SIDDHARTH SINGH BHADAURIA

Senior Cloud, Data & AI Engineer

London, UK • siddharthsbhadauria@gmail.com • +44 7448 722384

Portfolio: siddharthsbhadauria.github.io/portfolio/ • **GitHub:** github.com/siddharthsbhadauria • **LinkedIn:** linkedin.com/in/siddharthsbhadauria

SUMMARY

Senior Cloud and Data Engineer with 12+ years building cloud infrastructure, data pipelines and AI platforms on Google Cloud (GCP), AWS and Microsoft Azure, mostly for banks and other regulated clients. At Deloitte I currently build GCP logging, asset inventory and alerting infrastructure for JPMorgan Chase that handles 50M+ events a day. Before that I maintained the Terraform AWS EKS modules behind 3,600+ deployments and 700+ microservices. I also co-built Deloitte's internal GenAI platform used by 75,000+ staff.

SKILLS

- **Google Cloud:** BigQuery, Pub/Sub, Cloud Run, Cloud Functions, Cloud Scheduler, Cloud Logging (Log Router, sinks, log buckets), Cloud Monitoring, Cloud Asset Inventory, Cloud Storage, Cloud KMS, Private Service Connect, Vertex AI, Artifact Registry, Dataproc
- **AWS:** EKS, S3, DynamoDB, RDS, MSK, SQS, SNS, Lambda
- **Azure:** AKS, Container Apps, Functions, Data Factory, API Management, Blob Storage, Cosmos DB, SQL Server, Key Vault, Azure OpenAI
- **Infrastructure & DevOps:** Terraform, Kubernetes, Docker, GitHub Actions, GitLab CI/CD, Azure DevOps, ARM templates
- **Data Engineering:** Apache Kafka, Kafka Connect, Apache NiFi, dbt Core, DuckDB, AbInitio, IBM DataStage, PostgreSQL, Oracle
- **AI & ML:** RAG, pgvector, Qdrant, DeepEval, Gemini, Llama 3, Mistral 7B, scikit-learn, MLflow, FastAPI
- **Security:** CMEK with Cloud KMS HSM, Workload Identity Federation, PII masking, GPG key management
- **Languages & Tools:** Python, Go, SQL, PySpark, Bash, pytest, GitHub Copilot, Claude Code

EXPERIENCE

Senior Cloud Infrastructure & Observability Engineer | Deloitte (Client: JPMorgan Chase & Co.) — **London, UK | Jun 2026 - Present**

GCP logging, observability and security platform

- Built organisation-wide log routing in Terraform for audit, network flow and platform logs, fanning organisation-level sinks into 16 Pub/Sub topics that stream to the SIEM and the Kafka databus.
- Cut long-term log storage costs by about 94% with GCS lifecycle rules that move logs to Archive after 90 days and delete them at 540 days.
- Rolled out Cloud KMS HSM customer-managed keys (CMEK) with 30-day rotation across GCS, Cloud Logging, BigQuery and Pub/Sub.
- Designed an active/standby asset inventory export across two regions: 12 scheduled jobs a day on Cloud Run and Cloud Functions with a 2-hour freshness target.
- Moved the export service to Cloud Run v2 without downtime, using container digest pinning, Direct VPC egress and keyless Workload Identity Federation.
- Built an alert pipeline that sends Cloud Monitoring alerts to Kafka and ServiceNow through Pub/Sub, with delivery in under 60 seconds.
- Wrote Terraform blueprints that deploy unchanged across environments, each covered by automated BDD tests (pytest-bdd).

Parental Leave | Career Break — **London, UK | Oct 2025 - May 2026**

Senior Consultant, DevOps Engineer | Deloitte (Client: JPMorgan Chase & Co.) — **London, UK | Aug 2024 - Sep 2025**

Terraform AWS EKS and microservices platform

- Maintained the core Terraform AWS EKS cluster modules and the blueprints that connect services to S3, DynamoDB, RDS, MSK, SQS and SNS, cutting provisioning time by 65%.
- Scaled the EKS module to 3,600+ deployments running 700+ microservices for 20+ business lines.
- Reworked the cluster module into a multi-tenant model with per-tenant resource allocation and isolation.
- Led Kubernetes upgrades from v1.19 to v1.34 with written migration guides, across 50+ production releases with no downtime.
- Built Python and Go test suites that validate module changes before they reach deployment.

Senior Consultant | Deloitte (Internal: PairD GenAI Platform) — **London, UK | May 2024 - Aug 2024**

- Co-built Deloitte's internal AI assistant for 75,000+ staff, serving 100,000+ queries a day across OpenAI and Google Gemini models.
- Kept model traffic on private networks using Google Private Service Connect and custom DNS forwarding.
- Built RAG pipelines on Postgres pgvector and Qdrant and tuned prompt templates to improve retrieval accuracy.
- Converted Azure workloads to Terraform and ARM templates and deployed containers to Azure Container Apps and AKS.

Senior Consultant | Deloitte (Internal: GenAI LLM Evaluation) — **London, UK | Apr 2024 - May 2024**

- Built an evaluation pipeline with DeepEval scoring answer relevancy, contextual accuracy, completeness and latency.
- Benchmarked Mistral 7B, Llama 3 and Gemini 1.0 Pro on Vertex AI against a standard set of business prompts.
- Automated metric collection and charts in Python, cutting evaluation time from 2 weeks to under 4 hours.

Architect & Senior Developer | Deloitte (Internal: DataSAFE M&A Accelerator) **London, UK | Dec 2023 - Mar 2024**

- Built processing pipelines for 10TB+ of GPG-encrypted datasets with schema validation, integrity checks and automated decryption.
- Wrote Python tooling that finds and masks personal data (PII) across 1,000+ database columns using Fernet encryption.
- Built event-driven Azure Functions that unmask datasets only for authorised consumers.

Azure Data Engineer | Deloitte (Client: UBS AG) **London, UK | Jul 2023 - Aug 2023**

- Built migration and archival pipelines moving legacy Credit Suisse application data into UBS Azure across 1,000+ systems.
- Tripled ingestion throughput from PostgreSQL, Sybase and Oracle with Azure Functions and Azure Data Factory pipelines.
- Provisioned infrastructure with Terraform and set up CI/CD in GitLab and Azure DevOps.
- Designed file validation, GPG encryption and archival workflows that meet financial compliance and legal hold rules.

Senior Cloud Data Engineer | Deloitte (Client: Virgin Media O2) **London, UK | Jul 2022 - Apr 2023**

- Built pipelines loading HR records for 40,000+ employees from Oracle HCM, Workday and SuccessFactors into BigQuery.
- Protected the data with GPG encryption, role-based access and Python scripts for automated PII masking.
- Automated GPG key generation and rotation with shell scripts, cutting maintenance effort by 80%.

Data Engineer | Deloitte (Client: Motor Insurers' Bureau) **London, UK | Mar 2022 - May 2022**

- Built a data quality scoring platform on Azure for the MID, CUE and MIAFTR insurance databases (35M+ records).
- Provisioned Azure Data Factory, Azure SQL Server, storage accounts and Informatica Secure Agents with Terraform.
- Configured Informatica Cloud (IICS) profiling and validation rules that produce automated quality scorecards.

Senior Cloud & Data Engineer | Infosys (Client: HSBC) **London, UK / Pune, India | Jan 2015 - Nov 2021**

- CESOP reporting: built an AWS to GCP migration proof of concept for payment reporting, with Terraform for Dataproc and Cloud Functions.
- Streaming: built pipelines moving 10M+ messages a day from Apache Kafka to BigQuery through GCS with Apache NiFi.
- Change data capture: set up IBM CDC (DB2) and Attunity Replicate (Oracle) through Kafka Connect and AWS Lambda for transfers between AWS and GCP.
- Legacy migration: led transformation pipelines in IBM DataStage, Hadoop, Hive, Sqoop and Oracle for the Common Data Repository migration.

OPEN-SOURCE PROJECTS

Homelab Anomaly Detection (auto-datapulse + homelab-mlops) | siddharthsbhadauria.github.io/homelab-mlops/

- Self-hosted on a home NAS: telemetry every 15 minutes into DuckDB and dbt, anomaly models retrained every 6 hours with MLflow, scores served from FastAPI and alerts raised as GitHub Issues.

Cloud-FinOps-RAG | siddharthsbhadauria.github.io/cloud-finops-rag/

- Daily GitHub Actions pipeline that pulls prices for 60 LLMs across 8 providers and ranks them in DuckDB by benchmark score per dollar.

TF-Cost-Governor | siddharthsbhadauria.github.io/tf-cost-governor/

- Estimates monthly cost and security guardrails for AWS, GCP and Azure Terraform blueprints before apply, totalled in DuckDB daily.

Awesome-AI-Discoveries | siddharthsbhadauria.github.io/awesome-ai-discoveries/

- Python bot that logs a new agent, RAG or MCP repository three times a day to a searchable directory (270+ repos).

CERTIFICATIONS

- **Google Cloud Certified: Professional Machine Learning Engineer** (Dec 2025 - Dec 2027)
- **Google Cloud Certified: Professional Data Engineer** (Dec 2025 - Dec 2027)
- **Google Cloud Certified: Generative AI Leader** (Jan 2026 - Jan 2029)
- **Google Cloud Certified: Associate Data Practitioner** (Jan 2026 - Jan 2029)
- **Google Cloud Certified: Cloud Digital Leader** (Feb 2024 - Feb 2027)
- **Google Cloud Partner Specialist: Gemini Enterprise Agent Development** (Nov 2026 - May 2027)
- **Google Cloud Partner Specialist: Gemini Enterprise Deployment** (Nov 2026 - May 2027)
- **AWS Certified AI Practitioner** (Sep 2024 - Sep 2027)
- **HashiCorp Certified: Terraform Associate (003)** (Sep 2024 - Sep 2026)
- **AWS Certified Cloud Practitioner** (Jul 2022 - Feb 2028)
- **Microsoft Certified: Azure Fundamentals (AZ-900), Data Fundamentals (DP-900) and Azure AI Fundamentals (AI-900)** (2024)
- **Oracle Cloud Infrastructure: Certified Foundations Associate** (2024)

EDUCATION

Bachelor of Engineering in Electronics and Communication | RGPV University, India (2010 - 2014)