

PERFIL

Ingeniero de IA especializado en sistemas de agentes y plataformas agénticas para empresa. Diseño, desarrollo y despliegue la pieza completa: el framework de agentes, los MCP servers propios, RAG y memoria, el LLM gateway, los guardrails de seguridad y la infraestructura en Kubernetes que lo hace escalable y observable. Vengo del backend en Python y del DevOps (Kubernetes, Terraform, CI/CD, Azure), así que llevo los casos de uso de la idea a producción. En mi día a día desarrollo con agentes de código como Claude Code y Codex, aplicando harness engineering, skills y context engineering.

Lo que busco: Seguir creciendo profesional y personalmente mientras construyo soluciones tecnológicas de alto impacto: plataformas de agentes de IA que lleguen a producción y equipos capaces de mantenerlas.

EXPERIENCIA

AI Engineer · Plataformas de agentes & DevOps · NTT Data

sept. 2023 - hoy · Híbrido

Puesto: IST - Process & System Technician. Empecé con automatizaciones backend en Python, pasé a DevOps, Cloud y observabilidad, y hoy diseño, construyo y llevo a producción plataformas de agentes de IA para casos de uso corporativos.

Etapas: Backend y automatización -> DevOps y Cloud -> Observabilidad y AIOps -> Machine learning e IA generativa -> Agentic AI y plataformas de agentes

Agentes de IA y plataforma agéntica

- Arquitectura de una plataforma de agentes de IA corporativa con framework propio, modular e independiente del proveedor de modelos.
- Framework propio de agentes: base común para crear agentes nuevos con las mismas garantías de seguridad, trazabilidad y despliegue.
- Sistemas multi-agente: un orquestador coordina agentes especializados y reparte las tareas; los agentes se comunican con el estándar A2A (Agent-to-Agent).
- Agentes con LangChain: tool calling, cadenas y flujos sobre LLMs.
- MCP servers propios (Model Context Protocol) a medida de cada cliente, que exponen sus APIs, datos y sistemas internos como herramientas para los agentes.
- RAG: pipelines de ingesta de documentos, chunking, embeddings y búsqueda vectorial sobre el conocimiento de la empresa.
- Memoria de agentes a corto plazo (contexto de la conversación) y a largo plazo (información que debe persistir entre sesiones).
- LLM Gateway como punto único de acceso a modelos: abstracción multi-proveedor, enrutado de modelos y control centralizado de uso.
- Guardrails de seguridad para entornos corporativos: validación de entradas y salidas, detección de prompt injection, protección de datos sensibles y políticas de la empresa.
- Skills: capacidades empaquetadas y reutilizables que los agentes cargan cuando las necesitan.
- Prompt y context engineering para controlar qué contexto recibe cada agente y obtener respuestas consistentes.
- Modelos en local con Ollama para desarrollo y pruebas.

Plataforma, Kubernetes y DevOps

- Cada componente de la plataforma de IA (orquestador, agentes, LLM gateway, guardrails, RAG, memoria y MCP servers) desplegado como microservicio en su propio pod, con escalado horizontal independiente.
- Despliegue y mantenimiento de clústeres Kubernetes y OpenShift en producción y preproducción sobre Microsoft Azure.
- Recursos de Kubernetes por servicio: Deployments, Services, ConfigMaps, Secrets, health probes, límites de recursos y escalado automático.
- Infraestructura como código con Terraform y gestión de configuración con Ansible.
- Pipelines CI/CD para construir, probar y desplegar servicios e imágenes de contenedor.
- Gestión de contenedores e imágenes Docker en entornos productivos.
- Administración de servidores Linux, scripting en Bash y Nginx como servidor web y proxy inverso.

Observabilidad y operación

- Monitorización de aplicaciones, servidores Linux y clústeres Kubernetes con New Relic: dashboards y alertas.
- Autoremediaciones automáticas a partir de alertas de New Relic.
- Instrumentación con OpenTelemetry: trazas, métricas y logs de servicios y de aplicaciones de IA, incluidas las llamadas de los agentes a modelos y herramientas.
- Logging, métricas y trazabilidad para garantizar la estabilidad de los sistemas.
- Herramientas de monitorización con AIOps.

Automatización, backend y machine learning

- Automatizaciones en Python para procesos de negocio e infraestructura.
- Jobs y procesos programados con Rundeck y orquestación de flujos con Apache Airflow.
- APIs REST y microservicios en Python sobre PostgreSQL.
- Modelos de machine learning y redes neuronales con TensorFlow.
- Desarrollo e integración de soluciones con inteligencia artificial en aplicaciones y procesos existentes.
- IA generativa aplicada a procesos reales de negocio.

Desarrollo con agentes de código

- Desarrollo diario con agentes de código: Claude Code y OpenAI Codex.
- Harness engineering: instrucciones de proyecto, skills, subagentes, hooks y MCP para que los agentes de código trabajen con autonomía y de forma fiable.
- Verificación automática de lo que genera el agente con tests y linters.

Casos de uso y forma de trabajo

- Casos de uso completos: análisis con negocio, diseño de arquitectura, desarrollo, despliegue y operación.
- Herramientas a medida para necesidades muy específicas de cliente gracias a MCP servers propios.
- Preparado para montar una plataforma desde cero y coordinar a un equipo técnico en su construcción.
- Trabajo en equipo con metodología ágil.

Logros

- Construcción de una plataforma de agentes de IA modular y escalable: orquestador, LLM gateway, guardrails, RAG, memoria y MCP servers desplegados como servicios independientes en Kubernetes.
- Despliegue y operación de clústeres Kubernetes y OpenShift en entornos productivos sobre Microsoft Azure.
- Desarrollo con agentes de código (Claude Code, Codex) integrado en mi flujo de trabajo diario.
- Casos de uso end-to-end para cliente, desde el análisis con negocio hasta producción, con herramientas MCP desarrolladas a medida para necesidades muy específicas.
- Alertas y autoremediaciones con New Relic e instrumentación de servicios y agentes con OpenTelemetry.
- Automatización de procesos generales con Python, Rundeck y Apache Airflow.

Tecnologías: Agentic AI, LangChain, MCP, A2A, RAG, LLM Gateway, Guardrails, Claude Code, Codex, Python, Kubernetes, OpenShift, Docker, Terraform, Ansible, CI/CD, Azure, OpenTelemetry, New Relic, Ollama, TensorFlow, Rundeck, Apache Airflow, PostgreSQL, Linux, Git

Becario · NTT Data mar. 2023 – jul. 2023 · Híbrido

Prácticas del Grado en Ingeniería Informática y desarrollo de TFG full-stack con microservicios.

- Trabajo en equipo con metodología ágil.
- Desarrollo de automatizaciones en Python.
- Gestión de infraestructura con IaC (Terraform, Pulumi, Ansible).
- Planificación, desarrollo e implantación de un TFG full-stack con arquitectura de microservicios.
- Primer contacto con contenedores y Kubernetes en entornos productivos.
- Desarrollo de automatizaciones para procesos generales.

Tecnologías: Python, Kubernetes, Docker, Terraform, Pulumi, Ansible, Rundeck, React, JavaScript, Azure, Nginx, Linux, API REST, PostgreSQL, Git

Encargado informático · Cooperativa San Isidro Labrador, Las Labores sept. 2022 · Presencial

Responsable del sistema de medición del grado alcohólico de la uva y del sistema informático de tickets durante la campaña de vendimia.

- Medición del grado alcohólico de la uva.
- Gestión de los tickets de la cosecha en el sistema informático.

Becario · Telefónica abr. 2019 – jun. 2019 · Presencial

Prácticas del ciclo de Técnico Superior en Desarrollo de Aplicaciones Multiplataforma.

- Trabajo en equipo.
- Desarrollo e integración de herramientas web internas.
- Diseño y desarrollo de herramientas web internas.

Tecnologías: HTML, CSS, JavaScript, PHP, Git

PROYECTOS

Plataforma de agentes de IA · NTT Data · hoy · código privado

Plataforma corporativa para construir y operar agentes de IA, con framework propio y cada componente desplegado como servicio independiente en Kubernetes. Orquestación multi-agente con A2A. MCP servers propios por cliente. RAG, memoria, LLM gateway y guardrails. Observabilidad con OpenTelemetry. LangChain, MCP, A2A, RAG, Kubernetes, OpenTelemetry

MCP servers a medida · NTT Data · hoy · código privado

Servidores MCP propios que exponen a los agentes las APIs, datos y sistemas internos de cada cliente, con control de qué puede hacer cada herramienta. Herramientas para casos de uso muy específicos. Código propio, sin depender de integraciones cerradas. Desplegados como pods independientes. MCP, Python, Kubernetes

Observabilidad y autoremediación · NTT Data · código privado

Monitorización de aplicaciones, servidores Linux y clústeres Kubernetes con alertas que disparan autoremediaciones, e instrumentación con OpenTelemetry. Dashboards y alertas en New Relic. Autoremediaciones automáticas. Trazas, métricas y logs con OpenTelemetry. New Relic, OpenTelemetry, Kubernetes, Linux

Clústeres Kubernetes en Azure con IaC · NTT Data · código privado

Despliegue y mantenimiento de clústeres Kubernetes y OpenShift en producción y preproducción, con infraestructura como código y pipelines CI/CD. Terraform y Ansible. Pipelines CI/CD. Producción y preproducción sobre Azure. Kubernetes, OpenShift, Terraform, Ansible, Azure

TFG full-stack con microservicios · NTT Data · 2023

Trabajo de Fin de Grado desarrollado durante las prácticas: planificación, desarrollo e implantación de una aplicación full-stack con arquitectura de microservicios. Prácticas del Grado en Ingeniería Informática. Microservicios, contenedores e IaC. Python, React, Docker, Kubernetes, Pulumi

Iceflix · Grado en Ingeniería Informática

Sistema distribuido de streaming basado en microservicios. Proyecto del Grado en Ingeniería Informática. Python, ZeroC Ice y ZeroC IceStorm. Python, ZeroC Ice, IceStorm · github.com/Diego303/Iceflix

Infoge · Grado en Ingeniería Informática

Herramienta de información sobre localizaciones con APIs de mapas y de Twitter. Proyecto del Grado en Ingeniería Informática. Python y APIs externas. Python, APIs · github.com/Diego303/Infoge

Vídeo interactivo · Grado en Ingeniería Informática

Vídeo en el que el usuario interactúa con varios elementos. Proyecto del Grado en Ingeniería Informática. Angular. Angular · github.com/Diego303/Video-Interactivo

Shoppyp · Grado en Ingeniería Informática

Tienda temática. Proyecto del Grado en Ingeniería Informática. UWP y C#. C#, UWP · github.com/Diego303/Shoppyp

ERP · Ciclo DAM

Planificación de recursos empresariales en C#. Proyecto del ciclo de Desarrollo de Aplicaciones Multiplataforma. C#. C# · github.com/Diego303/ERP-DAM

STACK

CERTIFICACIONES

COMPETENCIAS

IDIOMAS

Agentic AI
MCP (Model Context Protocol), RAG, A2A (Agent-to-Agent), Sistemas multi-agente, LLM Gateway, AI Guardrails, Agentic AI, Memoria de agentes, Agent Skills, Prompt & Context Engineering, LangChain, Ollama

Desarrollo con IA
Claude Code, Harness Engineering, OpenAI Codex

Plataforma & DevOps
Kubernetes, Terraform, OpenShift, Azure, CI/CD, Docker, Ansible, Linux, Bash, Nginx

Observabilidad
OpenTelemetry, New Relic

Backend & datos
Python, API REST, Apache Airflow, PostgreSQL, Rundeck, TensorFlow, Java

Frontend y herramientas
JavaScript, React, Astro, Vite, HTML, CSS, Git, GitHub

CKAD: Certified Kubernetes Application Developer
The Linux Foundation · 2025

HashiCorp Certified: Terraform Associate (003)
HashiCorp · 2024

New Relic Full Stack Observability Practitioner
New Relic · 2024

Microsoft Certified: Azure AI Fundamentals (AI-900)
Microsoft · 2023

FORMACIÓN

Grado en Ingeniería Informática
Escuela Superior de Informática, Ciudad Real (UCLM) · 2019-2023

Técnico Superior en Desarrollo de Aplicaciones Multiplataforma (DAM)
IES Maestre de Calatrava, Ciudad Real · 2017-2019

Capacidad de aprendizaje
De backend a DevOps, de DevOps a machine learning y de ahí a agentes de IA en algo más de tres años.

Orientación a resultados
Llevo los casos de uso hasta producción, no me quedo en la prueba de concepto.

Compromiso
Aporto lo mejor de mí a cada proyecto y a mis compañeros.

Adaptabilidad
Me adapto a las necesidades de cada cliente, hasta escribir la herramienta que su caso necesita.

Trabajo en equipo
Trabajo en equipos empresariales con metodología ágil.

Visión de extremo a extremo
Del análisis con negocio a la arquitectura, el código, el despliegue y la operación.

Español
Nativo

Inglés
Lectura técnica fluida, conversación básica

INTERESES

Agentic AI
Sistemas multi-agente, MCP, A2A, Estándares abiertos

Desarrollo con IA
Agentes de código, Harness engineering

Platform Engineering
Kubernetes, Infraestructura como código, Observabilidad

CAPACIDADES

Agentic AI y plataformas de agentes

- Arquitectura de plataformas de agentes de IA para empresa con framework propio, modulares e independientes del proveedor de modelos.
- Sistemas multi-agente: orquestación, delegación de tareas y comunicación entre agentes con el estándar A2A.
- MCP servers propios para conectar agentes con las APIs, datos y sistemas internos de cada cliente.
- RAG: ingesta de documentos, chunking, embeddings y búsqueda vectorial sobre conocimiento corporativo.
- Memoria de agentes a corto y largo plazo para mantener el contexto entre conversaciones y tareas.
- LLM Gateway: punto único de acceso a modelos con enrutado multi-proveedor y control de uso.
- Guardrails de seguridad: validación de entradas y salidas, prompt injection y datos sensibles.
- Skills reutilizables que amplían las capacidades de los agentes sin tocar su núcleo.

Desarrollo con agentes de código

- Desarrollo diario con agentes de código: Claude Code y OpenAI Codex.
- Harness engineering: instrucciones de proyecto, skills, subagentes, hooks y MCP para que el agente trabaje con autonomía y de forma fiable.
- Prompt y context engineering para obtener resultados consistentes.
- Revisión y verificación automática de lo que genera el agente con tests y linters.

Platform Engineering y DevOps

- Cada componente de la plataforma de IA como servicio independiente en su propio pod, con escalado horizontal por componente.
- Despliegue y mantenimiento de clústeres Kubernetes y OpenShift en producción sobre Microsoft Azure.
- Despliegue de plataformas y aplicativos en clústeres Kubernetes de producción.
- Infraestructura como código con Terraform y Ansible.
- Pipelines CI/CD y automatización de despliegues.
- Gestión de imágenes y contenedores en entornos productivos.

Observabilidad

- Monitorización de aplicaciones, servidores Linux y clústeres Kubernetes con New Relic.
- Alertas, dashboards y autoremediaciones automatizadas.
- Instrumentación con OpenTelemetry: trazas, métricas y logs de servicios y de aplicaciones de IA.
- Logging, métricas y trazabilidad para garantizar la estabilidad de los sistemas.
- Monitorización de aplicaciones de inteligencia artificial.

Machine learning e IA generativa

- Diseño y desarrollo de modelos de aprendizaje automático y redes neuronales con TensorFlow.
- Desarrollo de soluciones con Python y librerías de IA.
- Integración de IA generativa en procesos reales de negocio.
- LLMs en local con Ollama para desarrollo y pruebas.

Backend y automatización

- Automatizaciones en Python para procesos de negocio e infraestructura.
- Diseño y desarrollo de APIs, servicios escalables y microservicios en Python.
- Programación avanzada en Python.
- Gestión eficiente de bases de datos y lógica de negocio.
- Base en Java y C#.

Casos de uso y liderazgo técnico

- Casos de uso completos: análisis con negocio, arquitectura, desarrollo, despliegue y operación.
- Herramientas a medida para necesidades muy específicas de cliente gracias a MCP servers propios.
- Preparado para montar una plataforma desde cero y coordinar a un equipo técnico en su construcción.

Frontend

- HTML, CSS y JavaScript.
- Frameworks modernos como React y Astro.
- Integración con APIs backend y cuidado de la experiencia de usuario.