

JOAN MARQUEÑO SUREDA

Developer · Python · LLMs · Agentes · Kubernetes

✉ joan.marqueno@gmail.com · [in linkedin.com/in/joanmarqueno](https://www.linkedin.com/in/joanmarqueno) · github.com/JoanMarqueno · 📍 Tarragona, España



Python · TypeScript · LangChain · Azure OpenAI · RabbitMQ · Kubernetes · KEDA · Redis · PostgreSQL · pgvector · Docker · Azure · AWS · Hono · Next.js · Flutter

RESUMEN PROFESIONAL

Desarrollador especializado en **Python** y **sistemas LLM en producción**. **3+** años construyendo una plataforma IDP enterprise con **+30 clientes activos**, **+1M de documentos procesados** y contratos de hasta **2M de documentos anuales**. Experiencia real en agentes con tool calling, orquestación event-driven, resolución de problemas de producción y arquitectura hexagonal en Kubernetes (Azure, AWS y on-premise).

EXPERIENCIA

Developer — Python & IA Aplicada

Jun 2023 – Actualidad

INETUM · Tarragona, España · Proyecto IIM-AI4CS

- Diseñé e implementé **+20 microservicios Python** especializados orquestados por un manager de flujos configurables y consumidos vía **RabbitMQ**: OCR, clasificación con Document Intelligence y LLM, extracción con GPT y tool calling, validación simple y cruzada, anonimización, transcripción multimedia con Whisper y conversión de formato.
- Agente GPT en producción con **tool calling** (crop de imagen, Bing Grounding vía LangChain) y **structured outputs** con strict mode y esquemas dinámicos por tipo documental; flujo **HITL** completo cuando las validaciones no superan el umbral de confianza.
- Resolví problemas reales de producción: token overflow por imágenes acumuladas en historial de agente, timeouts en lotes grandes, race conditions entre pods con distributed locking en **Redis**, y rate limits de Azure OpenAI con retry exponencial cogiendo el retry-after del header.
- Escalado automático con **KEDA** basado en profundidad de cola RabbitMQ, tenant isolation con colas y scalers dedicados por cliente, y observabilidad de costes de tokens por proyecto, modelo y servicio con dashboard propio y capping por variable de entorno.
- Plataforma desplegada en **Azure, AWS y on-premise** con arquitectura hexagonal, **Kubernetes, Keycloak** y configuración por cliente: **+30 clientes activos**, **+200.000 documentos/mes**, contratos de hasta **2M de documentos anuales**.
- **RAG** en MVP de gestión documental: chunking semántico con enriquecimiento de keywords y resumen, búsqueda vectorial con Azure AI Search y evaluación con golden datasets.

Assistant / Store / Multi Store Manager

Oct 2017 – Ago 2022

PANDORA · Barcelona / Tarragona, España

- Liderazgo de equipos, KPIs y gestión operativa bajo presión — experiencia hoy aplicada al trabajo con producto y negocio en proyectos de IA enterprise.

PROYECTOS PERSONALES

MysticMe — App móvil publicada en Google Play

2023 – Actualidad

- App de tarot y horóscopo desarrollada end-to-end de forma independiente. **Flutter · Dart · Supabase · Firebase Crashlytics**. Estrategia de crecimiento con TikTok/CapCut.

- Web orientada a captación de contactos. **React** · **TypeScript** · **Tailwind CSS** · **Vite**.

FORMACIÓN & CERTIFICACIONES

CFGS Desarrollo de Aplicaciones Multiplataforma

2022 – 2024

Institut Francesc Vidal i Barraquer · Nota: **9,80** — Matrícula de Honor

Microsoft Certified: Azure AI Engineer Associate

Junio 2024

SKILLS

IA & Agentes: LLMs, tool calling, structured outputs, HITL, visión multimodal, Whisper, LangChain, Azure OpenAI, Document Intelligence, RAG, pgvector.

Backend: Python, TypeScript, Hono, Zod, OpenAPI/Swagger, RabbitMQ, Redis, Peewee, Drizzle.

Cloud & Infra: Kubernetes, KEDA, Azure, AWS, on-premise, Docker, Helm, Jenkins, Grafana, Keycloak.

Frontend: Next.js, React, Tailwind CSS, Flutter.

Idiomas: Español y catalán nativos · Inglés conversacional.