Last Update: **28 Oct 2025** chuxiaoyu.github.io

XIAOYU CHU

Amsterdam Zuid, the Netherlands | Willing to Relocate

(+31) 06-4528-2610 | chuxiaoyu123@gmail.com | github/chuxiaoyu | linkedin/chuxiaoyu

Summary

Ph.D. researcher, AI engineer, and data scientist with extensive experience in turning large-scale data into actionable insights through statistical analysis, machine learning (ML), and natural language processing (NLP). Skilled in time series analytics, LLM applications, and LLMOps.

Skills

Programming: Python, SQL
Big Data: MongoDB, PySpark, MySQL, Spark, Jupyter, NumPy, Pandas, Matplotlib, Seaborn
AI&ML: PyTorch, LangChain, Prompt Engineering, RAG, TensorFlow, Scikit-learn, Keras
DevOps: Git, GitHub, Linux, Cursor, Docker, AWS, GCP, SLURM, Prometheus

Education

Vrije Universiteit Amsterdam Feb 2022 – July 2026 (expected)
Ph.D. Candidate in Computer Science *Amsterdam, the Netherlands*

Beijing Normal University Sep 2018 – Dec 2021
MSc. in Information Science, GPA: 89.60/100.00 (Top 20%) *Beijing, China*

Lanzhou University Sep 2014 – Jun 2018
B.A. in Accounting, GPA: 4.29/5.00 (Top 15%) *Lanzhou, China*

Experience

Ph.D. Candidate in Computer System Analysis Mar 2022 - Present
Vrije Universiteit Amsterdam *Amsterdam, the Netherlands*

- Developed a data pipeline leveraging 6 major **LLM APIs (GPT, Claude, Gemini)** to automate the **extraction and classification of structured information** from over 3,000 cloud incident reports (including AWS, Azure, and GCP), enabling large-scale analysis of service failures. [GitHub]
- Developed an **automated monitoring framework** to track the availability of major LLM services (e.g., ChatGPT, Claude), analyzed over 500 incidents to identify critical failure-recovery patterns. [GitHub] [FAILS]
- Analyzed **26GB** of performance, failure, and power usage data from ML workloads in a Dutch **HPC datacenter** (SURFLisa), collected via **SLURM** and **Prometheus**. [GitHub] [PDF]
- Published 6 papers in high quality venues; presented 4 talks at international conferences and professional events.
- Mentored 10+ BSc. and MSc. theses, course, and honors projects, resulting in 4 publications. [Link]

NLP Research Engineer Intern Mar 2021 - May 2021
Orange Labs Beijing *Beijing, China*

- Prototyped a **deep structured semantic model** to improve the accuracy of personalized news recommendation.
- Optimized **MongoDB** operations and contributed to **API design** for data access, enhancing system usability.

Data Analyst Intern Sep 2020 - Dec 2020
51Talk Online Education *Beijing, China*

- Extracted and analyzed operational data via **SQL**, improving efficiency in data-driven decision-making.

Projects

Peer Review Texts Mining and Application for Literature Retrieval | *PyTorch, Elasticsearch* MSc Thesis

- Collected a structured Chinese peer review dataset and proposed a systematic text analysis framework;
- Developed a multi-label classification approach using the Chinese BERT-wwm pretrained model with a softmax classifier to automatically classify peer review text;
- Implemented a prototype literature retrieval system using Elasticsearch, integrating peer review information to improve retrieval accuracy.

FAILS: Automating Analytics of LLM Service Incidents | *Flask, OpenAI API*

- Led the development of the first open-sourced framework for automatic collecting and analyzing incident reports on different LLM services and providers; [FAILS]
- Providing 17 types of failure analysis, allowing analyzing temporal trends, and service reliability metrics;
- Leveraging advanced LLM tools to assist in data analysis and interpretation, enabling users to gain observations and insights efficiently.

Selected Publications [Google Scolar]

- **Xiaoyu Chu**, Shashikant Ilager, Yizhen Zang, Sacheendra Tallaui, Alexandru Iosup. "Leveraging LLMs for Structured Information Extraction and Analysis of Cloud Incident Reports". Under Review.
- Sacheendra Talluri, Dante Niewenhuis, **Xiaoyu Chu**, Jakob Kyselica, Mehmet Cetin, Alexander Balgavy, and Alexandru Iosup. "Cloud Uptime Archive: Open-Access Availability Data of Web, Cloud, and Gaming Services." Revision TPDS. [Paper]
- Sándor Battaglini-Fischer*, Nishanthi Srinivasan*, Bálint László Szarvas*, **Xiaoyu Chu**, Alexandru Iosup. "FAILS: A Framework for Automated Collection and Analysis of LLM Service Incidents". ACM HotCloudPerf2025. [Paper]
- **Xiaoyu Chu**, Sacheendra Talluri, Qingxian Lu, Alexandru Iosup. "An Empirical Characterization of Outages and Incidents in Public Services for Large Language Models". ACM ICPE2025. [Paper]
- **Xiaoyu Chu***, Daniel Hofstätter*, Shashikant Ilager, Sacheendra Talluri, Duncan Kampert, Damian Podareanu, Dmitry Duplyakin, Ivona Brandic, Alexandru Iosup (*Equal contributions). "Generic and ML Workloads in an HPC Datacenter: Node Energy, Job Failures, and Node-Job Analysis". IEEE ICPADS2024. [Paper]

Selected Supervision [Link]

- 2025, Msc Thesis, Yiren Bai. "Understanding Service Reliability of Large Language Models: An Empirical Characterization on Operator and User Reports". [Thesis]
- 2025, Bsc Thesis, Maja Bińkowska. "DataViz: A Business Data Visualization System Using LLMs". [Thesis]
- 2024, Bsc Thesis, Yizhen Zang. "Enhancing Operational Data Synthesis and Predictive Analysis in HPC Clusters Using Large Language Models". [Thesis]
- 2023, Msc Thesis, Shekhar Suman. "ODAbler: Design and Evaluation of an Operational Data Analytics Framework for Energy-efficient management of Workloads in a Data Centre Simulator OpenDC". [Thesis]

Awards

- ACM/ICPE Travel Grant, ICPE, May 2023, 2025
- CSC Scholarship for Ph.D., VU, Mar 2022 - Feb 2026
- Summer School Scholarship, Technion, Jul 2017 - Aug 2017
- CSC Exchange Scholarship, KNU, Sep 2016 - Jun 2017
- 1st Class Merit Scholarship, BNU, Sep 2020
- 2nd Class Merit Scholarship, LZU, Sep 2015, 2016