

Yagmur Gulec

Full Stack Data Engineer | Towards Data Science Contributor

Sherbrooke, Québec, Canada | yagmur.gulec@outlook.com | [Portfolio](#) | [GitHub](#) | [LinkedIn](#)
+1 819 588 75 93 | Permanent Resident of Canada

Professional Summary

Full-stack developer and data engineer with experience building cloud-native pipelines, containerized applications, and applied machine learning models. Skilled in AWS serverless architectures, Terraform, and computer vision (YOLO-based detection). Passionate about making complex technical topics accessible—regular contributor to Towards Data Science, where I publish tutorials on data engineering and ML workflows.

Key Technical Skills

AI/ML: Python, PyTorch, computer vision (object detection)

Cloud: AWS Serverless (Lambda, API Gateway), Terraform, Docker, CI/CD

Programming & Tools: Python, Java, Bash, Jupyter, Asyncio, Pandas

Frontend: React

Experience

Machine Learning Developer Intern

M2M, Remote

May 2025 – Jun 2025

- Enhanced YOLO-based wildfire detection with domain-specific data and augmentation pipelines.
- Applied image processing techniques for data augmentation.

Software Developer

Riipen Beyond the Cloud and Level UP, Remote

Feb 2024 – Mar 2025

- Developed a prediction market data dashboard and deployed to AWS Lambda (Python, Dash, SAM).
- Automated provisioning of a Jenkins server with Terraform on AWS EC2 behind Nginx.
- See feedback & projects: [Riipen Profile](#).

P.h.D Researcher (Mechanical Engineering)

Université de Sherbrooke, Québec

Feb 2019 – Sep 2021

- Built object-oriented C++ models for vapor bubble dynamics in boiling liquid.
- Automated large-scale simulations on HPC (Calcul Québec) with Bash/Linux.
- Implemented ImageJ Java plugin to measure bubble sizes from experimental images.

Personal Projects

Serverless Job Queue for SQL-Based Label Statistics and Dataset Partitioning

[GitHub](#) | [Website](#)

- Designed and deployed a serverless data-processing pipeline for object detection dataset analytics (Pascal VOC).
- Enabled automated stratified splits of 20k+ annotated images for ML training

- **Use Case: Balanced dataset generation for computer vision models.**
- **API Gateway + Lambda** to receive SQL job requests and insert status into DynamoDB.
- SQS to queue jobs asynchronously.
- Worker Lambda to process SQL queries using Amazon Athena, compute stratified train/validation splits, and update job status.
- Infrastructure provisioned with **Terraform**
- Stack: AWS Lambda, SQS, DynamoDB, Athena, API Gateway, Terraform, Python, Athena, Glue Catalog, Docker

NOAA Climate Data Visualization and Analysis Pipeline

[GitHub](#) | [YouTube Demo](#)

- Delivered interactive geospatial visualizations with React/Deck.gl to analyze climate datasets
- Built an end-to-end pipeline from NOAA API to a geospatial full-stack web app.
- Parallel ingestion with rotating API keys (Bash, GNU Parallel); async cleaning (Asyncio, Pandas).
- RESTful backend (Spring Boot, Java) serving GeoJSON; frontend (React, TypeScript, Deck.gl).
- Containerized frontend, backend, and database (Docker).

Publications

How to Import Pre-Annotated Data into Label Studio and Run the Full Stack with Docker Towards Data Science

August 29, 2025

- Tutorial about how to import pre-annotated datasets into Label Studio

End-to-End AWS RDS Setup with Bastion Host Using Terraform

[GitHub](#) | [Towards Data Science](#)

Jul 28, 2025

- Automated secure MySQL RDS deployment with a bastion host using modular Terraform.

Clustering Eating Behaviors in Time: A Machine Learning Approach to Preventive Health

[GitHub](#) | [Towards Data Science](#)

May 8, 2025

- Implemented Modified Dynamic Time Warping (MDTW) for temporal dietary pattern clustering (Python).
- Added CI with GitHub Actions and Pytest for robust testing.

Education

Bishop's University, Sherbrooke, QC — Master's in Computer Science

Jan 2022 – May 2024

GPA: 3.9/4.0 Courses: Data Mining, Database Design (SQL, PostgreSQL), Deep Learning

Language Skills

English: Full professional proficiency | **French:** Professional proficiency