

# Mostafa Ghadimi

Senior Data & AI Engineer · Tech Lead · Platform & AI

Istanbul, Türkiye · Open to Remote

+98 912 072 7175 | mostafa.ghadimi76@gmail.com | mostafaghadimi.github.io | github.com/mostafaghadimi | linkedin.com/in/mostafaghadimi

## Summary

Senior Data & AI Engineer (6+ yrs) building large-scale data platforms across e-commerce, fintech, real estate, and PropTech — streaming pipelines (Kafka, Debezium, dbt, Airflow), warehouse modelling (ClickHouse, Pinot, PostgreSQL), and the Kubernetes / GitOps backbone that runs them. Designed multiple analytics platforms from blank page to production, ran hiring for the data team (20+ technical interviews), authored engineering practices still in use, and led small teams as senior IC or technical lead. Also ships LLM / agentic systems in production (LangChain, LangGraph, LangSmith) as a second specialty.

## Skills

<b>Languages</b>	Python, SQL, Bash
<b>Data Engineering</b>	Apache Airflow, dbt, Apache Spark, ETL/ELT, CDC, data quality, warehouse modeling (Data Vault, Star/Snowflake)
<b>Streaming</b>	Apache Kafka, Strimzi, Debezium, Kafka Connect, Schema Registry, KRaft
<b>Databases &amp; Analytics</b>	ClickHouse, Apache Pinot, PostgreSQL, MSSQL, Oracle, Metabase, DataHub
<b>AI &amp; Agents</b>	LangChain, LangGraph, LangSmith, FastAPI, Celery, OpenAI, tool calling (@ <b>tool</b> ), skill registries / on-demand knowledge bases, agent middleware ( <b>before_agent</b> / <b>after_agent</b> ), multimodal / vision LLMs, MCP, prompt engineering
<b>Cloud &amp; IaC</b>	AWS (EC2, ECS, ECR, CloudWatch), Kubernetes (RKE/RKE2, Kubespray), Docker, Helm, Terraform, Ansible
<b>DevOps &amp; Observability</b>	ArgoCD, GitLab CI, GitHub Actions, HashiCorp Vault, Prometheus, Grafana, Elasticsearch
<b>Leadership &amp; Process</b>	Hiring & interviewing (20+), 1:1s, mentoring, architectural decisions, agile (Scrum, OKRs), technical writing

## Experience

### Vendoroo, Senior AI Engineer | Remote, San Francisco

Nov 2025 - May 2026

AI agents for property management — automating tenant communication, maintenance triage, and vendor coordination.

- Ship the tenant-facing maintenance-triage agent serving **millions of messages** across **thousands of properties** and **1,000+ property managers**, with monthly agent volume scaling **~3x in 6 months (LangChain, LangGraph, FastAPI)**.
- Modularised the agent graph — conversation policy split from supporting middleware (input cleaning, context batching, reply formatting) — so each layer is independently testable and hot-swappable.
- Added a vision pre-step via **before\_agent** middleware that converts tenant-submitted photos (**100K+** processed) into structured context; multimodal handoff plugged in without changes to downstream nodes.
- Restored **LangSmith** tracing across the LLM fallback path (previously a black box), authored the next-milestone target architecture and decision log, and rebuilt new-engineer onboarding so day-one clone-and-run works across all repos.

### SnappPay, Data Engineering Tech Specialist | Remote, Tehran

Jun 2024 - Oct 2025

Iran's leading buy-now-pay-later (BNPL) fintech, part of the Snapp Group super-app ecosystem — millions of consumers and merchants across the country.

- Promoted from Senior Data Engineer to **Tech Specialist in 5 months**; owned hiring for the data team (**20+ technical interviews**), authored the interview rubric and calibration process still in use, ran weekly 1:1s, and wrote the onboarding playbook.
- Cut cluster provisioning from a multi-day manual ritual to **one command** via an **Ansible** toolkit — ClickHouse cluster, monitoring, and managed PostgreSQL with a zero-downtime upgrade path.
- Built the team's CI/CD framework as reusable **GitLab** pipeline components — every new repo ships DAG delivery, Spark jobs, semantic releases, and proxy-aware Docker builds out of the box.
- Delivered the team's **dbt** project (staging + marts: transactions, orders, carts, credit, user dim) on Airflow + ClickHouse, turning ad-hoc SQL into reviewable, testable analytics code.
- Built a data trust layer — generic DQ framework (completeness / freshness / uniqueness / consistency, chunked + retry-aware), schema-drift alerting to MS Teams via Kafka, daily credit-warehouse backfill — plus the streaming and GitOps backbone (**Strimzi** Kafka on K8s, **ArgoCD**).

### IRITCO, Senior Data Consultant | Remote

May 2024 - Apr 2025

Iranian IT consultancy specialising in data platforms, data governance, and ICT solutions for enterprise and public-sector customers.

- Led **5 engineers** through the greenfield build of a national aid-distribution analytics platform on **ClickHouse**.
- Ran **4 architectural decision sessions** (R&D, prototype, presentation, decision log) covering analytics-DB choice, ETL vs. ELT, and on-prem vs. cloud — evidence-backed sign-off rather than instinct.
- Retired a legacy **Oracle** reporting stack with a Python ingestion pipeline (SQLAlchemy, pandas) that mirrored operational tables into the new warehouse on a known schedule.
- Established the engineering practice the team kept — meta-repo orchestration, Git conventions, code-style guide, mandatory code review, on-call runbook.

## Mofid Securities, Data Architect & Senior Data Engineer | Tehran, Iran

Mar 2023 - May 2024

One of Iran's largest stock brokerages and a top retail-trading platform, with millions of investor accounts.

- Designed and built the brokerage's **first analytics data platform** from a blank page — owned architecture, tool selection, and core framework code across analytics DB, orchestration, GitOps, secrets, BI, and monitoring.
- Automated the migration from a legacy **SQL Server** warehouse to **ClickHouse** with a custom change-data-capture framework that generated destination tables on the fly and handled decimal precision, binary columns, Jalali dates, and late-arriving updates with no manual intervention.
- Delivered a no-code data-quality framework (table-level YAML config, separate query paths for SQL Server and ClickHouse) so onboarding a new table to monitoring is a config change.
- Bootstrapped the **Kubernetes** platform bare-metal to production: **RKE** clusters, **ArgoCD** delivery, ingress for Airflow / Grafana / Longhorn, systemd-managed Kafka cluster (KRaft) feeding the CDC pipeline; packaged Metabase as a reusable **Helm** chart.
- Authored the engineering methodology the team adopted (2-week sprints, 20-min standups, grooming / planning / retro, OKR-driven backlog) and wrote it up as a living team doc.

## Offerland, Cloud Engineer | Remote, Vancouver, Canada

Apr 2023 - Nov 2024

Vancouver-based real-estate marketplace connecting home buyers, sellers, and agents through data-driven tools.

- Cut **AWS spend by ~40%** with no service impact via right-sizing, Reserved Instance commitments, S3 lifecycle policies, and a sweep of orphaned compute and log retention.
- Migrated production from a single EC2 + Docker-Compose host to **AWS ECS** behind an Application Load Balancer with an ECR image-promotion pipeline — removed single-host outage as a failure mode.
- Delivered the company's observability stack (**Prometheus**, **Alertmanager**, **Grafana** dashboards, **CloudWatch** metrics) with on-call routing so issues surfaced before customers noticed.
- Owned the IaC repos provisioning the platform (**Terraform** / **Ansible** modules, GitLab-CI templates, AWS auth and backup pipelines) and unified secret management to a single source of truth.

## Digikala, Data Engineer → Senior Data Engineer | Tehran, Iran

Jan 2021 - Apr 2023

Iran's largest e-commerce platform — “the Amazon of Iran” — among the most-visited websites in the MENA region, serving tens of millions of users.

- Co-built **Backbone**, the internal platform that gave data scientists a one-button path from notebook to production — **15+ DS services** (search, recommendation, query processing, trending, spell correction, synonym, Snowplow) on a shared **Spark + Airflow + Pinot** stack, no Kubernetes YAML required.
- Delivered **up to 10x query speed-up** on billions of rows of clickstream / order OLAP workloads by introducing **Apache Pinot** with hand-tuned star-tree indexes and rewriting the heaviest dashboards.
- Tuned production Kafka brokers (compression, page-cache, ACLs, broker config), reducing consumer lag and disk pressure on the busiest topics.
- Brought **GitOps** (**ArgoCD** + Kustomize + Kaniko/Dagger CI) into the data org, replacing manual **kubect1 apply** with reviewable PRs; provisioned production K8s clusters via **RKE** and **Kubespray**.

## Earlier Experience

May 2019 - May 2022

- **Lions Digital Solutions** (*Software Engineer*, Vancouver, 2021–2022) — Built a sales-analytics web app for Amazon sellers with marketing / CRM data ingestion (Amazon, Google Ads / Analytics, Ontraport, Close, Oncehub, Thinkific, HelloSign) and the CI/CD that ships it.
- **Hoosh Pardaz** (*Algorithmic Trader & Data Scientist*, Tehran, 2020) — Researched Crypto ETFs and built backtested trading strategies for the Canadian market under Dr. Soleymani Baghshah and Dr. Behrouzi.
- **Wize Analytics** (*Infrastructure & Data Engineer*, Tehran, 2019) — Delivered the data infra (message bus, search index, document store, developer toolchain) for a Sharif-University NLP startup advised by Dr. Rabiee and Dr. Ramezani.

## Education

---

2022 - 2024	<b>MSc in Algorithms and Computations</b> , University of Tehran — relevant coursework: Distributed Systems, Advanced Algorithms, Machine Learning	Tehran, Iran
2016 - 2020	<b>BSc in Computer Engineering</b> , Sharif University of Technology	Tehran, Iran

## Awards

---

2019	<b>1st place, Fintech Challenge</b> , 3rd ICT Challenge, Sharif University of Technology	Iran
2016	<b>Top 0.002% nationally (rank 61 of 80,000)</b> , National Universities Entrance Exam (Konkour)	Iran

## Languages

---

<b>English</b>	Professional proficiency
<b>Persian</b>	Native proficiency