MARKO VUKOVIC

+1-647-231-1120 | mail@mvk.vc | site.mvk.vc

SKILLS

Languages: Elixir, Python, TypeScript, Ruby, Rust, SQL
Libraries: Phoenix, LiveView, FastAPI, Pandas, PySpark, React, Rails, pgvector
Infrastructure: Azure, GCP, Databricks, Fly.io, Supabase, RabbitMQ, Postgres, Docker

PROJECTS

Replicant

Crowd-sourced LLM inference server with an OpenAI-compatible API.
Technologies: Elixir, Phoenix, LiveView, TypeScript, NodeJS, Electron, React, RabbitMQ, Postgres

Rager Cloud

Logging service for the Rager library used to build workflows with generative models.
Technologies: Elixir, Phoenix, LiveView, TypeScript, D3, Postgres

EXPERIENCE

Data Engineer — **Remote**
Authentica Solutions — *Aug 2024 – Aug 2025*

- Developed and automated data pipelines in Azure Databricks using Python, Pandas, and PySpark to ingest, normalize, and export education data between external systems, along with verification tests and deployment automation using the Databricks SDK.
- Built a new FastAPI backend service to be the primary data source for the main frontend application, enabling secure querying and aggregation of Databricks data.
- Crafted internal tools and POCs, including an Elixir and Phoenix LiveView prototype to validate new workflows and APIs used for migration to Databricks.

Full-Stack Engineer — **Remote**
Admintel — *Aug 2023 – Jan 2024*

- Led development of an AI-driven legal research tool using Phoenix LiveView deployed on Fly.io, with RDS Postgres and pgvector for embeddings, Oban for background processing, and S3 for document storage.
- Built and deployed a FastAPI RAG service using LangChain and the OpenAI API, enabling real-time chat and structured generation from documents to power features such as custom flashcards.
- Identified the need for and implemented monitoring for RAG workflows using Langfuse and created real-time analytics dashboards with Elixir LiveBook to track user activity and application metrics on Fly.io.

Data Scientist — **Toronto, Canada**
Royal Bank of Canada — *July 2019 – September 2021*

- Improved data quality by building a TensorFlow embedding pipeline in Python to cluster and flag duplicate IT risk issues, removing 200+ duplicates upfront and reducing ongoing manual review efforts.
- Unlocked insights from 18,000+ unstructured IT risk issues by building an XGBoost model in Python for multi-class text and metadata classification, enabling automated reporting by category.
- Automated risk scoring for 300+ monthly IT risk assessments by developing a CatBoost regression model in Python, improving efficiency and consistency in impact and likelihood evaluations.

EDUCATION

Dockyard Academy — **Remote**
Full-stack development with Elixir — *Completed July 2023*

Toronto Metropolitan University — **Toronto, Canada**
BSc. (Honours) – Financial Mathematics, Economics Minor — *Completed May 2019*