Contact

yongpengfu0011@gmail.com

www.linkedin.com/in/yongpeng-fu
(LinkedIn)

github.com/YongpengFu
(Personal)

Top Skills

Cloud Computing

AI Engineering

Software Development

Languages

Mandarin Chinese (Native or
Bilingual)

English (Full Professional)

Certifications

Access Basics for Excel Users

Learning R (2013)

The Complete Hands-On
Introduction on Apache Airflow

WHMIS, Lab Safety, Biosafety

Dagster Essentials

Publications

Engineering allosteric interactions
between domains of periplasmic
trehalases

Determining the IgG concentrations
in bovine colostrum and calf sera
with a novel enzymatic assay

Development of a blood calcium test
for hypocalcemia diagnosis in dairy
cows

A Novel Cell-Penetrating Antibody
Fragment Inhibits the DNA Repair
Protein RAD51

YONGPENG FU

Staff Data/AI Engineer

Calgary, Alberta, Canada

Summary

I'm a Data & AI Engineer with expertise in building scalable
pipelines, intelligent systems, and agentic workflows that connect
data to actionable insight.

Skilled in Python, Java, Apache Spark, Kafka, Hadoop HDFS, AWS
S3, LangChain and MCP. I design reliable data architectures and
integrate LLM frameworks and AI orchestration tools to develop
context-aware agents and automation solutions.

I focus on writing clean, efficient, and well-documented code that
drives clarity, performance, and maintainability. With over 4 years
of experience, I bridge data engineering and AI system design to
deliver robust, production-ready solutions for modern data-driven
applications.

Experience

RBC Borealis
Staff Data/AI Engineer
July 2025 - Present (4 months)
Calgary, Alberta, Canada

1. Leading AI Engineering System for mortgage retention

2. Aiming for a 3% increase in mortgage retention and a projected $21M

revenue uplift within 3 years

RBC
2 years 3 months

Senior Data Engineer
September 2023 - July 2025 (1 year 11 months)
Calgary, Alberta, Canada

1. Transformed disorganized big data into a scalable and evolving system

2. Led API development for data ingestion

Developer, Coop

May 2023 - September 2023 (5 months)
Calgary, Alberta, Canada

1. Revamped legacy shell scripts, achieving a 1,800× performance boost
2. Automated cleanup and AWS S3 archival of EDL data, saving $100K in storage costs
3. Designed and implemented real-time event ingestion pipelines to minimize downtime

Cenovus Energy
Data Scientist, Intern
September 2022 - May 2023 (9 months)
Calgary, Alberta, Canada

1. Developed a Flask- and Dash-based web application to automate the company's capital forecasting process, achieving an 80% reduction in processing time and a 20% improvement in accuracy.
2. Built and deployed 2 automation tools for the financial operation team using GUI and web application
3. Developed unit testing for continuous deployment in the codebase
4. Created one PowerBI dashboard to monitor all employee function hierarchy
5. Technology used: Python, R, SQL, PyQt, Dash, PowerBI, GitLab, Heroku, Posit Rstudio

Canadian Tire Corporation
Data Scientist, Intern
May 2022 - August 2022 (4 months)
Alberta, Canada

1. Created algorithms to consolidate 8,000+ inventories across 340 stores each month, streamlining the process, reducing data errors by 90%, and delivering $500K incremental value.
Reduced the manual time from 4 hours to 5 minutes by developing a GUI tool (Tkinter)
2. Developed an algorithm to consolidate more than 8000 articles across 340 stores efficiently
3. Created a time-series model to predict the inventory selling curve
4. Collaborated with other 2 interns to centralize all automation tools in one GUI application
5. Technology used: Python, R, SQL, Tkinter, Pandas, Dask, Matplotlib, ARIMA (time series), GitHub

Data for Good YYC

Data Analyst, Volunteer
September 2021 - May 2022 (9 months)
Calgary, Alberta, Canada

• Helping new volunteers develop a strong commitment to community volunteering projects
• Supporting project teams' formation from finding the right members to initiating the project
• Actively contributing to the technical knowledge pool of programming, database, and statistics

University of Calgary
4 years 10 months

Bioinformatician / Research Assistant
August 2017 - May 2022 (4 years 10 months)
Calgary, Alberta, Canada

Project 1: Leading Scientist in Project of "Fast Evolving Trehalase based on 16srRNA Phylogeny"
• Built up pipeline from scratch for discovering evolutionarily fast-growing microbe based on 16srRNA phylogeny
• Utilized multiple databases like SILVA, NCBI, UniPort to cast a broad net search
• Accelerated the project execution by incorporating Supercomputer into the study via Linux system
• Standardized the study with thousands of lines of R code from Model selection, statistics to tree construction
• Identified several potential bacterial strains that otherwise would have taken years to isolate from a wet lab

Project 2: Built "point of care" Biosensor prototype based on commercial Glucometer to detect certain cow diseases.
• Developed new diagnostic methods for testing signals of infection in blood, milk or saliva samples
• Conducted extensive DNA manipulation and enzymatic assay for selecting fast-activity enzymes

Bioinformatician
January 2020 - April 2021 (1 year 4 months)
Calgary, Alberta, Canada

• Established a complete protocol for RNA-seq and Microbiome analysis using a great deal of Data Science techniques including:

1. Sequence Quality Control, Data Trimming, Filtering, Sequence Aligning, Gene Quantification, Differential Expression analysis using DESeq2 package in R for RNA-seq

2. and additional Sequence Dereplication, Model Fitting, OUT and Phylogenetic Tree construction, plot exploratory analysis, Supervised Learning, DADA2 package for Microbiome.

• For RNA-seq project, about 50 genes out of a total of 50k genes were first time identified to be strongly responsive to the gut infection in C57/BL6 mouse model after the challenge of Citrobacter rodentium

• For microbiome analysis, species population dynamics were illustrated in C57/BL6 mouse model before and after Citrobacter rodentium infection

University of Calgary Continuing Education
2 years 1 month

Trainee in Business Intelligence and Analytics
January 2019 - January 2020 (1 year 1 month)
Calgary, Alberta, Canada

• Teamed with domain experts to design interview questions, understand business needs, gather business requirements, and translate them into dimensional model both conceptually and logically

• Designed and built a physical normalized database and data warehouse for improved ingestion and processing

• Utilized ETL to create actionable reports in Power BI dashboard, saving 10 hours of manual work each week

Sql Server Database Administrator
January 2018 - January 2019 (1 year 1 month)
Calgary, Alberta, Canada

• Managed, monitored, and troubleshoot Microsoft SQL databases and servers

• Ran ad hoc queries for requested changes, updates, and modifications to database structure and data

• Used stored procedures, triggers, and views to provide structured data to stakeholders by combining millions of rows of data from 20 disparate data sources

University of Saskatchewan
Master Student Researcher
January 2013 - February 2017 (4 years 2 months)
Saskatchewan, Canada

• Plan and manage a 3-year project in studying the relationships between protein structure and function using phage display technology in conjunction with high-throughput screening and sequencing
• Look after all aspects of project development like initial design, planning, proposal development,
data collection, progress report writing, and collaboration with external partners
• Collaborated with multiple external Labs to analyze synthetic antibodies, monitor bacterial growth, and collect samples
• Conducted intensive literature review, QC, reported on academic research, and chaired weekly lab meetings

Chinese Academy of Sciences
Research Assistant
May 2011 - September 2011 (5 months)
Beijing City, China

• Performed HPLC to obtain electrophoregrams for each of three pharmaceutical drugs.
• Trained in analytical techniques like polyacrylamide gel electrophoresis, Western blotting
and Flow Cytometer/FCM.
• Kept accurate records for experimental progress and reported to the supervisor.

Education

University of Calgary
Master in Data Science, Data Science and Machine Learning · (September 2021 - September 2023)

University of Calgary
Certificate, Database Administrator · (2018 - 2019)

University of Saskatchewan
Master of Science (M.Sc.), Health Sciences · (2013 - 2017)

Beijing Institute of Technology
Bachelor's degree, Health Science · (2008 - 2012)