

# CHENGXUAN (SHELTON) XIA

669-666-4614 | xcx3205@gmail.com | <https://www.linkedin.com/in/chengxuanxia/>

## SUMMARY

Over 4 years of AI and engineer work experience. Proficient in tasks like searching, recommendation, text generation and LLM API function development, with full-stack skills including real-world data mining, engaging with cross-functional teams, solving ambiguous problems, and driving impactful results.

## TECHNICAL SKILLS

- Languages: Python, Java, R, Matlab, SQL, JavaScript
- Development & Deployment: Git, Docker, Github, Spark, Linux/Unix, Sublime Text, SVN, Deep Learning
- Framework & tools: PyTorch, TensorFlow, Selenium, Langchain, Pandas, AWS, Flask, Kubernetes, SageMaker, Oracle

## PROFESSIONAL EXPERIENCE

**GenAI Engineer** 12/2024 - Present

Walmart Global Tech

- Developing EndGenius, an **AI Agent** system for Walmart engineers that automates test script generation, support **Code Copilot** and troubleshooting suggestions. Enhanced the **Embedding** method, significantly improving accuracy
- Building a **Routing Agent system** that dynamically integrates additional agents and functions, enhancing automation and expanding overall capabilities

**Co-Founder** 02/2024 - Present

Shelton Angelina LLC

Sunnyvale, CA

- Developing products and websites as a full-stack, providing privacy-preserving AI solutions tailored for small businesses as an AI expert

**Backend Engineer, Machine Learning** 09/2023 - 06/2024

Petal.org

Sunnyvale, CA

*Medical Document Template Generation*

- Developed Approvo.ai, using **Retrieval-Augmented Generation(RAG)** and **LLM** to discover predicate candidates and provide medical device recommendation and template generation from FDA clinic trials
- **Fine-tuned LLM** for medical **document understanding** and formal writing, with user-provided textual records and question answering steps. The product averaged saving doctors 5hrs/week from trials writing

*Chatbot Development Platform*

- Developed GPT-Trainer, using **Langchain** to facilitate users to integrate pre-made **function APIs** and **customize function calls** for utilization within designed **agents and chatbots**
- Constructed **AI-Augmented support agent** and agent sharing framework, which cut the customer service cost by 50%

**Capstone Project** 03/2023 - 12/2023

Adobe Inc.

Santa Clara, CA

- Implemented a web-based user interface using JavaScript to enable **prompt engineering**, facilitating multi-turn interactions and offering multi-choice user guidance
- Developed a multimedia system utilizing a LLM as backend and **Stable Diffusion** as generation model to provide enriched prompts, with a focus on assisting novice users in maximizing the benefits of Text-to-Image generation AI

**Machine Learning Engineer** 06/2023 - 09/2023

BearHug AI Inc.

Sunnyvale, CA

- Built an **AutoGPT agent** utilizing **Langchain** and **Selenium** for interactive operation and manipulation of dynamic web content.
- Implementing an optimized **web page parsing** method, reducing reaction time from 2 minutes to 40 seconds

**Software Engineer, Machine Learning** 09/2020 - 07/2022

Beijing Paoding Technology Co.

Beijing, China

*Auto Data-Annotation System for Financial Documents*

- Trained a **weak-supervision Transformer model** for keyword-based label generator to generate 6 million high-quality labeled data, **saving** the company **6 months'** work and cost of **\$20,000** from human annotation

## EDUCATION

M.S. **Natural Language Processing (GPA: 3.9/4.0)** University of California, Santa Cruz

M.S. **Computer Science (GPA: 3.7/4.0)** University of Chinese Academy of Sciences

B.S. **Physics** University of Chinese Academy of Sciences