

Benjamin Johnston

Columbus, OH · (614) 843-3974 · bejaunst@gmail.com · linkedin.com/in/bejaunst
github.com/cetio · cetio.github.io

TECHNICAL SKILLS

Systems Programming C, C++, C#, x86 Assembly, D

Scripting & Automation Python, Ruby, JavaScript, TypeScript, SQL

AI & LLM RAG, LLM Integration, AI Agents, Model Evaluation, Structured Output, Tool Calling, Embeddings, Vector Search, MCP

Protocols & APIs W3C WebDriver, OAuth 2.0, REST, SSE Streaming, DDC/CI, JSON Schema, WebSocket, SDK Design

Infrastructure & Tools Linux, Git, CI/CD, Docker, CMake, Selenium, SIMD, GitHub Actions

EXPERIENCE

Software Engineer

Jul 2021 – Present

Independent client engagements

Columbus, OH

- Architected and delivered SDKs, automation tooling, and API integrations for individual clients, including an AI-assisted grading platform with LMS OAuth that reached ~95% grading accuracy and cut instructor workload in half.
- Authored Intuit, the first native LLM SDK for D (8,400+ lines, 67 modules) with structured output, tool calling, and multi-provider support; contributed 120 commits to fern-pl/fnc (Fern Native Compiler), serving as the second-highest contributor.
- Designed and maintain 10+ public open-source libraries across C#, C, C++, and D, spanning systems programming, protocol implementation, and developer tooling with ~53,000 lines of code and 700+ commits.

Automation Engineer (RPA & QA)

Nov 2025 – Present

Handshake

Columbus, OH

- Cut manual annotation and triage effort by 95% across 400+ contractor hours by building AutoHandshake!, a Ruby/JavaScript Selenium automation suite with UI-stall recovery, multi-tab orchestration, and full audit trails.
- Evaluated 4,000+ multimodal AI responses against structured quality rubrics for correctness, reasoning accuracy, and domain understanding, earning the Handshake Model Validation 2 — Expert credential.
- Evaluate LLM quality on coding and knowledge-based tasks using head-to-head model comparisons, benchmark evaluations, and structured quality criteria, analyzing failure modes to inform model-quality improvements.
- Mentored contractors on platform workflows and rubric interpretation, raising evaluation consistency team-wide and reducing rework from misaligned quality standards.

AI Evaluation Engineer

Dec 2025 – Jun 2026

Snorkel AI

Columbus, OH

- Engineered constraint-based training workflows for Claude and ChatGPT across 2 contracted projects targeting complex programming and logic reasoning tasks.
- Designed rigorous evaluation frameworks to systematically assess model reasoning capabilities, driving continuous improvements in code generation and logical deduction accuracy.
- Validated training-set integrity through Docker-based CI/CD, catching defects pre-release and protecting downstream model quality.

Search Quality Evaluator

Apr 2025 – Aug 2025

RWS Group

Columbus, OH

- Assessed search relevance and LLM output quality against detailed rubrics across Google query simulations and third-party models, identifying failure patterns that informed ranking and UX improvements.

Biochemistry Intern

Columbus Alternative High School

2022 – 2024

Columbus, OH

- Managed safety data sheets (SDS/MSDS) and lab compliance documentation across ~12 science classes, ensuring all materials met district and state safety standards.
- Prepared lab materials and set up experiments for chemistry and biology courses, supporting hands-on instruction for hundreds of students across multiple grade levels.
- Graded assignments, lab reports, and exams using structured rubrics, applying consistent quality standards that paralleled later work in AI response evaluation.

SELECTED PROJECTS

Intuit | D | github.com/cetio/intuit

2025 – Present

- Authored the first native LLM SDK for D (8,400+ lines, 67 modules), mapping chat completions, structured output, SSE streaming, and tool calls onto typed D structures.
- Supports OpenAI, Claude, and Qwen with OpenAI-compatible routing, SIMD-accelerated embeddings, and automatic JSON schema generation for type-safe native tool use.

semi-sleuth | Ruby on Rails 8 | Private

2026

- Built an MCP (Model Context Protocol) server enabling AI agents to query and organize research data through typed schemas with a Rails 8 dashboard for human review.
- Full-stack product: Hotwire/Turbo, Stimulus, TailwindCSS, SQLite→PostgreSQL, Docker deployment, multi-tenant architecture.

UndeadInterop | C# | github.com/cetio/UndeadInterop

2024

- Built a Windows syscall research library in C# that resolves syscall IDs from PE export tables and invokes them directly via runtime-generated x86_64 shellcode stubs, bypassing the ntdll usermode layer entirely.
- Implemented 8 hook detection classifications (inline jumps/calls, stack pivots, debug interrupts, forwarded exports, tamper detection) by disassembling function prologues with the Iced x86 decoder; includes a complete NTSTATUS enum (339 values).

VEXAPI | C/C++ | github.com/cetio/VEXAPI

2024

- Reverse-engineered the VEX Robotics V5 private firmware API from ARM64 disassembly of VEXos using Ghidra, producing 8,650 lines of documented C/C++ headers across 39 files with 481 firmware memory offsets.
- Documented 34 public API headers covering 20+ device types (motor, IMU, vision, optical, display) and 5 private headers with jump table and boot sequence signatures; cited by the Rust vex-sdk crate.

Cassowary | C# | github.com/cetio/Cassowary

2024

- Authored a .NET runtime introspection library (7,400+ lines, 54 files) exposing CoreCLR internals as first-class managed types for reflection, allocation, and layout inspection without native code or P/Invoke.
- Modeled 15 CLR internal structures (MethodTable, EEClass, MethodDesc, FieldDesc) and 15 generic type signature models; referenced by Novus, one of the largest .NET reflection libraries.

Nimbus | C | github.com/cetio/nimbus

2025

- Built a Linux CLI (2,500+ lines) for unified display configuration controlling brightness, contrast, gamma, and color temperature across both sysfs backlight and DDC/CI over I2C backends.
- Implemented function-pointer-based polymorphic display abstraction with runtime capability detection, supporting 5 VCP feature labels plus arbitrary hex codes; includes Ruby integration test suite.

Selenium | D | github.com/cetio/selenium

2026 – Present

- Authored the only maintained W3C-compliant Selenium WebDriver client for D (4,800+ lines, 28 modules), with typed APIs, driver auto-detection for 4 browsers, and full session lifecycle management.
- Implemented a Bridge/Driver architecture with concurrent session support; published on the DUB package registry with Selenium Grid support.

- Built a contracted D library that parses and exposes game data as strongly typed objects across 12+ modules, with filesystem caching, version invalidation, and configurable cache policies.
- Depends on cetio/selenium-sdk for WebDriver-driven data extraction; includes a separate client application rendering 10,000+ tiles and 5,000+ units via WebSocket feeds in under a second.

OPEN-SOURCE IMPACT

48 public repositories, 60 total projects, ~53,000 lines of code, 700+ commits across D, C#, C, C++, and Ruby.

5 DUB packages published (Intuit, Multidasm, Tern, Akashi, Webull-SDK).

120 contributions to fern-pl/fnc (Fern Native Compiler) — second-highest contributor to an upstream compiler project.

VEXAPI cited by the Rust vex-sdk crate, demonstrating external real-world impact.

EDUCATION

Columbus State Community College

AAS, Software Development

Expected Dec 2026

Columbus, OH

CERTIFICATIONS

Handshake Model Validation 2 — Expert AI response evaluation across 4,000+ multimodal samples

MS Career Essentials Generative AI (NASBA)