datos Calidad de datos
Gobierno de datos datos
Integración de datos Perfil de datos
modelo de datos Carga de datos



Entendamos los data pipeline

Generalización de un data pipeline



Entendamos los data pipeline

Generalización de un data pipeline





**¡Gestión ante el
cambio inesperado!**



Algunas ejemplos de incidentes



Almacenamiento

- Cambios en los modelo de datos
- Permisos de acceso



Procesamiento

- Acceso a los datos
- Transformaciones
- Infraestructura y dependencias
- (Parametrización y modularidad)?



Diseño y definición

- Logica de negocio
- Gobierno
- organizacional
- Solicitudes de información



Parece fácil pero es “riesgoso” en realidad

Caso: “ajuste de esquema de tabla”

Tipo	Actividad	Automatización	Riesgos
Definición y diseño	Qué datos utilizar	Indicar qué columnas deben utilizarse ahora.	No hay disponibilidad de datos de negocio
Almacenamiento (Inicio)	Acceso de datos	Ajustar columnas a utilizar para procesamiento	Possible interrupción del proceso
Procesamiento	Transformar los datos	Gestionar el tipo de dato	Dar datos no valiosos para el negocio



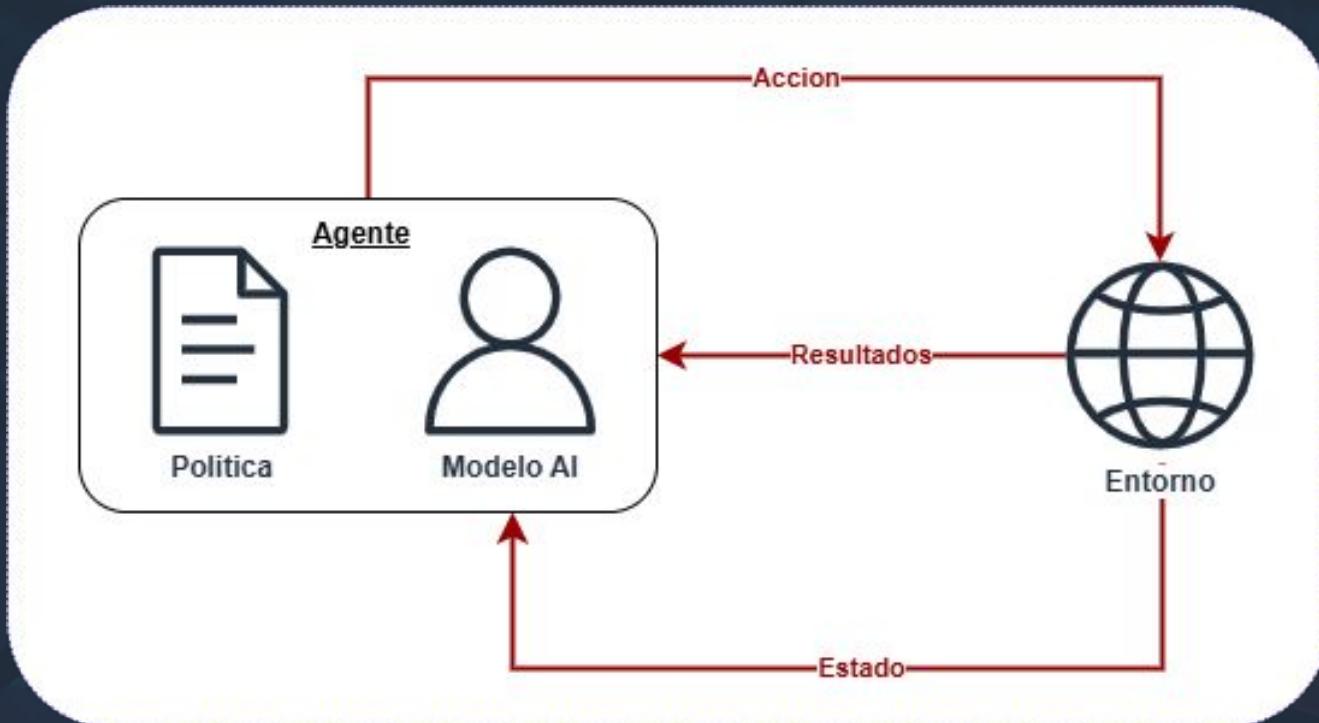
El conocimiento es
corazón en la
automatización



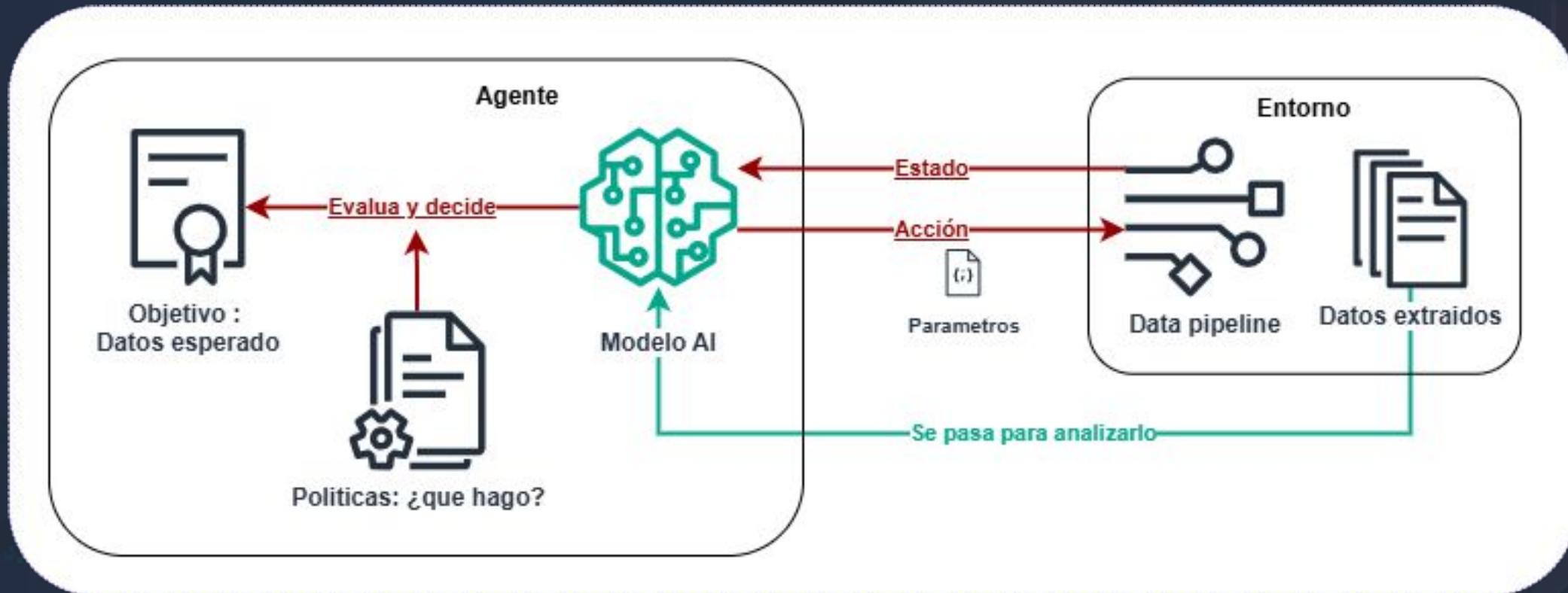
**La gestión de riesgos
por toma de decisión es
corazón en la
automatización**



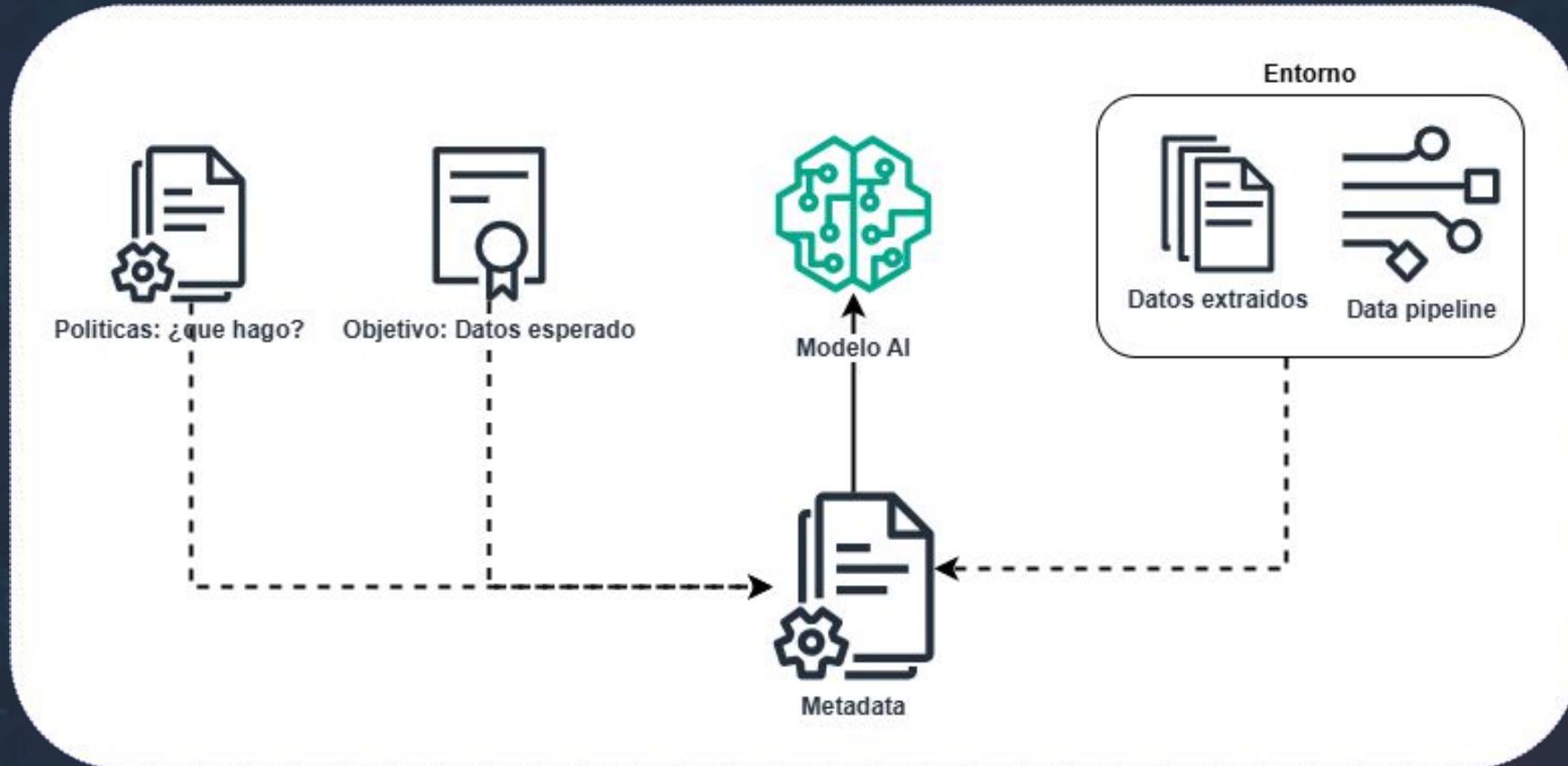
Agentes & AI: el fundamento



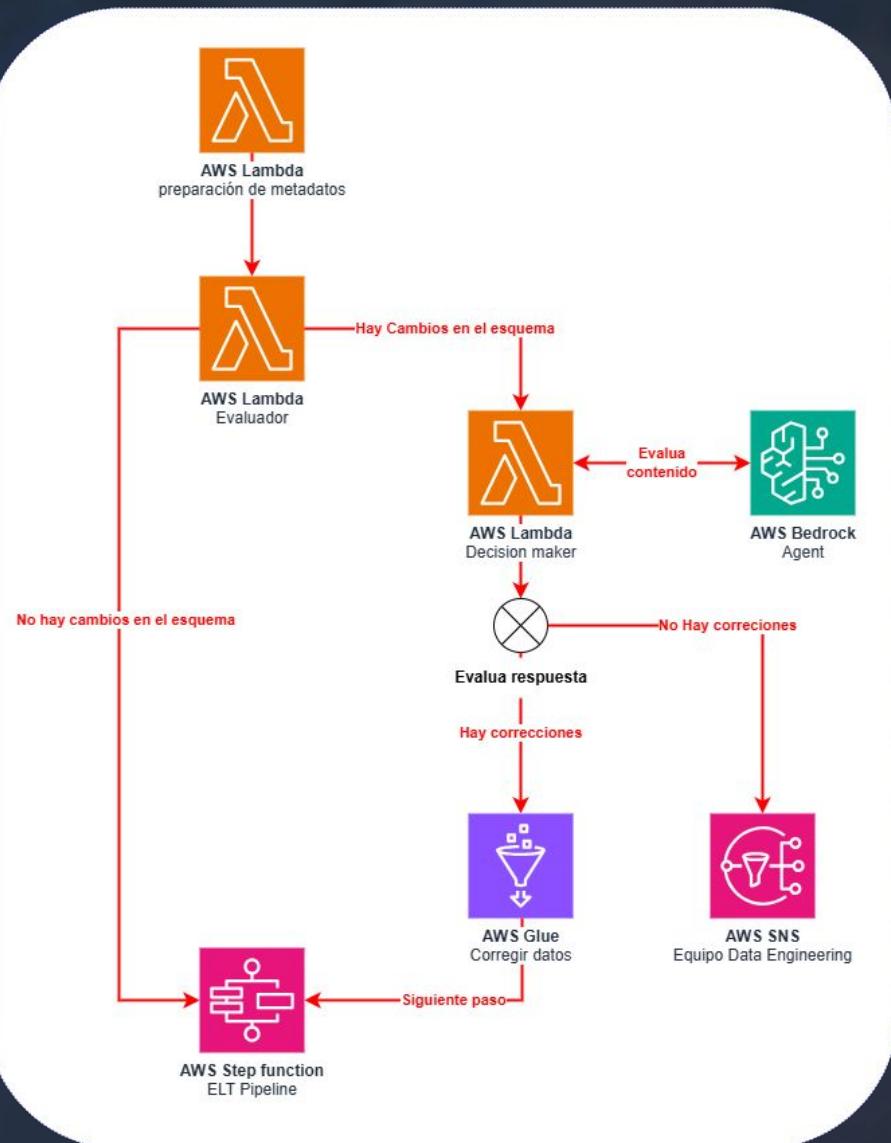
Agentes & AI: definición y estrategia



Agentes & AI: la clave



Agentes & AI: el flujo



Lección aprendidas



El riesgo ante el negocio

En la forma del uso de los datos con el negocio guia los riesgos



Usa bien la IA en el pipeline

Usar la AI en el lugar y forma adecuada es ser eficiente.



Monitoreo everywhere

Saber qué cambios se realizaron y en donde es lo más importante





Carlos Fernando Chicata

Data & Solution cloud engineer
TRD

Un curioso con bases en ciencias de la computación sobre temas relacionados con ingeniería de datos y privacidad de datos; que paso por algunas experiencias laborales diversas y en aumento.

¡Miembro de AWS User Group Arequipa y Perú! y tuve en su momento varios insignias de Top Community Linkedin en su momento.

Conejos, viajar (en lo posible), memes, comidita, certificaciones y compartir conocimientos & experiencias con flow peruano son temas importantes en mi vida.

Spoiler: Disculpen por el cambio de acento 😊





¡Gracias!



Por favor, completa la encuesta de la sesión.





28 de junio de 2025
Bogotá, Colombia

