Santiago Saldivar

631-655-5962 | santiagosaldivar19@gmail.com | linkedin.com/in/santiagosaldivar | github.com/santi1999 | santiagosaldivar.com

EDUCATION

Fordham University   Bronx, NY
B.S. in Information Systems   *Graduated: Spring 2023*

Suffolk County Community College   Riverhead, NY
A.S. in Accounting — Norman R. McConney, Jr. Award Recipient   *Graduated: Spring 2020*

EXPERIENCE

Data Engineer / Software Developer   Mar. 2023 – Present
Gov Glance Foundation   *Remote*
- Developed and maintain 300+ web crawlers with 99% reliability for government data collection.
- Built scalable ETL pipelines using Python, PostgreSQL, and MongoDB.
- Deployed full-stack metasearch engines with Django and SearXNG used AI endpoints to generate answers to queries.
- Achieved 7x faster search performance of 30 gb of records from 7s to 1s by combining metadata indexing and strategic data normalization and lexeme-based pre-processing.
- Implemented vector embeddings on markdown data for Q&A system with local LLM indexing and citation-based answers.
- Managed cloud infrastructure (DigitalOcean) and implemented CI/CD workflows.
- Applied text similarity techniques such as Cosine Similarity, Levenshtein Distance, and Jaccard Similarity to enhance data extraction and accuracy along with reducing hard-coding of HTML selectors.

Assistant Project Manager / Lineman   Jan. 2022 – Mar. 2023
Sound Aircraft Services   *East Hampton, NY*
- Oversaw JIRA tickets for internal software, coordinated QA fixes and feature updates.
- Supported aircraft fueling operations and real-time logistics.

PROJECTS

Local Building Codes PDF Q&A Chatbot | *Python, Supabase, LangChain, SQLite, Local LLMs*   2025
- Built UI for self-hosted AI system that will parse PDFs and returns page-cited answers.
- Parsed byte data of PDF to reduce overhead on AI Indexing of legal documents and town codes for compliance and public use.

Government Metasearch Engine | *SearXNG, Django, MongoDB, NGINX*   2023
- Customized UI/UX and backend of SearXNG to unify search across government datasets.
- Improved frontend performance and added API integration with scraped sources.

Things to Do in The Hamptons | *Next.js, Supabase, AI Parsing*   2024 – Present
- Built a local event aggregator with current Web extraction stategies
- Implementing heuristics to reduce noise on a webpage for easy retrieval of main content
- Added admin panel and SEO-focused article rendering with SSR.
- Created an Instagram and TikTok following to create community and collaboration to get real users and engagement on the platform
- Currently 1000 unique users on Things to Do in The Hamptons

TECHNICAL SKILLS

Languages: Python, PostgreSQL, JavaScript (TypeScript), MQL, Bash, HTML/CSS
Frameworks/Libraries: React, Next.js, Flask, Django, LangChain, FastAPI, Tailwind, Supabase
Tools & Platforms: Git, DigitalOcean, Postgres, MongoDB, Supabase, Vercel, NGINX, JIRA, Docker
AI/ML: scikit-learn, OpenCV, Tesseract, local LLMs, OCR pipelines
Web Crawling: Scrapy, BeautifulSoup, Requests, AsyncIO, Newspaper3k, Trafilatura
Other: PDF parsing, CI/CD, SEO optimization, Server-Side Rendering (SSR)