

Jack Kowalik

Troy, MI · jackkowalik20@gmail.com · (248) 404-5724
authorize.earth · docs.authorize.earth · linkedin.com/in/jackkowalik

Experience

Founding Software Engineer, Authorize Earth

Troy, MI Mar 2024 – Present

Seed-stage identity verification and compliance platform for high-risk merchants. Designed the data, ML, and API layers as the only engineer.

Data pipelines and validation

- Built two screening systems on authoritative government sources: politically exposed persons from 37 feeds across 207 jurisdictions (794K current records), and sanctions entities from 25 government lists, reconciling divergent formats into one schema with daily refresh. - [authorize.earth/products/sanctions-benchmarks](#), [authorize.earth/products/pep-benchmarks](#)
- Built TIN validation for 180 countries against personal and business formats. - [authorize.earth/products/tin-benchmarks](#)
- Built adverse media screening that matches extracted name, DOB, and location against negative news and public criminal cases, using Google Programmable Search with Puppeteer: 98.5% article surfacing rate on 274 known adverse subjects. - [authorize.earth/products/adverse-media-benchmarks](#)
- Wrote document extraction on VLM-based OCR with field-level output into typed schemas: 97.3% single-pass extraction on a 6,400-sample benchmark. - [authorize.earth/products/benchmarks](#)
- Maintained per-product benchmarking pipelines and evaluation infrastructure, with public benchmarks posted for OCR extraction, sanctions and PEP screening, TIN validation, age estimation, and adverse media.

Infrastructure and delivery

- Built the TypeScript/React SDK for client identity verification and the merchant dashboard for management of regular tasks. Also built the PHP backend that communicated with FastAPI services.
- Served every model as a Dockerized FastAPI service on RunPod GPU instances, sized by VRAM footprint; measured each service's marginal cost from hourly GPU rates and inference time: ~\$0.01 GPU compute per verification.
- Shipped a REST API with webhook delivery backed by Azure Cache for Redis and Azure Key Vault for secrets.
- Developed the platform's first generation under Plaid's Reseller Partner program for account and identity verification before narrowing to identity verification on in-house models. Passed their audit for PII and bank account data handling.
- Published a Shopify transaction-security app to the App Store, integrated in production by ~50 external merchants.

Fraud detection and biometrics

- Trained a high-resolution six-model document forgery detection ensemble covering digital injection, screen recapture, print copies, consumer substrates, face manipulation, and text-field tampering.
- Built facial age estimation for age gating; zero-shot evaluation on 97,698 public FairFace faces measured 0.95 under-18 AUC, held within 0.02 AUC across seven race groups. - [authorize.earth/products/age-estimation-benchmarks](#)
- Ran 1:N face matching on per-merchant FAISS HNSW indexes, plus a presentation attack detection model trained on a consented, paid data collection study. Also built 1:1 selfie matching, comparing a photo against the ID document image.

Projects

Sanctions ETL - Open-source PHP ETL syncing government sanctions lists into JSONL or MySQL, with CI and a PHPUnit suite: [github.com/jackkowalik/sanctions-etl](#)

Doc value extraction - Open-source Python pipeline that extracts value fields from identity documents. EasyOCR collects all string boxes, Qwen2.5-VL-7B semantically separates value vs label fields, then fuzzy matching aligns them: [github.com/jackkowalik/docvalue-extract](#)

Education

Oakland University

2024–2025

Business Administration (left to build Authorize Earth full time).

Skills

Languages: Python, TypeScript, PHP, SQL

Web: React, Webpack

Data and backend: REST APIs, ETL and normalization pipelines, Docker, FastAPI, Azure Cache for Redis, Azure Key Vault, Google Programmable Search, Puppeteer

ML / CV: PyTorch, CNN and transformer architectures, FAISS HNSW, VLM extraction, LLM entity resolution

Domains: identity verification, sanctions and PEP screening, document forensics, biometric matching