MARIANO DÍAZ

Senior Full Stack Engineer | AWS & Generative AI

TypeScript · Node.js · React / Next.js · React Native · PostgreSQL · AWS · Terraform · LLMs · RAG · AI Agents

lth.nano@gmail.com | linkedin.com/in/diaznano | +54 9 381 412-6387

Tucumán, Argentina (GMT-3) · Disponibilidad para trabajo remoto y relocalización

PERFIL PROFESIONAL

Ingeniero de software full stack con más de 7 años de experiencia construyendo productos web y mobile y llevándolos a producción en AWS. Trabajo en formato forward deployed: junto con cada cliente entiendo el proceso, diseño la solución y la implemento integrada a sus sistemas hasta producción. Desarrollo backends en Node.js (NestJS, Express) y Ruby on Rails, frontends en React y Next.js y apps en React Native, desplegados en EC2, Lambda, API Gateway, RDS PostgreSQL, S3 y CloudFront, con CI/CD en GitHub Actions y Terraform. Desde febrero de 2024 pongo IA generativa en producción: agentes con RAG, tool calling y MCP que atienden unas 17.000 conversaciones mensuales entre 35 clientes, un pipeline de documentos con OCR multimodal y una arquitectura multiagente. Hoy reparto mi tiempo entre desarrollo y liderazgo técnico de un equipo de unas 10 personas; antes fui Team Lead en TaDa, una app con más de 10M de descargas.

EXPERIENCIA PROFESIONAL

Xetro · Remoto, Argentina / España

Consultora de software e IA donde trabajo desde 2019. A través de Xetro tuve asignaciones de staff augmentation en las empresas que figuran más abajo.

Director de Proyectos de Software e IA

Noviembre 2024 – Actualidad

Rol 50/50 entre desarrollo hands-on y liderazgo técnico de un equipo de unas 10 personas (desarrollo, diseño y paid media), para una cartera de 35 clientes de salud, turismo, logística, gastronomía y retail en Argentina y España.

- **Forward deployed con clientes:** relevo el proceso de cada cliente (salud, turismo, logística, gastronomía y retail), defino la solución y la implemento integrada a sus CRMs, bases de datos de sistemas de gestión y medios de pago, acompañando hasta producción.
- **Agentes de IA en producción:** agentes conversacionales sobre WhatsApp Business API con tool calling, memoria conversacional y derivación a un asesor según confianza y riesgo: unas 17.000 conversaciones mensuales entre 35 clientes (500 a 1.000 por cliente, medidas en Chatwoot y el CRM propio), con cerca del 80% cerradas por el agente sin derivación a un asesor.
- **Calidad y costo de la IA:** elijo y combino modelos (Claude, OpenAI, Gemini) según costo, latencia y precisión; los datos que cambian (precios, stock, turnos) se consultan siempre a la base de datos, con reglas de abstención. Migré el contexto a RAG sobre Pinecone y pgvector: cerca de 15% menos de costo de API (tokens) por conversación, medido mes contra mes.
- **IA documental y multiagente:** pipeline de extracción y validación de documentos con OCR multimodal (GPT-4o-mini) y un agente extractor (Claude Sonnet), unos 100 documentos por mes; arquitectura multiagente de una plataforma de logística; agentes conectados a herramientas y datos internos con MCP y function calling; automatizaciones en n8n.
- **Despliegue en AWS:** proyectos en NestJS, Express, Ruby on Rails y Next.js sobre EC2 detrás de Application Load Balancers, con Route 53 y SSL; RDS PostgreSQL con snapshots; imágenes y documentos en S3, servidos con CloudFront cuando hacía falta.
- **Serverless, autenticación y mobile:** funciones AWS Lambda en JavaScript (Node.js) y APIs con API Gateway; login de usuarios con Amazon Cognito; apps React Native con AWS Amplify, notificaciones push y carga de imágenes a S3.
- **CI/CD, infraestructura y observabilidad:** pipelines en GitHub Actions para desplegar en AWS; parte de la infraestructura en Terraform y contenedores con Docker; monitoreo con CloudWatch (logs, métricas y alarmas), Sentry y Grafana, y auditorías automáticas diarias de las respuestas de los agentes. Reduje costos consolidando servicios en menos instancias.
- **Producto e integraciones:** plataforma de logística de cargas en Rails 7 organizada en 12 módulos; integraciones con Mercado Pago, Chatwoot y Meta Conversions API.

Team Lead React

Marzo 2023 – Febrero 2025 · Promovido a Director en noviembre de 2024

- Lideré un equipo de hasta 10 personas en una aplicación web de gestión de turnos para una empresa internacional: planificación, estimación, code review y Scrum.

- Desarrollé con React, Redux y APIs REST.
- En febrero de 2024 puse en producción los primeros agentes de IA para clientes.

Entrevistador Técnico | Eureka Labs

Agosto 2025 – Junio 2026 · Remoto · Contrato temporal, en paralelo a Xetro

- Evalué a cerca de 30 candidatos para posiciones JavaScript (Node.js, React, Next.js y React Native) con criterios de evaluación y ejercicios de live coding propios, y redacté las recomendaciones de contratación.

React Native Developer | Bankingly

Junio 2024 – Noviembre 2024 · Remoto · Staff augmentation a través de Xetro

Plataforma de banca digital para instituciones financieras.

- Desarrollé flujos críticos de banca digital (login, transferencias, pagos y biometría) con React Native, TypeScript, Redux y Jest.

React Native Developer | Cultura IT

Noviembre 2023 – Junio 2024 · Tucumán, Argentina · Staff augmentation a través de Xetro

GoodSeed Notebook: app de registro de cultivos que reemplazó un proceso manual en papel.

- Implementé una arquitectura offline-first con sincronización de datos y escribí módulos nativos de cámara para Android e iOS.

Team Lead React Native / React Native Developer | Onetree

Febrero 2021 – Marzo 2023 · Remoto · Staff augmentation a través de Xetro · Team Lead desde febrero de 2022

Contexto del producto: TaDa, app de delivery de bebidas con más de 10M de descargas en más de 10 países.

- Ingresé como desarrollador y al año pasé a liderar uno de los dos equipos de 5 desarrolladores del proyecto.
- Refactoricé el checkout completo, implementé la analítica con RudderStack y mejoré la performance junto con backend (React Native, Redux, GraphQL).

Desarrollador Web y Mobile | Diveria, Kubikware y Xetro

Junio 2019 – Septiembre 2021 · Remoto · Staff augmentation a través de Xetro y proyectos de Xetro

- Diveria: app web de logística para Avancargo (Next.js, NestJS). Kubikware: apps React Native Modamily y Breast Advocate. Xetro: Workast y Forbes8. Todas publicadas en App Store y Google Play.

FORMACIÓN E IDIOMAS

Ingeniería en Sistemas de Información | Universidad Tecnológica Nacional (UTN) · 2013 – 2019 · Estudiante avanzado: resta 1 materia

Analista de Sistemas Informáticos (título intermedio) | Universidad Tecnológica Nacional (UTN) · 2019

Intercambio académico | Université de Technologie de Belfort-Montbéliard, Francia · 2017

Idiomas: Español (nativo) · Inglés (B1, intermedio) · Francés (B1, certificación DELF) · Catalán (básico, en curso)

CAPACIDADES CLAVE

- **IA generativa:** LLMs (OpenAI, Claude, Gemini), agentes de IA, sistemas multiagente, RAG, embeddings, Pinecone, pgvector, prompt engineering, tool calling, MCP, guardrails, human-in-the-loop, OCR multimodal, n8n
- **AWS:** EC2, Lambda, API Gateway, S3, CloudFront, RDS PostgreSQL, Application Load Balancer, Route 53, SSL, Cognito, Amplify, CloudWatch; en entornos de otros equipos: VPC, DynamoDB, ElastiCache, EKS
- **DevOps y observabilidad:** Terraform, GitHub Actions, Docker, Git, Sentry, Grafana, RudderStack; también Railway, Supabase y Neon
- **Backend:** TypeScript, Node.js, NestJS, Express, Ruby on Rails, REST APIs, GraphQL, webhooks; Python para automatización y scraping
- **Frontend y mobile:** React, Next.js, Redux, Tailwind CSS, Hotwire, React Native (iOS y Android)
- **Bases de datos:** PostgreSQL, MySQL, MongoDB, Redis, Supabase, Neon, optimización de consultas
- **Integraciones:** WhatsApp Business API, Chatwoot, CRMs, Mercado Pago, Meta Conversions API
- **Liderazgo:** liderazgo técnico de equipos de 5 a 10 personas, code review, estimación y planificación de sprints, Scrum, mentoring, entrevistas técnicas, relación con clientes, trabajo directo con clientes