

Ihor Ts

Senior Senior Data and ETL Engineer

SUMMARY

- Data engineer with 18+ years in distributed systems and data-intensive platforms, now focused on cloud ETL/ELT (Extract-Transform-Load and Extract-Load-Transform) pipelines and lakehouse architecture.
- Recent work: rebuilding a legacy on-premise SQL/shell/Python pipeline estate as medallion-layered Databricks/Delta Lake pipelines for a global investment bank, and building AWS-native batch and streaming pipelines (EMR/Spark, S3, Lambda, Step Functions, Kinesis) for a global financial market-data provider.
- Strong SQL and relational database design; dimensional modelling and data warehousing; data quality, lineage and governance in strictly regulated environments.
- EU-resident in Sofia.

TECHNICAL SKILLS

Main Technical Skills	Python (5 yr.), SQL (5 yr.), Databricks (1 yr.), AWS (5 yr.)
Programming Languages	Groovy, Java (10 yr.), JavaScript, Python (5 yr.), TypeScript (1 yr.)
Java Frameworks	Apache Spark, Spring, Spring Boot (3 yr.), Struts 2 (5 yr.)
Scala Frameworks	Apache Spark
AI & Machine Learning	AWS SageMaker, AWS SageMaker (Amazon SageMaker), NLP, PandasAI (2 yr.)
Python Libraries and Tools	Gensim, Pandas, PySpark (2 yr.), pytest
Salesforce Ecosystem	Google Ads
Social Media Marketing	Google Ads
Data Analysis and Visualization Technologies	ADLS, Apache Airflow, Apache Spark, Databricks (1 yr.), Data Lineage, Data Quality, ETL, Lakehouse, Pandas, PandasAI (2 yr.), Presto S3 (2 yr.)
Databases & Management Systems / ORM	Apache Spark, AWS DynamoDB (2 yr.), Cassandra (2 yr.), CouchDB, Database Design, Data Lake, Data Warehousing, Microsoft SQL Server (1 yr.), MongoDB (2 yr.), MySQL (1 yr.), NoSQL, Oracle Database (2 yr.), PostgreSQL (5 yr.), Query, Redis (2 yr.), SQL (5 yr.), SSIS
Cloud Platforms, Services & Computing	AWS (5 yr.), Azure, Serverless
Amazon Web Services	Amazon EMR (2 yr.), AWS Auto Scaling (1 yr.), AWS Cloudformation (2 yr.), AWS DynamoDB (2 yr.), AWS EC2 (2 yr.), AWS ECS (Amazon Elastic Container Service), AWS



	Elastic Kubernetes Service (EKS) (2 yr.), AWS EMR, AWS Kinesis (2 yr.), AWS Lambda (2 yr.), AWS S3, AWS SageMaker, AWS SageMaker (Amazon SageMaker), AWS SNS (2 yr.), AWS WorkSpaces
Azure Cloud Services	Azure Cloud Functions (2 yr.), Azure Databricks, Databricks (1 yr.)
Industry Domain Experience	Advertising, Banking, Enterprise, Fintech, retail, Telecom
UI/UX/Wireframing	3D Modelling (1 yr.)
Deployment, CI/CD & Administration	Active Directory, CI/CD (2 yr.), DevOps, GitLab CI, istio (1 yr.), Jenkins, New Relic
Web/App Servers, Middleware	Apache HTTP Server
Scripting and Command Line Interfaces	Bash (5 yr.), Shell Scripts
Logging and Monitoring	Datadog (2 yr.)
Methodologies, Paradigms and Patterns	Distributed systems, microservices, Refactoring, REST
Virtualization, Containers and Orchestration	Docker (2 yr.), Kubernetes (1 yr.), Terraform (2 yr.)
Version Control	Git (5 yr.)
QA, Test Automation, Security	Integration testing, pytest
Platforms	IOT, Unity
Message/Queue/Task Brokers	Kafka (1 yr.), RabbitMQ (2 yr.)
Third Party Tools / IDEs / SDK / Services	Parcel, WebStorm
Operating Systems	Red Hat Enterprise Linux, RHEL (Red Hat Enterprise Linux)
Other Technical Skills	CDC, Databricks Notebooks, Data Platform, Data Processing, Delta lake (1 yr.), ETL/ELT, Monitoring, Performance Tuning, Reactive, Spring WebFlux, Step functions, test coverage, Word2Vec, Zipkin

KEY STRENGTHS – CLOUD ETL/ELT & DATA PLATFORM

- **Enterprise ETL/ELT at scale:** designed and operated batch and streaming pipelines moving millions of records per week from operational source systems into curated analytical datasets – landing/staging/curated (bronze/silver/gold) layering, incremental and idempotent loads (MERGE), data cleansing, profiling, reconciliation and quality gates.
- **Dual-cloud delivery:** Azure – Databricks, Data Lake Storage Gen2, Delta Lake, Unity Catalog; AWS – S3, EMR (Spark), Lambda, Step Functions, Kinesis, SQS/SNS,



DynamoDB, SageMaker, CloudFormation. Comfortable mapping equivalent services across both stacks.

- **Legacy-to-cloud migration:** migrated an undocumented on-premise ETL estate (Autosys, shell, SQL, Python) into cloud pipelines with parallel-run reconciliation, controlled cutover and zero disruption to downstream consumers – the same discipline a database-migration or SSIS-modernisation programme requires.
- **Structured, semi-structured and unstructured data:** relational extracts (Oracle, SQL Server, PostgreSQL), XML/JSON/CSV landing files, Parquet/Delta storage, large free-text corpora for NLP feature generation, and binary media assets streamed through Kafka into a corporate data lake.
- **SQL and data modeling:** advanced SQL (window functions, CTEs, execution-plan tuning), relational schema design, dimensional modeling – star schemas, grain declaration, fact table types, SCD (Slowly Changing Dimension) Type 2 – and curated marts serving BI (Business Intelligence) and data-science consumers.
- **Reliability and performance:** Spark job tuning (partitioning, skew handling, broadcast joins, small-file compaction), pipeline monitoring and alerting, SLA-bound production support and on-call ownership on regulated and peak-load systems.
- **Governance and security:** auditability and lineage for every derived field, PII (Personally Identifiable Information) handling and masking, access control, and change management inside a banking compliance regime.

WORK EXPERIENCE

Senior Data Engineer - EPAM Systems

UBS (global investment bank), Compliance analytics platform)

Duration: Oct 2025 – Jun 2026

Summary: Bank-wide e-mail and instant-messaging metadata transformed into conversation-level features consumed by finance analytics teams in a strictly regulated environment with full auditability, change control and data-quality gates.

Responsibilities:

- Re-architected legacy shell/SQL/Python batch ETL into Databricks (PySpark) pipelines over a bronze/silver/gold Delta Lake design focusing on lineage, observability and testability.
- Built curated/gold layer datasets by joining pipeline output with Oracle reference data for downstream reporting and modelling.
- Processed weekly batches of millions of communications metadata records with validation and reconciliation checks for banking audit trail.
- Migrated Autosys-orchestrated pipeline estate from RHEL 7 to RHEL 8 with controlled cutover and no disruption.
- Collaborated with analytics and compliance stakeholders to translate business requirements into pipeline logic and data contracts.

Technologies: Azure Databricks, Delta Lake, Apache Spark (PySpark), Python, SQL, Oracle, Bash, Autosys, RHEL 7/8, Git

Senior Software Engineer - EPAM, insurance client (Customer portal backend)

Duration: Mar 2025 – Sep 2025



Summary: Development of backend REST services and third-party system integrations for a customer portal in the insurance domain.

Responsibilities:

- Developed REST services and integrated third-party systems using Java and Spring Boot.

Technologies: Java, Spring Boot

Senior Software Engineer - Hitachi and Disney Advertising (AdTools microservice platform)

Duration: Nov 2023 – Oct 2024

Summary: Microservice platform serving ads across Disney's linear, digital and streaming properties integrating external providers such as Google Ads.

Responsibilities:

- Built event-driven services ingesting ad-delivery event streams via Amazon Kinesis for reporting and revenue reconciliation.
- Developed and integrated microservices on Kubernetes/Istio and implemented Google Ads integration.
- Managed infrastructure with Terraform and ensured code quality through reviews, refactoring, and integration test coverage.
- Monitored production observability using Datadog and Zipkin.

Technologies: AWS (Kinesis, DynamoDB), Java 11–17, Spring Boot, Kubernetes, Istio, Docker, Terraform, Datadog

Senior Software Engineer - client: Tesco (Legacy microservices migration)

Duration: Dec 2022 – May 2023

Summary: Migration of legacy microservices to a reactive stack for one of the world's largest grocery retailers.

Responsibilities:

- Migrated legacy microservices to Spring WebFlux reactive stack.
- Maintained MongoDB-backed services and performed integration testing.

Technologies: Spring WebFlux, MongoDB

Senior Software Engineer - Evri (Hermes), Parcel-logistics microservices

Duration: Jan 2021 – Dec 2022

Summary: Development and support of high-throughput parcel-logistics microservices on AWS ECS, handling peak-season production loads.

Responsibilities:

- Built and supported microservices using Spring Boot, RabbitMQ, Cassandra, and Redis.
- Provided production support under peak-season load conditions.

Technologies: Spring Boot, RabbitMQ, Cassandra, Redis, AWS ECS



Senior Software Engineer - Refinitiv (LSEG) , Content Source Directory

Duration: May 2020 – Dec 2020

Summary: Self-service management platform for news and content sources for a global financial market-data provider.

Responsibilities:

- Designed and implemented serverless services on AWS Lambda with SQS/SNS messaging.
- Contributed to component architecture design and coordinated team work.

Technologies: AWS Lambda, SQS/SNS, Java 11, MySQL

Senior Software Engineer (Data/ML pipelines) - Refinitiv (LSEG), NLP data pipelines for financial news and content intelligence.

Duration: Dec 2018 – Mar 2020

Summary: Development of NLP data pipelines over large text corpora to support financial news and content intelligence using AWS cloud services.

Responsibilities:

- Migrated on-premise data processing workloads to AWS Lambda and Step Functions with infrastructure as code using CloudFormation.
- Built EMR Spark jobs for preparing and transforming large unstructured text corpora for model training.
- Replaced Gensim Word2Vec implementation with Amazon SageMaker.

Technologies: AWS EMR, SageMaker, Lambda, Step Functions, CloudFormation, S3, Java 11, Python 3

Senior Software Engineer - Adidas (Kafka-decoupled microservices for media asset ingestion)

Duration: 2018

Summary: Built microservices decoupled by Kafka to ingest and store media assets and associated metadata into the corporate data lake.

Responsibilities:

- Developed Kafka-based microservices for large-scale distributed file system ingestion and storage of media assets and metadata.

Technologies: Kafka

Senior Software Engineer / Team Lead. Schad (Industrial IoT), Predictive and reactive maintenance analytics

Duration: Nov 2016 – Dec 2017

Summary: Led development of predictive and reactive maintenance analytics for industrial facilities using stream and batch processing of sensor data.



Responsibilities:

- Led a team of 4 engineers building analytics solutions using Apache Spark and Apache Storm.
- Handled stream and batch processing of sensor data for industrial IoT applications.

Technologies: Apache Spark, Apache Storm, Java, Spring, CouchDB, MS SQL Server

Senior Java Developer / Team Lead - easyfinancial (Canada) (Online consumer-lending platform)

Duration: Sep 2013 – Oct 2016

Summary: Development and leadership on a regulated online consumer-lending platform including prequalification, credit adjudication, document verification, e-signing and payments.

Responsibilities:

- Led a team of 4 engineers on a regulated financial product.
- Owned production monitoring and 24/7 support.
- Developed server-side lending workflow functionality using Spring and PostgreSQL.

Technologies: Java, Spring, PostgreSQL

EDUCATION

- **Lviv Polytechnic National University**
B.Sc., Computer Systems Automation and Metrology
2004 – 2008

