AARON BARLOW

865.804.6746 | abarlow505@gmail.com | github.com/aroswift | aaronbarlow.dev

EXPERIENCE

Oak Ridge National Laboratory — National Center for Computational Sciences (NCCS)

HPC Software Engineer | Jun 2020–Present

Scope: Build researcher self-service and fleet operations software across **27 HPC clusters** (incl. Frontier), enforcing **policy-as-code** and secure access for a **$700M+** compute footprint.

- Enabled **privacy-preserving federated learning at exascale** on **Frontier (first U.S. exascale system; TOP500 #2)**, unlocking cross-institution healthcare data training on AMD GPUs. Brought NVFLARE to ROCm/MI250X with PyTorch (ROCm), produced HIP-compatible builds, packaged deterministic ROCm containers, and integrated with Slurm with mTLS inside secure enclaves. Validated multi-node training rounds previously not possible on Frontier. Now used to build **privacy-preserving medical foundation models** on Frontier.
- Owned and scaled **myOLCF** (researcher self-service & monitoring) used by **thousands (q̃k)** across **1,000+ projects** with **99.9%+** availability. Shipped **12** FY2025 releases. Governs access, allocation, and policy across open, moderate, and secure enclaves.
- Replaced Jbuilder with a JSON API serializer and explicit Redis caching for hot read paths, taking the 100k-row directories endpoint from s̃72s to **178–206 ms** on cached hits (0.206s avg) (≈**1320×**; **−99.92%**) and improving first-hit from **298.5s** to **246.8s** (**−17%**). Eliminated view-rendering cost (s̃52s → õms) and cut ActiveRecord time (ĩ9.2s → **8.8ms**). Stabilized fleet-wide polling and removed timeouts during cluster syncs.
- Extended the Netflix-originated **fast_jsonapi** fork (**jsonapi-serializer**) via ‘lib/library_extensions/jsonapi_serializer.rb to support legacy root envelopes for V3/V4, runtime cache flags/TTLs, and per-record invalidation; enabling progressive rollout across endpoints.
- Accelerated build, startup, and test cycles by migrating to Vite and optimizing CI. Builds **2 min → 9 s (-92%)**, startup **30 s → <200 ms (-99%)**, tests **90 s → 6 s (-93%)**. Cut developer idle time and sped up review cycles.
- Built the **Smart Facility metrics platform** (backend in Crystal/Amber) ingesting compute/data/I/O/efficiency metrics; pre-aggregated dashboards deliver **100k-record** views with low latency and serve **5̃0k time-series points in 8̃7 ms**. Benchmarked ̃10.4k req/sec per CPU core (9̃6 µs/req) for simple JSON. Shifted heavy computation to background jobs and index tables. Used by lab leadership to guide next-gen procurements (RAM/IO/GPU/storage priorities) and to benchmark domains, flagging inefficient Slurm jobs.
- Standardized **GitOps delivery** with Kustomize + Argo CD on Kubernetes (GitLab runners), enabling declarative configs, drift correction, and one-click rollbacks across environments.
- Unified day-to-day operations across **27 HPC clusters** (incl. **Frontier**) via a central **policy-as-code** service. Automated project provisioning, access control, and scheduler policy across SLURM/LSF enclaves. Backbone software for a **$700M+** compute fleet and mixed open → secure environments.
- Automated project lifecycle and data-integrity workflows, eliminating manual back-office corrections for project/application records.
- Zero post-deploy incidents over the past 3 years by expanding Cypress/feature coverage and tightening parameter validation.
- Mentored **3** interns/junior engineers; coached **100+** students through PEARC/SC workshops. Authored runbooks, SLOs, and incident playbooks.

Bank of America — Consumer, Small Business & Wealth Tech

ML Engineer Intern | Jun–Aug 2019

- Built an **NLP entity extraction service** (names/phones/addresses/accounts/amounts) at **96% F1**, supporting **$20M+/yr** in savings through automation

Oak Ridge National Laboratory — National Center for Computational Sciences (NCCS)

Software Developer Intern | May 2015–May 2019

- Year-round development of HPC-centric services, applications, and BI tools. Shipped production features across internal portals.

- Automated supercomputer access communications via a policy-aware email system. Standardized messaging and reduced manual steps.
- Built a WordPress/REST plugin to sync & display HPC metrics on olcf.ornl.gov. Improved data freshness and reduced update toil.

SKILLS

- **Languages:** Ruby, Python, Go, TypeScript (Node), SQL, Bash
- **Back-end:** Rails, FastAPI/Flask, gRPC, REST, Redis, Postgres
- **Infra/Obs:** Docker, Kubernetes, Kustomize, **Argo CD**, **Slurm**, CI/CD, Prometheus/Grafana, OpenTelemetry, **ROCm**, HashiCorp Vault, MinIO, KubeDB, Stash, SonarQube
- **ML/FL Infra:** NVFLARE, PyTorch (ROCm), HIP
- **Agentic Systems:** multi-agent orchestration (**CrewAI** for Anthologia & prediction-market system), autonomous pipelines, headless uploaders, queueing/retries

EDUCATION

East Tennessee State University — Bachelor of Science, Computer Science | Dean's List | GPA **3.94/4.00** • May 2020

Activities: **ACM (President, 2019–2020)**; **Ethical Hacking (Vice President, 2018–2019)**

SELECTED PROJECTS

- **Automatic movie creation (agentic pipeline)** — multi-agent workflow generates script → storyboard images → 10-sec clips (Runway/Cling) → stitches, QC, and **auto-publishes**.
- **Anthologia (AI video storytelling)** — multi-agent pipeline (**CrewAI**) turns concepts into **20-second** videos with **4 scenes**, using DALL·E 3 + Veo 3.0 (fallback Kling), ElevenLabs narration/SFX, and LatentSync lip sync; typical **4–12 min** per production with automated QC/fallbacks.
- **Project Cadenza (agentic music + video)** — agentic pipeline creates artists/albums, writes lyrics (GPT), generates songs (Suno/Udio), masters audio, produces thumbnails/art, and **auto-uploads to YouTube** with metadata/scheduling; **published 150 videos**; prepares ISRC/metadata for distributor workflows.
- **Policy-driven data synthesizer (training automation)** — takes a policy/goal and **auto-creates synthetic training data** → trains/refits a model with quality guardrails.
- **Automatic product creation (e-commerce)** — AI agents (GPT) and custom orchestration + Flux Pro image gen → vectorize/upscale → SEO titles/tags → **auto-publish** via Printful/Etsy and Redbubble; **152 products created** with **57k+ product views** to date; **cost ≈ $0.31 per product**; database-driven workflow with retries and headless uploaders.
- **Automatic music video creation** — audio beat-map → storyboard → video assembly → captioning → render; **478 views** with **10.8 hours watched** on recent releases.
- **Automatic podcast creation** — multi-channel system (12+) with **Claude 3.5 Sonnet** scripts (7k+ words) → ElevenLabs TTS → audio assembly → **auto-publish** to Spreaker with metadata/scheduling; produces **6+ episodes per run** and has shipped **192 episodes**, totaling **1.8k downloads** and **50+ hours listened**.
- **Automatic book creation** — multi-agent pipeline (CrewAI) orchestrates GPT-4o-mini writing, Flux Pro/Replicate illustrations with vision-based QC, then compiles **print-ready PDF/EPUB** via ReportLab/LaTeX; metadata persisted in Firebase; typical cost **<$5–10/book**.
- **AI prediction-market trading system (Kalshi)** — multi-agent analysis (**CrewAI**) + real-time WebSocket feeds, vector DB, and Kelly sizing; selects market and executes end-to-end in **<6 minutes** (dev stage), with Dockerized services and a live dashboard.
- **HOA management platform** — multi-tenant Rails+React application with modules for dues, violations, residents, and public sites; role-based access, reporting, and React/Vite front-end with Inertia; built for production but not yet deployed to live HOAs.

TALKS & COMMUNITY

- **CUG 2025** — ”Employing a Software-Driven Approach to Scalable HPC System Management.”
- **NLIT 2024** — ”Employing DevOps in HPC Operational Management.”
- **ORNL Pathways to Computing Internship Program** — Workshop Chair, 2022–present.
- **PEARC Conference** — Student Program Committee Chair, 2021–present.