Wilson (Wei-Sheng) Wang

Sr. Backend Engineer

josixwang@gmail.com
linkedin.com/in/josixwang

josix.tw
github.com/josix

TECHNICAL EXPERIENCE

Sr. Backend Engineer / Appier, Inc.   **Taipei, Taiwan   Mar 2023 — Now**

LLM Prediction Service

- Led the development of an LLM-powered service, implementing a robust prompt engineering pipeline with Haystack and FastAPI. Integrated OpenTelemetry, Google Cloud Trace for comprehensive monitoring of usage, costs, and prediction performance. Delivered a clean, well-documented codebase using tools like ruff, pre-commit, and Pydantic.
- Designed and implemented an LLM batch prediction framework using Airflow's dynamic task mapping and deferrable operators, optimizing scalability for large-scale user segmentation with Trino. Successfully integrated OpenAI batch operations, ensuring compliance with SLA requirements.

Data Connectors

- Led end-to-end development of a scalable error handling microservice supporting 7 connection types (S3 bucket, S3 file, GCS, SFTP, BigQuery, Google Sheets). Delivered Airbyte API integration and successfully deployed the enhancement, improving system reliability and troubleshooting efficiency.
- Resolved a longstanding pod creation issue in Airbyte data integration that previously required manual restarts. Setup OpenTelemetry metrics collection to proactively alert on failed connection creation, enhancing system observability.

Segmentation System

- Designed and implemented integration across 6 feature components using the Feature Flag Decorator pattern, ensuring code maintainability and enabling ease of unit testing while accelerating feature delivery.
- Designed offline event segmentation and data preprocessing workflows leveraging Celery, Trino, and GraphQL. Accelerated feature delivery by 60% and achieved 100% key results completion, including custom segmentation conditions and a unified data model.

Data Engineer / Tresl, Inc.   **Palo Alto, CA   Oct 2022 — Nov 2022**

airbyte

- Decoupled the customer data extraction and loading workflow from Airflow by using Airbyte. Reduce the complexity of the DAGs and the CPU usage of Airflow by around 20%.
- Built the ETL pipelines that extract data from various sources, load them into the GCS and send them to BigQuery for further analysis.

ML Engineer / Miso Technologies, Inc.   **San Francisco, CA   May 2021 — May 2022**

miso-webhook

- Responsible for building the ETL pipelines for ingesting data from Shopify API and sending them into internal data API in around 80 requests per minute from Shopify Webhook API.
- Reduce response time by 80% by using multi-layer caching consisting of memory cache and Redis cache to avoid taking too much time on fetching authentication data from the internal API.

string-index

- Responsible for building a key-value store for mapping productID/userID into integer and ndarray at a performance around 4 ms per 11 million query words, which is the core component used in search and recommendation models.
- Reduce the memory usage by 85% with 11 million tokens input by switching the hash map into DAWG to record the tokens. Persist the tokens in the disk and achieve the purpose of sharing memory through mmap.
- Decoupled the string-index component from the code base to make it easier to maintain and extend. Build a standalone package with complete usage documentation, 100% unit test coverage, and a unified coding style conforming to the black, pylint rules.

python-package-template

- Shorten development environment setup time by 60% by using Cookiecutter. Supporting virtual environment management tools like pipenv/poetry, pre-commit hooks, linting, testing, formatting commands, and GitLab CI/CD, Dockerfile configurations.

Wilson (Wei-Sheng) Wang

Sr. Backend Engineer

josixwang@gmail.com
linkedin.com/in/josixwang

josix.tw
github.com/josix

ACTIVITIES

Open Source Contributions

@apache/airflow

- Implemented the integration of OpenAI Batch API support in Apache Airflow, including creating a new hook, operator (deferrable compatible), and trigger, along with comprehensive unit testing and documentation updates. (#41554)
- Implemented a tags filter feature for DAGs list in Apache Airflow, enhancing the UI as per AIP-38 Modern Web Application. (#43303)
- Resolved issues in Airflow Connection by preventing an AttributeError when setting up MongoDB connection and fixing an InvalidURL error in Airbyte hooks, enhancing the reliability and stability of the platform (#43024, #38860)

@airbytehq

- Implemented the new source connector that ingests 5 different resources from ClickUp API through the low-code CDK with the passed acceptance test and complete documentation. (#17770)
- Reported one extensibility issue that occurred when running acceptance test in python 3.10+ (#17855)

@python

- Co-organized the python documentation localization meetup regularly and increased the amount of translated documents by around +40%

PyCon Taiwan   **2019 — Present**

Web Squad Lead   *2022 — 2024*

- Coordinated around 8+ team members to play different roles like UI/UX designers, developers, and reviewers. Make sure the internal communication is smooth by building meeting minutes, onboarding docs, and feature docs to ensure the conversation is public and easy to access.
- Monitor work in the kanban board, make sure to deliver the features requested by the Program Team, Sponsor Team, Registering Team, etc. Distribute the tasks to the team members, and follow up on the progress.

Infra Squad Lead   *2021 — 2023*

- Coordinated around 7 members to fit into the infra squad and help them understand and contribute to the projects by mentoring and 1-1 meetings.
- Maintaining 4 projects [1] for supporting the Program Team, Web Team, and Committee to build an automatic workflow through CLI tools.

Developer   *2019 — Present*

- Contributed to the pycon.tw and pycontw-frontend, including implementing the homepage, about, and sponsor page with the card, modal components, and the flexbox, grid layout.
- Responsible for reviewing the opened PRs. Make sure that code uses the unified naming convention, and shared components, and check the functionalities and appearance work correctly. Giving the feedback following the conventional comments guideline.

EDUCATION

M.S. in Computer Science, *National Chengchi University*, *Taipei, Taiwan*   Feb.2020 — Nov. 2022

- Published *Improving Complete Cold-Start Recommendation via Content-Based Preference Graph Convolution Networks*, which is a **graph-based hybrid recommendation model used to resolve the complete cold-start recommendation issue** that occurred in real-world recommendation systems like ticket-selling platforms and news platforms.

B.S. in Computer Science, *National Chengchi University*, *Taipei, Taiwan*   Sep. 2014 — Feb. 2020

SKILLS

Programming Language	Python
Data Engineering	Airbyte, Kedro
Backend Development	FastAPI, Django, Redis, ElasticSearch
Developer Experience	invoke, pylint, isort, black, mypy, pre-commit, Commitizen, pytest, loguru, MkDocs
DevOps and Other Tools	Docker, Git, GitHub Actions, GitLab CI, AWS

---

[1]session-video-publisher, hackmd-archive-worker, talk-preview-imag-builder, and pycon-archive-past-website