

Yasir RAZA

✉ yasirabdali6@gmail.com | 🌐 github.com/Yasirrazaa | 🔗 linkedin.com/in/yasir-raza | 📍 Islamabad



☎ +92 314 5363379 | 📦 pypi.org/user/yasirabdali6 | 🏠 kaggle.com/yasirabdaali | 🌐 yasirrazaa.github.io

SUMMARY: Python backend engineer with 1.5+ years of remote experience on US-based teams, building API and agent-serving infrastructure: **FastAPI** services, WebSocket servers, serverless GPU workers, MCP/HTTP control planes, async PostgreSQL, and S3 storage. Author of two published **PyPI** packages (10K+ downloads).

TECHNICAL SKILLS

APIs & SERVICES: FastAPI, WebSocket, REST design, MCP servers, HTTP control planes, token authentication, async request handling

DATA & STORAGE: PostgreSQL (asyncpg), Alembic migrations, Redis, S3, ETL pipelines, caching

ASYNC & SYSTEMS: asyncio, thread pools, lazy resource management, subprocess isolation, streaming (ijson), multiprocessing

INFRA & DEPLOYMENT: Docker (multi-stage, CUDA images), docker-compose, RunPod serverless, CI/CD (GitHub Actions), Make, Linux

LANGUAGES: Python, TypeScript, SQL, Bash

WORK EXPERIENCE

TEAMLIFT — UNITED STATES (REMOTE)

Junior Data Scientist

Jan 2024 - Dec 2024

- Scraped **thousands of AI tools** from public tool directories, building the structured dataset that powers **Teamo**, Teamlift's recommendation engine.
- Built an automated **verification pipeline** (HTTP status checks, screenshot capture, freshness validation) that kept listings accurate and trustworthy — eliminating stale and duplicate entries at scale.
- Preprocessed and normalized thousands of tool records into model-ready features for the recommendation engine.
- Worked directly with US-based stakeholders remotely, shipping dataset improvements on tight iteration cycles.

Data Science Intern

Jul 2023 - Dec 2023

- Built an automated pipeline collecting **thousands of job postings** from dice.com with **deduplication** — hardened with retries, rate limits, and robust parsers.
- Extracted skill-demand signals from job descriptions at scale; trained **time-series models** to forecast demand, informing product and market decisions.

TECHNICAL PROJECTS

Kausal — Production AI Agent Platform (private)

2026

- Extended the ii-agent framework in **pure Python** (no LangGraph): custom **RAG** pipeline, layered memory, **OpenTelemetry** observability (tracing, metrics, SigNoz), realtime **voice/video via LiveKit**, **LiteLLM** multi-model routing, agent skills, and **S3** storage — backed by PostgreSQL/Redis, Docker microservices, and a **7,275-test** suite.

Chatterbox-Optimized — Serverless TTS API

2025

- Shipped a production **RunPod serverless** worker (rp_handler + CUDA Dockerfile) for text-to-speech: up to **4.5× faster inference** (RTF 0.13), long-text chunking with stitched audio, Whisper-validated candidate selection, and a locally testable API harness.

Cerberus — Multi-Agent Browser Control Plane

2026

- Built an HTTP API + dashboard (**127.0.0.1:9867**) with token authentication, a multi-agent isolation gateway with worker leases and routing, parallel instance management, persistent per-profile session storage, and a **60-tool MCP bridge** — verified by 86 automated checks.

NotebookLLM — MCP Server & Async Execution 2025 - 2026

- Published a **PyPI** package + MCP server (**10K+ downloads**): async Jupyter kernel execution with lazy per-session kernel lifecycle and thread-pool responsiveness, streaming ingestion for 10 MB+ notebooks, and atomic crash-safe writes.

Forge — The API Compiler for the Agentic Era 2024 - 2026

- Emitters producing **Python SDKs** (Pydantic-validated), CLIs, MCP servers, and **17 output types** from one OpenAPI 3.x IR; multi-registry publishing pipeline and 13 compile gates in CI.

Archon 2.0 — Agent Security Platform 2026

- Runtime **defense proxy** (OpenAI-compatible drop-in) with HMAC identity, rate limiting, per-agent policy, output redaction; Postgres-backed policy registry with SCIM v2 provisioning and OIDC SSO; Helm chart + web UI fleet dashboard.

EDUCATION

B.Sc. Software Engineering, National University of Modern Languages (NUML), Islamabad Aug 2026

CERTIFICATIONS

Datasets Expert · Notebooks Expert (13 medals; Datasets rank 731/11,277, best #109), Kaggle **2024 - 2026**

Associate Data Scientist in Python · Data Analyst with Python (Career Tracks), DataCamp **2025 - 2026**

Generative AI Fundamentals · Intro to LLMs · Intro to Generative AI · Intro to Responsible AI, Google Cloud Skill Badges **2023**

Python · Intro to Programming (Professional Certificates), Kaggle Learn **2025**

Advanced LLM Bootcamp, NUST & Sky Electric **2024**