

Yi Xin

Phone: 249-874-5096 **Email:** jonzy.xin@outlook.com **Location:** Niagara Falls, ON, Canada **Portfolio:** <https://www.niagaradataanalyst.com> **LinkedIn:** <https://www.linkedin.com/in/osmond-xin-92a736308> **GitHub:** <https://github.com/Osmond-Xin>

Work Authorization: Master of Data Analytics completed in Canada; PGWP-eligible with ~3 years of open work authorization (dates to be confirmed). No LMIA or employer sponsorship required. Available for a full-time start from August 2026.

PROFESSIONAL SUMMARY

Data and AI-application engineer who builds production systems end to end — from data pipelines and AWS cloud infrastructure to applied LLM/agent workflows and the dashboards that make results usable. Recent hands-on work includes a ~20-node LangGraph agent, a LangChain RAG pipeline, an IIoT predictive-maintenance backend, and an award-winning geospatial analysis of 275,156 trees — built alongside a Master of Data Analytics (GPA 4.13/4.3) in Canada. Earlier engineering and architecture work, including an AWS solution featured as an official case study, is the foundation these projects build on.

PROJECTS

job-hunt — Open-source LangGraph Job-Search & Application System

Personal Project (open-source) | GitHub: <https://github.com/Osmond-Xin/job-hunt> | 2026 — ongoing

Three cooperating subsystems that take a job posting from discovery to a reviewed, ready-to-submit application, with a human in the loop before anything is sent.

- **Evaluation graph** — a LangGraph StateGraph (~20 typed nodes) that extracts a job description, gates it, scores fit across weighted dimensions, and generates a tailored CV, cover letter, and report; LLM calls run across the Claude and OpenAI APIs with structured outputs and graceful fallback.
- **Discovery service** — scans direct ATS APIs (Greenhouse / Lever / Ashby) and the Brave WebSearch API, de-duplicated against a local tracker.
- **Application assistant** — Playwright automation for Workday and LinkedIn Easy Apply with a two-step, fill-only human-in-the-loop flow (never auto-submits behind unmet gates), coordinated via file-lock IPC sentinels and a structured JSONL event log.
- Covered by 460+ pytest tests; a single mode: `student` | `full` switch cascades across discovery, scoring weights, and narrative.

Tools: Python, LangGraph, LangChain, Playwright, Claude API, OpenAI API, Brave WebSearch API, pytest

LangChain RAG Pipeline

Personal Project | GitHub: <https://github.com/Osmond-Xin/LangChainRAG> | April 2026

Retrieval-Augmented Generation demo built with LangChain, showing end-to-end document ingestion, vector retrieval, and source-attributed answer generation over a real corpus.

- Designed a full RAG pipeline: document loading → chunking → embedding → vector store → retrieval → LLM synthesis
- Integrated FAISS and Chroma vector backends for semantic similarity search over document corpora
- Orchestrated retrieval chains with LangChain and the OpenAI / Claude APIs, surfacing source attribution back to the originating passages
- Built for extensibility: modular loaders, configurable chunking strategies, swappable vector stores

Tools: Python, LangChain, FAISS, Chroma, OpenAI API, Claude API, Vector Embeddings

Matrix-Sync — IIoT Predictive-Maintenance Backend for Industrial Pumps

Personal Project | GitHub: <https://github.com/Osmond-Xin/industrialmatrix> | 2026 — ongoing

End-to-end industrial-IIoT backend that ingests high-frequency sensor telemetry from a mining centrifugal pump, persists it to a time-series store, and is architected for a closed-loop predictive-maintenance feedback flow (edge sampling → feature extraction → cloud ingestion → anomaly detection → adaptive baseline).

- Built a Go edge agent that publishes deterministic pump telemetry over MQTT (Mosquitto), with a Sparkplug B upgrade path (NBIRTH/DBIRTH device metadata, bbolt-backed offline replay)
- Implemented a TypeScript/Node.js ingestion service that subscribes to MQTT topics and writes to PostgreSQL/TimescaleDB with idempotent migrations and structured pino logging
- Authored a 7-ADR design pack covering broker choice (Mosquitto vs EMQX), TimescaleDB hypertable layout, protocol tradeoffs (JSON → Sparkplug B), and MVP scope
- Reproducible local stack via Docker Compose (Mosquitto + TimescaleDB) and a Makefile

Tools: Go, TypeScript, Node.js, Python, PostgreSQL/TimescaleDB, MQTT, Sparkplug B (Protobuf), Docker Compose

Climate Vulnerability Assessment & Urban Forest Replanting Strategy

H.E.A.D. Competition 2026 — 1st Place | Mohawk College & McKeil School of Business | March 2026

End-to-end geospatial data analysis assessing climate-change impacts on Hamilton's urban forest, producing science-based replanting recommendations across 15 municipal wards.

- Processed and taxonomically matched **275,156 city-owned trees** via the GBIF Backbone Taxonomy API, achieving a **98.9% match rate**
- Built a full ETL pipeline in Python/Pandas; applied the IPCC AR5 statistical framework to project a **+4.47°C temperature increase by 2080**
- Identified Ward 3 as the highest-risk zone (vulnerability score: 1.423) and recommended 10 climate-resilient replacement species
- Delivered a 5-module interactive GIS dashboard and competition poster; invited to the **CityLAB Winter 2026 Project Showcase**

Tools: Python, Pandas, GBIF API, IPCC AR5 statistical modeling, GIS visualization, ETL pipeline

EXPERIENCE

Data & ML Engineer (Internship) — FindGrant

Jan 2026 – Mar 2026 | Remote, Toronto, ON, Canada

- Engineered a real-time stock-trading data pipeline integrating live market APIs, EMA/ADX momentum signals, and options-expiry factors, validated through a custom-built backtesting engine.

- Deployed an interactive Streamlit/Python dashboard monitoring live strategy signals and portfolio performance; the strategy returned a 10% average monthly ROI in backtesting.

Software Engineer & Technical Consultant (Freelance)

Oct 2022 – Dec 2024 | Remote, Beijing/China (*Clients include: Freedom Property Investors, and others*)

- Built end-to-end data solutions and RESTful APIs for diverse clients, accelerating data processing by 40%.
- Developed conversational bots on the OpenAI and Claude APIs, plus the data-collection and visualization platforms behind them.
- Processed large volumes of interaction data with Python (pandas, matplotlib) to extract business metrics and user-behavior patterns.
- Maintained a 92% client-satisfaction rate by aligning technical deliverables with business specifications and scalability requirements.

Platform Product Manager — Beijing Bailongma Yunxing Technology Co., Ltd.

Aug 2021 – Sep 2022 | Beijing, China

- Spearheaded the optimization of a ride-hailing SaaS fleet management system through comprehensive data analysis, boosting fleet operational efficiency by 25% and reducing operational costs by 30%.
- Automated core CRM, asset management, and financial workflows, decreasing processing time by 20% and improving overall user satisfaction to 90%.
- Conducted extensive user interviews and data analysis to identify key pain points; redesigned driver rating system using data-driven methods, improving rating accuracy and satisfaction.
- Executed comprehensive market research and directed cross-functional initiatives across 5+ teams, driving a 40% increase in platform adoption.

Technical Director — Iqidao (Beijing) Cultural Communication Co., Ltd.

Dec 2014 – May 2021 | Beijing, China

- Owned product design and system architecture end-to-end across the 7-year tenure — defined product requirements, user flows, and success metrics while architecting the underlying data models, service composition, and AWS cloud infrastructure.
- Designed and implemented a company-wide data collection, tracking, and visualization system, integrating multi-source data and developing interactive dashboards for stakeholders.
- Architected AWS cloud solutions for scalable data processing — **officially featured as a case study on the global AWS website**; sustained 99.9% system availability.
- Directed a 10+ member cross-functional engineering team, reducing development costs by 35% and improving operational efficiency by 50% through strategic technical planning.

EARLY CAREER (2005–2014)

- **App Development Director** | Trends Group | Dec 2013 – Dec 2014
- **Technical Cooperation Manager** | Pangu Cultural Communication | Mar 2013 – Dec 2013
- **Development Manager** | Entrepreneurship Startup | Mar 2012 – Feb 2013
- **Java Development Manager** | IZP (China) Network Technologies | Nov 2009 – Mar 2012
- **Senior Software Engineer** | Beijing Yisiteng Technologies Co., Ltd. | Jan 2008 – Nov 2009
- **Software Engineer** | Beijing Zhongxun High-Tech Co., Ltd. | May 2005 – Dec 2007

EDUCATION

Master of Data Analytics — University of Niagara Falls Canada *Jan 2025 – 2026 (coursework completed)* | *Niagara Falls, ON, Canada* | *GPA: 4.13/4.3*

- **President's Academic Distinction List** (Spring 2026) — university-wide honor for a term GPA of 4.00+ on a full course load, awarded by the President and Vice-Chancellor.
- Applied advanced machine learning models (Random Forest, XGBoost) to diverse datasets, extracting actionable business insights.
- Authored comprehensive technical reports and developed interactive visualization dashboards to translate complex data into clear stakeholder presentations.

Bachelor of Computer Science and Technology — Beijing University of Posts and Telecommunications *2001 – 2005* | *Beijing, China*

SKILLS

Data Science & Analytics: Python (Pandas, NumPy, Scikit-learn), SQL, Machine Learning (Random Forest, XGBoost), Quantitative Backtesting, Technical Analysis, Exploratory Data Analysis (EDA)

Data Visualization: Tableau, Streamlit, GIS Visualization, Interactive Dashboards

Software Engineering: Java/J2EE, Go, Node.js, React, JavaScript, RESTful APIs, Microservices Architecture

AI & LLM: LangChain, LangGraph, OpenAI API, Claude API, Agentic Pipelines, RAG (Retrieval-Augmented Generation), Vector Databases (FAISS, Chroma), LLM Orchestration, Prompt Engineering

Cloud & Database: AWS (Lambda, API Gateway, DynamoDB), PostgreSQL/TimescaleDB, MySQL, Oracle, Docker, Git, System Design

Product: Product Strategy, 0-to-1 Product Development, Cross-functional Collaboration, Agile/Scrum

CERTIFICATIONS

- AWS Certified Solutions Architect – Professional | 2021
- AWS Certified Solutions Architect – Associate | 2021
- PMP Certified (Project Management Professional) | 2014