

Muhammad Kashif Irshad

AI Integration & Full-Stack Engineer | React, Python, RAG, APIs

Lahore, Pakistan (UTC+5) | Remote consulting and project delivery | US, UK and Canadian client experience

dev@irshados.com | WhatsApp +92 311 100 1172

linkedin.com/in/mki72 | github.com/mkirshad | irshados.com

PROFILE

Engineer with 18+ years in software, combining production AI integration, React/TypeScript applications, Python APIs and data engineering. Build and operate a multi-tenant document assistant with approximately 117,000 embedded chunks, and deliver CRM automation and internal SaaS features for a US client. Completed long-running UK CRM engagements and a nine-month Canadian data integration project. Own architecture, implementation, deployment and production support.

Availability: Weekly capacity, start date and time-zone overlap by agreement. Based in Pakistan for remote delivery.

CORE SKILLS

AI and automation: OpenAI APIs, RAG, pgvector, embeddings, structured outputs, tool calling, OCR, retrieval evaluation, token budgets, Claude Code.

Application engineering: Python, Django REST Framework, FastAPI, React, TypeScript, Node.js/Express, REST APIs, webhooks, OAuth, PostgreSQL, Redis.

Data and cloud: BigQuery, dbt, Airflow/Composer, Snowflake, Power BI, Google Cloud Run, Cloud Build, AWS, Azure, Docker, CI/CD.

SELECTED CONSULTING AND PROJECT DELIVERY

Employer and client identities withheld for confidentiality. Selected work spans consulting, team delivery and earlier employed roles.

AI-enabled CRM and internal SaaS platform

Jul 2025 - present | Remote contract | Full-stack development lead

- Build AI and CRM features in a production SaaS platform using FastAPI on Cloud Run, React/TypeScript, BigQuery and Firestore; integrate Outlook and business APIs.
- Implement OpenAI-assisted SMS/email drafting with channel length limits and meeting-transcript workflows that create assigned CRM tasks with source-based de-duplication.
- Develop role-based access, JWT authentication, filtered data grids and dashboards; maintain scheduled actions, duplicate-send prevention and retries. Deploy with Cloud Build and Secret Manager.

Document assistant and AI CRM products

2025 - present | Independent product development

- Built and operate **Irshad AI Employee**, a multi-tenant assistant serving private customer documents and a reference library of 946 statutes and 2,503 High Court judgments, with approximately 117,000 embedded chunks.
- Implemented Django REST, PostgreSQL/pgvector, vision-model OCR and retrieval controls across PDFs, Word files and scans. Support English, Urdu and Roman Urdu through WhatsApp, a web widget and a portal, with tenant document isolation and usage metering.
- Designed source citations and refusal behaviour for unsupported questions. In the published evaluation, gold-source accuracy was **74.5% versus 60.0%** for basic dense retrieval; the manual audit found zero fabricated citations.
- Built **Irshad AI CRM**, a React/TypeScript sales workspace with pipelines, activity, customer/product/inventory views and an OpenAI analyst over PostgreSQL/BigQuery data. Available as a live demo.

Product proof: [Document assistant demo](#) | [AI CRM](#)

Regulatory registration and environmental monitoring

Nov 2024 - Jun 2025 | Full-stack project delivery

- Delivered a public-facing regulatory registration system using React, Django and PostgreSQL: public registration, approval stages, role-based access and an installable PWA. Also developed an air-quality monitoring dashboard.

INTERNATIONAL CLIENT DELIVERY

Data integration and application delivery

Oct 2019 - present | Selected consulting and team-delivered projects

- **Canada, nine months:** Integrated QuickBooks Desktop SDK, Fishbowl inventory and Excel with BigQuery. Built .NET extraction utilities, historical and nightly loads, stored-procedure cleaning/de-duplication, and Power BI reporting. Completed with 5-star feedback.
- **UK, nine- and twelve-month engagements:** Delivered a CRM and a data approval dashboard using React and React Native; both received 5-star feedback. Also integrated Firebase analytics into BigQuery and Data Studio for a UK client.
- **US, eight-plus months through a delivery team:** Built marketing data workflows with Stitch/Fivetran, BigQuery, dbt and Airflow/Composer; applied three-sigma and k-means anomaly checks and maintained logging and alerts.
- Delivered AWS reporting workflows using Redshift, Glue, S3, Lambda and React, alongside API integration and data warehouse assignments.

Freelance work and client feedback: [upwork.com/freelancers/mkirshad](https://www.upwork.com/freelancers/mkirshad)

EARLIER PROJECT DELIVERY

Analytics warehouses for property and hospitality

Nov 2019 - Oct 2020 | Data engineering through a delivery team

Delivered analytics warehouse and data pipeline work across web/app analytics, entertainment and hospitality. Used Snowflake, Azure Data Factory/Data Lake, Databricks, PySpark/Scala and Power BI; validated pipelines through sampling and SQL checks.

Smart-metering software and data architecture

Dec 2017 - Oct 2019 | Technical leadership and architecture

Led software and data architecture for smart-metering systems, including meter data collection, warehouse/reporting, web services and delivery coordination. Worked with PostgreSQL, OBIEE, ODI and SSIS.

Healthcare analytics and business workflow platforms

May 2010 - May 2016 | Database, warehouse and application engineering

Designed databases and data warehouses for healthcare analytics, franchise business workflows and cloud applications. Delivered ETL migrations, reporting and a Laravel learning management system; franchise systems supported 150+ clubs in the US and Canada.

Earlier projects: Healthcare database administration and BI (2008-2010); leasing-system database engineering (2007-2008); transaction-processing databases (2006-2007).

RESEARCH AND PUBLIC CODE

Author of preprints on SSRN and Google Scholar. [SSRN author page](#)

Grounded or Silent: Citation-Faithful Legal Question Answering for Pakistani Law. Research preprint, 2026. PakLegalQA: 360 questions spanning answerable, unanswerable and false-premise cases, with a Roman-Urdu twin set.

DOI: [10.5281/zenodo.22305084](https://doi.org/10.5281/zenodo.22305084) | [Benchmark and code](#) | [Google Scholar](#) | [SSRN](#)

Auto Care AI: Android vehicle-cost application with a Django/Cloud Run backend, fuel logs, GPS trips and Google Drive backups. [Public code and documentation](#)

EDUCATION AND SELECTED CREDENTIALS

BS (Hons) Computer Science | National University of Computer and Emerging Sciences (NUCES-FAST) | 2001-2005

Upwork Skill Certification - React Front-End Development (2024); **Oracle BI Foundation Suite 11g Certified Implementation Specialist** (2016).

Selected professional learning: AI Engineering Essentials, LinkedIn Learning (2024); Microsoft Azure Databricks for Data Engineering, Microsoft/Coursera (2024); Developing Back-End Apps with Node.js and Express, IBM/Coursera (2024).