KUMAR CHANDRACHOODA

AWS Data Architect | S3 Lakehouse with Kinesis, Glue (PySpark) and Lake Formation | Python & SQL | 3x AWS Certified
+44 7767 964982 | kumarchandrachooda@gmail.com | kumarchandrachooda.com | linkedin.com/in/kumar-chandrachooda
Right to work: UK Skilled Worker visa valid until 11 February 2027, I would need a certificate of sponsorship in future.

SUMMARY

I have hands on experience in designing and architecting data and cloud platforms that other teams and agents can rely on. Over 8 years I've moved from .NET microservices and their data stores on Azure to leading AWS cloud and data architecture at PebblePad.

I architected the analytical data product on the medallion (bronze, silver, gold) lakehouse, from Amazon Kinesis stream ingestion through ETL in AWS Glue (PySpark) to Apache Iceberg tables on Amazon S3 Tables with AWS Lake Formation row level tenant isolation, all defined in AWS CDK and released through CI/CD. It turns 100 TB+ of legacy data (RDS PostgreSQL and self hosted MongoDB) into centralised, self service analytics so as to cut typical report time from 3-5 minutes to under 30 seconds.

Governance is built in, with GDPR data lifecycle management covering tombstoning, masking, audit trails and retention driven removal across subsystems. Using spec driven, multi agent workflows with Claude Code, I delivered the data services 3-5x faster than the team's previous hand built delivery. I coordinate technical work, resources and teams across the business and keep non technical stakeholders accurately informed.

EXPERIENCE

Senior Software Engineer, PebblePad Sep 2023 - Nov 2026

- Architected the AWS native analytical data product on the medallion architecture (Kinesis ingestion, AWS Glue (Spark) ETL, Amazon S3 Tables on Apache Iceberg, Lake Formation row level tenant isolation, all modularised in AWS CDK).
- Delivered the analytics service in six months, turning 100 TB+ of legacy data (RDS PostgreSQL and self hosted MongoDB) into an AWS native, AI first analytics offering under UK GDPR compliance.
- Recent records serve from an Amazon Redshift gold layer, and anything older than a year is archived to S3 Tables to keep the cost of dynamic reporting down while keeping it under 30 seconds, down from 3 minutes.
- I took it through Head of Platform and InfoSec approval and also trained 2 senior and 3 junior developers, and mentored 5 engineers and 2 apprentices while modernising a monolith into a .NET Core modular architecture.
- Architected the production change data capture pipeline that feeds the lakehouse. Three independent Debezium Server instances on ECS Fargate capture RDS PostgreSQL and self hosted MongoDB (self signed TLS) into Amazon Kinesis, with EFS backed offsets, JMX Prometheus metrics, Quarkus health endpoints, single task scheduling against dual slot contention and a Kinesis → Firehose → S3 staging zone → Glue landing pattern.
- The Glue ETL moves 84 interdependent tables per database through extract → tenant enhancer → transform → load into S3 Tables, with the tenant isolation as the namespace, and Glue Workflows trigger each stage from bronze through silver to gold.
- Architected and built the C# and .NET 10 control plane for the multi tenant analytics product, with Clean Architecture, MediatR CQRS, EF Core for writes and Dapper for reads. One EF Core migration set provisions every tenant under database, schema or shared database isolation, with NUnit tests written first for authorisation and tenant isolation.
- Delivered the platform's GDPR controls. Tombstoning lets a person be forgotten without losing audit trails, data is masked as it is removed, and a storage limitation job on ECS Fargate removes data past its retention limit across every subsystem, an event driven compliance tool with Lambda checks and EventBridge transport.
- Led agentic development with agent tools and skills that carry the platform's infrastructure guidelines and deployment process, and multi agent workflows from specification through planning, development and testing, delivering data services 3-5x faster than the team's previous hand built delivery.
- Lead implementation engineer for the CI/CD and plug and deploy Infrastructure as Code modules in AWS CDK with C#, now used by nearly 20 services. Set the build and deploy standards and put infrastructure under central human approval.
- Lead engineer on the security first, zero trust API framework, written in C# and .NET 10, that every new AWS web service starts from, cutting new service setup from 3 weeks to 2 days. Services run on Lambda or ECS behind API Gateway, secured with Amazon Cognito and PKCE, Lambda and JWT authorizers, WAF and KMS, and monitored through CloudTrail, GuardDuty and Security Hub.

Senior Product and Platform Engineer, Accenture Jul 2021 - Aug 2022

- Architected and built microservice platforms on Azure with their data pipelines and stores, delivered through a reusable Terraform library (App Service, Functions, APIM, AKS, Cosmos DB, Event Hubs, Front Door, SQL Managed Instance) scaled to a hub and spoke model across regions and environments, with CI/CD on Azure DevOps. Achieved Azure Solutions Architect Expert and DevOps Engineer Expert here.

Senior Software Engineer, Imaginea Technologies (now part of Accenture) Dec 2017 - Jun 2021

- Modernised legacy monoliths into .NET Core microservices (APIs, data models, integrations, analytics and reporting) and mentored 11 engineers and 5 apprentices, earning MCSA Web Applications and MCSD App Builder along the way. Ranked in the top 5% of Imaginea's developers by my reporting manager, with client work that led to new business wins.

Programmer Analyst Trainee, Cognizant Feb 2017 - Nov 2017

- Built ASP.NET Web Forms and MVC applications and proofs of concept, and earned the Microsoft Technology Associate (MTA).

DATA & PLATFORM PROJECTS

Multi Tenant Analytical Data Platform (Architect and Developer) The analytical data product on the medallion architecture, with AWS native bronze and silver layers (Kinesis, AWS Glue (PySpark), Apache Iceberg on Amazon S3 Tables) and a gold layer on Amazon Redshift and multi tenant PostgreSQL fed through the API framework.
Stack: PostgreSQL, C#, .NET 10, EF Core, Dapper, MediatR, Amazon Kinesis, AWS Glue, Amazon S3 Tables, Debezium, Redshift, Athena, Glue Workflows

KC Star (Architect) An inhouse analytical data platform on PostgreSQL, built on a row based star schema where new dimensions and measures need no DDL, SCD2 history with SHA-256 version hashes and time travel, data quality scores and lineage on every row, pg_cron batch processing with partitioning and monitoring, dependency driven recalculation and permission aware measures.
Stack: PostgreSQL, PL/pgSQL, pgcrypto, pg_cron, Docker

EARLIER PROJECTS: INSURANCE AND ENTERPRISE PLATFORMS

- **Insurance document platforms (Document Generator, Smart Comm):** Solution Architect and Lead Developer. Set up the Master Data Management (MDM) and Operational Data Store (ODS) for policy document generation with InRule validation, event driven on Azure Functions, .NET 6, Terraform (Azure) and API Management.
- **Insurance quoting and agency systems (Affinity, FirstRate, Inzurezone):** Team Lead and Assistant Architect on the migration from AS400 and web forms to microservices with single sign on across platforms. Mentored the team and landed 2 large migration projects.

EDUCATION

- MSc Advanced Computer Science (Cloud Computing), University of Leeds, Sep 2022 - Sep 2023, with Distinction
- B. Tech Computer Science Engineering, JSS Academy of Technical Education, 2016

CERTIFICATIONS

- AWS Certified Machine Learning Engineer - Associate, AWS Certified Data Engineer - Associate, AWS Certified Solutions Architect - Associate
- Microsoft Certified: Azure Solutions Architect Expert, DevOps Engineer Expert and Azure Developer Associate. MCSA Web Applications. MCSD App Builder (Programming in C#, Querying Data with T-SQL)

SKILLS

- **Data engineering:** ETL pipelines, streaming and batch processing, data modelling, data warehousing, medallion lakehouse (bronze, silver, gold), star schema, slowly changing dimensions (SCD2), data quality (cleansing and conformance), data lineage, Apache Spark (PySpark), Apache Iceberg, change data capture (CDC, Debezium), self service analytics, Master Data Management (MDM), Operational Data Store (ODS)
- **AWS data & analytics:** Amazon Kinesis, AWS Glue, Amazon S3, Amazon S3 Tables, AWS Lake Formation, Amazon Redshift, Amazon RDS, Amazon DynamoDB, AWS Lambda, ECS Fargate, EventBridge, SQS, IAM, KMS, Secrets Manager
- **Databases & storage:** PostgreSQL (ltree, jsonb), SQL Server (T-SQL), MongoDB, DynamoDB, Cosmos DB, OpenSearch, ClickHouse, Valkey/Redis
- **Data governance & security:** data governance, zero trust, multi tenant isolation, Lake Formation row level security, data lifecycle management, GDPR (right to erasure, tombstoning, data masking, audit trails, retention), ISO 27001
- **Infrastructure & DataOps:** Infrastructure as Code (AWS CDK / CloudFormation, Terraform on Azure), reusable infrastructure modules, CI/CD pipelines (Azure DevOps, GitHub), Git, Docker, Kubernetes (AKS), event driven and serverless architecture, OpenTelemetry observability
- **Languages, frameworks & AI:** Python, SQL, T-SQL, C#, TypeScript and PowerShell. .NET 10, ASP.NET Core Minimal APIs, EF Core, Dapper, NUnit and MassTransit. Amazon SageMaker, Amazon Bedrock, RAG, vector search, Claude Code

