Senior Fullstack Engineer (m/w/d)

Full-time I Onsite in Berlin Mitte or Potsdam-Babelsberg

We have two offices, one in Berlin Mitte and one in Potsdam-Babelsberg. You can work from either location.

About PLACED

PLACED is building the agentic sales OS to transform Europe's staffing industry. We help agencies place candidates radically faster and win more business. We have just closed our first institutional funding round and are growing fast with leading staffing firms in Germany. We're here to redefine how an entire industry operates. Join us now in building the next European tech champion.

Your role

We are seeking experienced Senior Full Stack Engineers to drive our product forward. You will own the full development lifecycle, from building responsive Next.js interfaces to architecting high-performance FastAPI and Hono backends. You will design robust PostgreSQL data models, deploy production-grade AI features, and lead our strategic backend migration from Python to TypeScript.

Our tech stack

- Backend: Python 3.13, FastAPI (primary production API, 123+ endpoints); TypeScript 5, Hono, Bun (new backend in active development)
- Frontend: TypeScript, React, Next.js 16
- Database: PostgreSQL 17 (pgvector for vector search), Redis (caching, async flags)
- Infrastructure: AWS (ECS/Fargate, S3, SES, ECR, RDS), GCP (Vertex AI, Document AI, Cloud Translation), Docker
- AI / LLM: OpenAI (embeddings, chat), Google Gemini / Vertex AI, Anthropic Claude, LiteLLM (multi-provider routing), Instructor (structured outputs), Langfuse (LLM observability)
- Auth: WorkOS (AuthKit, JWT)
- Observability: Sentry, Prometheus, Langfuse, structlog
- Tooling: GitHub Actions CI/CD, UV (Python), Bun, Ruff, mypy, ESLint, pre-commit, pytest, Drizzle ORM, SQLAlchemy 2.0 (async)

What you'll do

- Architect and implement responsive, high-performance user interfaces with Next.js 16 and React
- Design and build secure, high-performance RESTful APIs in FastAPI (Python) and Hono (TypeScript/Bun)
- Drive the active migration from our Python backend to the TypeScript/Hono architecture
- Model and optimize relational schemas in PostgreSQL 17, including pgvector embeddings for semantic search
- Build and maintain Redis-backed caching layers and async task patterns
- Orchestrate multi-provider LLM calls via LiteLLM, integrating OpenAI, Gemini/Vertex AI, and Claude
- Instrument LLM pipelines with Langfuse for observability, cost tracking, and quality evaluation
- Containerize services with Docker and deploy to AWS ECS/Fargate and GCP (Vertex AI, Document AI)
- Build and maintain CI/CD pipelines in GitHub Actions with 100% diff-coverage enforcement
- Research, design, and implement autonomous AI agents that orchestrate tasks and workflows using LLMs and custom policies
- Ensure fault tolerance, high availability, and scalability across distributed systems; identify and mitigate bottlenecks
- Partner with product, design, and QA to deliver end-to-end features
- Conduct code reviews and champion best practices

What we are looking for

- 5+ years of professional experience in full-stack development
- Proficiency with Next.js, React, and TypeScript, including SSR, app router, and server components
- Deep experience building Python APIs with FastAPI (or equivalent async framework)
- Strong TypeScript skills, with experience in Hono, Express, or similar server frameworks
- Strong SQL skills and hands-on experience with PostgreSQL, including advanced features (partitioning, indexing, full-text search)
- Experience with vector databases or pgvector for embedding storage and semantic search
- Familiarity with Redis for caching and async coordination
- Solid AWS expertise (ECS/Fargate, S3, SES, ECR, RDS); GCP exposure a plus
- Experience with containerization (Docker) and CI/CD pipelines (GitHub Actions)
- Demonstrated knowledge of distributed systems principles (consistency models, backpressure, graceful degradation)
- Excellent communication skills and ability to work asynchronously

Bonus

- Prior exposure to agentic AI frameworks or building autonomous agents
- Experience with LiteLLM, Instructor, or other LLM orchestration libraries
- Hands-on Langfuse or other LLM observability tooling (Helicone, Braintrust, etc.)
- Familiarity with GCP services: Vertex AI, Document AI, Cloud Translation
- Experience with WorkOS or similar auth platforms
- Background in observability (Prometheus, Sentry, CloudWatch)
- Familiarity with infrastructure-as-code (Terraform)
- Contributions to open-source projects or technical communities

What you get

- Competitive package: Salary + VSOP
- Team: Work in a small, focused and experienced team with high talent density.
- Room to grow: Own projects and take on more responsibility as you deliver.
- Health & learning: Urban Sports Club membership and a €1,000 annual learning budget.
- Late evenings: Dinner and an Uber home when you stay until 8pm or later.
- Everyday support: Monthly budget for Uber or cleaning services.
- Live close to the office: Extra gross salary for living within a 10 or 20 minute walk.

Apply

View this role and apply on JOIN

getplaced.de/careers