KE XIAO

📞 206-861-6267 ✉ luckyxiaoke@gmail.com 🔗 douglasxiao.github.io in linkedin.com/in/ke-xiao-b9b74b128/

PERSONAL SUMMARY

Tech Lead at Tiktok with 10 years of experience in Front-end Software Engineering, specializing in Tiktok shop E-commerce seller fulfillment, Azure Machine Learning, AI studio UI and infrastructure development, and collaboration with OpenAI. Skilled in web interface design, integration with large-scale systems, parallel computing, and machine learning models.

WORK EXPERIENCE

Tiktok, ByteDance   Mar 2024 − Present

Senior Frontend Software Engineer (Web, Lynx)   *Bellevue, WA, USA*

- Directed Tiktok shop E-Commerce seller fulfillment team, set up Web, iOS, Android development framework, pushed US fufillment team to work close-looply.
- Directed projects like US local to local, cross border fulfillment, and defined UIs across clients to share API and business layers, to expedite development cadence across 3 platoforms.

Azure, Microsoft   Oct 2018 − Mar 2024

Senior Frontend Software Engineer (Web)   *Redmond, WA, USA*

- Optimized Azure Machine Learning's developer experience, significantly boosted productivity for 100 developers across 2 organizations. Conducted the build tool transition from Webpack to Vite, and adopted Hot Module Replacement for instant updates. Improved 5x faster of web site cold start and refresh, modified eslint rule to instantaneous file-saving and instant reflection of changes, earned recognition and trust from the leadership team.
- Directed Retriever-Augmented Generation project, created intuitive interfaces with React and TypeScript, built uploading data component with JavaScript local storage and Azure blob/file storage, created document chunking and text embedding functionalities with Langchain, built vector index and storage with FAISS and Azure Cognitive Search, bumped the data to prompt flow to Azure OpenAI service and OpenAI connections to retrieve answers for users.
- Led the platform development of Azure Machine Learning Finetuning, drove discussions across teams including PM, designer and services. Built up Azure resources connections with Meta Llama-2 and Hugging Face to support models importing to AML for finetuning and evaluations, built up UI with React and TypeScript to support user setting LoRA parameters, connected user data and model with job scheduler service to distribute the training on serverless compute.
- Orchestrated the seamless integration of Azure Notebooks, integrated the 3rd party JavaScript bundle into CDN and consumed from Azure workspace resources provider, facilitated the embedding of it to the UI with React and TypeScript. Utilized the react-virtualized library to efficiently render large number(10K+) of files in Notebooks file explorer.
- Designed and implemented the Microsoft News article displaying portal with React and ASP.NET, defined the API with content-recommending service, obtained best matching article and linked it to the next button. Embed the articles into Bing search content, and improved the rankings for these articles, leveraging the Bing API to achieve this.

Amazon   Oct 2017 − Jun 2018

Full Stack Software Engineer (Web,Java)   *Vancouver, BC, Canada*

- Designed and implemented a backup payment method service with Java, communicated with Paystation services to get payments methods availabilities, and schedule the payment card automatically to improve user shopping experience. Integrated the payment information to the retail website with Amazon UI library, HTML, JavaScript and Perl.
- Streamlined the process of importing UpdateEverywhere payment service metrics into AWS Redshift by building up a pipeline to bump data to AWS SNS and SQS. Led the metrics analysis and tool development for lowering the costs and burden of whole team on-call efforts.

VMware   Apr 2014 − Sept 2017

Full Stack Software Engineer (Android)   *Beijing, China*

- Pioneered the design of a real-time audio-video input system for an Android remote login app using Java, C++, and the Android SDK, and introduced a module to capture device inputs. Implemented an Acoustic Echo Cancellation (AEC) module in Java, using the Least Mean Squares algorithm for echo-free communication.
- Led the design of a clipboard redirection virtual channel from an Android app to a cloud virtual desktop, utilizing Java for a rich text converter and Android SDK and C++ for an image caching system with the LRU algorithm.

Technical Skills

Programming Languages: TypeScript, JavaScript, HTML, CSS, LESS, SCSS, Java, Python, C++
Dev tools: Git, PostMan, Fiddler, Azure, AWS, Docker, Kubernetes
Frameworks: React, Jest, Enzyme, React Testing Library, ASP.NET
AI & Tools: Langchain, BERT, GPT, model fine-tuning, prompt engineering, CNNs, Keras, PyTorch, Transformer
Databases: MongoDB, Redis, Spark

Education

Beijing University of Posts and Telecommunications, Beijing, China   Sept 2011 – Apr 2014
Master of Computer Science   *ACM silver rewards*
Central South University, Hunan, China   Sept 2007 – Jun 2011
Bachelors of Computer Science