JOHN CHE

Johnche335 [at] gmail [dot] com | 608-960-2949

RECRUITER SUMMARY

Software Engineer with 3+ years of experience building scalable, data-driven systems. Key accomplishments include:

- Reduced threat detection time by 30% via a real-time ETL pipeline processing 5TB of daily security logs.
- Accelerated data throughput by 25%, enabling the onboarding of 150+ banking clients through automated data migration.
- Influenced $500K+ in cybersecurity investment by developing executive dashboards from complex data.
- Saved $30K annually by re-engineering a critical health data process with a unified SQL database.

EDUCATION

Master of Science (MSc), Computer Science | University of Wisconsin–Madison
Focus: Distributed Systems, Machine Learning, Security & Privacy, Big Data Systems

Bachelor of Science (BSc), Computer Engineering | University of Wisconsin–Madison

PROFESSIONAL EXPERIENCE

DATA ENGINEER (RESEARCH) | Universities of Wisconsin System Administration | Madison, WI |Aug. 2024-Present

- Automated anomaly detection across 13 campuses; engineered event streaming pipeline using Python, .NET, and REST APIs.
- Enabled $500K+ in cybersecurity investment decisions by developing Tableau dashboards that translated security data into actionable insights for executives.
- Reduced threat detection time by 30% by building a real-time ETL pipeline that consolidates 5TB/day of security logs for 165,000+ users across 13 distributed campuses.
- Secured critical data access by managing vendor (OmniSOC) and cross-departmental relationships.

TEACHING ASSISTANT, Software Engineering | University of Wisconsin-Madison | Madison, WI | Aug. 2023-May 2024

- Improved technical skills by 60% for 800+ students by designing CI/CD-enabled programming assessments in Java and Git.
- Achieved a 93% student satisfaction rate through personalized technical mentorship and project feedback.
- Enhanced industry proficiency by creating 4 original projects focused on version control, automated testing, and Agile workflows.

DATA ENGINEERING INTERN | Yapstone Inc. | San Francisco, CA | Jun. 2022-Aug 2022

- Accelerated data throughput by 25% enabling the onboarding of 150+ banking clients by building a pipeline (Python, Snowflake, AWS Lambda) that migrated PII from 50,000+ contracts, enabling sales teams with critical client insights for revenue retention.
- Delivered first-ever C-suite client dashboards by integrating Snowflake with Looker, providing real-time visibility for sales team.

DATA SERVICES SPECIALIST, HealthTech | Wisconsin Dept. of Health Services | Madison, WI | Jun. 2021-May 2022

- Saved $30K annually by re-engineering a critical process, replacing error-prone spreadsheets with a unified HIPAA-compliant SQL database serving 3,000+ clinics. Improved data reliability by 27% by automating task management eliminating manual errors.
- Secured project approval by conducting 15+ stakeholder interviews and authoring the business requirements document with a clear cost-benefit analysis.
- Reduced repetitive support requests by 75% by engineering a comprehensive user guide and support system for 15 users.

KEY PROJECTS

Multi-Label NLP Classification Pipeline | Feature Extraction at Scale

- Reduced computational overhead by 40% while retaining 95% of BERT's accuracy by fine-tuning DistilBERT on 1.3M documents across 82K+ categories.
- Tech: Python, PyTorch, Hugging Face, Docker

Distributed ML Training Pipeline | Multi-GPU Cloud Infrastructure

- Cut model training runtime by 75% by leading a 3-person team to deploy a distributed training system on GCP for a 300M-parameter transformer using Spark and PyTorch.
- Tech: Python, PyTorch, Spark, HDFS, GCP, Docker

PageRank Implementation on GCP | Scalable Data Pipeline for Graph Algorithms

- Built a scalable, fault-tolerant data pipeline by deploying the PageRank algorithm on a Docker Swarm cluster using Apache Spark and HDFS.
- Tech: Python, Apache Spark, Hadoop (HDFS), Docker, GCP

Network Intrusion Detection System | ML for Cybersecurity

- Developed a random forest classifier to detect threats by processing 150K+ labeled network requests from honeypot servers and addressing class imbalance with SMOTE.
- Tech: Python, scikit-learn, pandas, Docker

LEADERSHIP & SERVICE

Secretary | Association of Cameroonians in Madison (ASCAM) | Sep 2023 – Present
● Managed a $5K programming budget and coordinated community events for 80+ members.

Student Advisory Board Member | UW–Madison Dept. of Computer Science | Jan 2024 – June 2025
● One of 5 graduate representatives advising the department chair on initiatives affecting 6,000+ students.
● Helped design student spaces in the new $25M Computer Sciences building.

Member | Wisconsin AI Safety Initiative | Aug 2023 – June 2025
● Collaborated on interdisciplinary research to curate catalogues of trustworthy machine learning resources.

ACADEMIC RESEARCH

Che, J. (2025). Assessing the Challenges Facing Passkey Adoption: A Qualitative Study of User Barriers and Platform Design. Manuscript in preparation for ACM CHI or USENIX Security.
Che, J., Priyadarshi, A., & Kandasamy, S. (2024). Training Multilingual Transformer Models in Resource-Constrained Environments. Technical Report, CS 744, UW–Madison.
Alimaa, M., Che, J., & Chen, K.-Y. (2024). Extracting Structured Skills from Unstructured Job Listings Using Transfer Learning. Technical Report, CS 774, UW–Madison.

HONORS & AWARDS

● **UWSA Fellowship (2024)**: 1 of 7 statewide recipients for research excellence and leadership
● **King-Morgridge Scholarship (2019–2023)**: Full tuition merit scholarship (Top 6 of 800 international applicants)

TECHNICAL SKILLS

Languages: Python, SQL, Java, R
Data Engineering: ETL/ELT, Data Modeling, Stream Processing, Apache Spark, Hadoop (HDFS)
Cloud & Databases: AWS (S3, Lambda, SageMaker), GCP, Snowflake, MySQL, PostgreSQL, MS SQL, Elasticsearch (NoSQL)
ML/AI Frameworks: PyTorch, TensorFlow, scikit-learn, Hugging Face, NLP
DevOps & Tools: Docker, Docker Swarm, CI/CD, Git/GitHub, Databricks, Linux, Agile/Scrum
Visualization: Tableau, Looker, Matplotlib, Seaborn

Languages: English (Native), French (Fluent)