Ajay T Shaju

Artificial Intelligence & Data Science Engineer

☎ +91 9567117161 | ✉ ajaytshaju@gmail.com | 🌐 sites.google.com/view/004ajay | ⌽ github.com/004Ajay | 🔗 linkedin.com/in/ajay-t-shaju

Skills

Programming Languages:	Python, LaTeX
Operating Systems:	Linux, Windows
Libraries/Frameworks:	Pandas, OS, Numpy, BeautifulSoup, Request, Matplotlib, TensorFlow, Open-CV, Playwright, LangChain
Technical Tools:	VS Code, Jupyter Notebook, Docker, MS Office, Google Suite, Photoshop, Figma, Excalidraw, Mermaid
Abilities:	Communication, Empathy, Sociability, Problem Solving, Observation, Leadership, Time Management

Experience

ITC Life Sciences & Technology Centre (LSTC) | Data Science Trainee
Peenya, Bengaluru, India | 2nd Dec 2024 - Present

- Developed a multi-node **Streamlit Barcode Generator App replacing proprietary software**, enabling **parallel barcode printing** across 4 laptops (projected 6-10) and **cutting asset verification time by 75%**—allowing the **audit target to be met weeks ahead of schedule**.
- Created a **local Retrieval Augmented Generation (RAG) system** for the **OpenWebUI-Ollama AI platform** to answer Ayurveda-related queries, integrating multilingual **PDF-to-Markdown extraction via Docling** and **structure-aware chunking** and feeding into a **Chroma vector database**—achieving strong user test performance.
- **Implemented Git-based version control** for a research group lacking structured workflows, **improving code organization and collaboration quality by 60%** through streamlined version tracking and branching.
- **Helped peers in modern DS/SWE tools and practices**, **increasing team knowledge retention and adoption of best practices**.
- **Created a personal technical knowledge base** by documenting commands, gists, and workflows on a website, **reducing repetitive web searches and command retyping by 50%**.
- **Automated IT security checks** across Linux-based work computers with Bash scripting and cron jobs, **cutting redundant IT security workload by 4–5 hours monthly**.
- **Built a web scraping pipeline** for three complex websites, **boosting data collection efficiency by 50%** through parallelization and optimized request handling.
- **Engineered a high-speed PDF data extraction system**, **achieving 40-page document parsing in 7 seconds with >95% accuracy**, drastically improving document processing efficiency.

Projects

Inkwell: AI & Retrieval-Augmented Generation (RAG)-Based Personal Writing Assistant *June 2025*
AI Writing Assistant *Source Code - GitHub*

- Designed an **alternative to fine-tuning LLMs** by combining *RAG* with a custom **"Style-Augmented Generation"** approach to **mimic personal writing style** from diverse blog documents.
- Implemented an AI assistant that prioritizes generating text in the **author's style over generic context-based responses, moving beyond typical RAG pipelines.**

Automated Financial Tracker using Gmail, Zapier, and Google Sheets *October 2024*
Personal Finance Dashboard Automation *Medium Blog*

- Collected, cleaned, and organized **950+ rows of financial transaction data** using specific steps in **MS Excel and Python**.
- Automated personal finance tracking by **integrating Gmail and Google Sheets via Zapier** to streamline financial data logging.
- Designed a Google Sheets dashboard with **Donut and Bar Charts, enabling real-time insights** across categories.

Marks2CSV: A tool to convert marks in answer sheets to CSV format using OCR *January 2023 - June 2023*
Mark Entry Automation | B.Tech Mini Project *Source Code - GitHub*

- **Addresses the time-consuming task of mark data entry** of our teachers in the college
- Developed a customized software, resulting in **time savings of 5+ hours per week** with accuracy by 98%
- **Presented and Accepted at ICTEST24 Conference** organized by GMEC, Thrikkakara, Ernakulam. Kerala, India

Education

7.8/10	**Bachelor of Technology (Artificial Intelligence & Data Science, AD)**, *SJCET, Palai, Kerala, India*	*2020 - 2024*
84%	**Higher secondary**, *Sanjos Senior Secondary School, Palai, Kerala, India*	*2020*
85%	**Matriculation**, *Sanjos Senior Secondary School, Palai, Kerala, India*	*2018*

Certificates

2024	**Career Essentials in Data Analysis by Microsoft-LinkedIn** , ETL, Analysis, Visualization of Data	*LinkedIn*
2023	**British Council English Score Certificate**, Core Skills CEFR C1 (539) and Speaking CEFR B2 (428)	*British Council*
2022	**Python For Data Science**, Intermediate course to do Data Science with Python	*NPTEL IIT-M India*
2021	**Joy of Computing Using Python**, Introduction to Python Programming	*NPTEL IIT-M India*

Languages

Malayalam	Native proficiency (Mother Tongue, Public Speaking)
English	Professional proficiency (CEFR Tested, Contribution to Blender FOSS Repository)