

# Clinton Judy

Boalsburg, PA | [clinton@judy.io](mailto:clinton@judy.io) | [github.com/judy](https://github.com/judy) | [linkedin.com/in/clintonjudy](https://www.linkedin.com/in/clintonjudy) | [clinton.judy.io](https://clinton.judy.io)

## SUMMARY

---

I build production AI products end to end. At Zipline I designed and shipped Zippy, the company's flagship AI assistant: a retrieval system now answering frontline retail workers across 50+ brands and 500K+ conversations. I shipped contextual, atomic-fact retrieval to paying customers a year before Anthropic published the same core idea as Contextual Retrieval. That work rests on 15+ years as a Rails engineer and SRE operating at scale, and today I'm building a from-scratch agent runtime with its own memory and emotional substrate.

## WORK EXPERIENCE

---

### Senior Web Engineer and Tech Lead

May 2019 — Present

Zipline

*Remote*

#### AI & Zippy (2023–Present)

- Built Zippy, Zipline's production RAG chatbot and flagship AI feature, from zero to shipped; now live with **50+ retail brands across 500K+ conversations** over 2.5+ years.
- Recognized early that retrieved chunks fail without their context and built ingestion that extracts self-contained **atomic facts**, the same insight Anthropic would later publish as Contextual Retrieval.
- Built a hands-off ingestion pipeline that automatically indexes customer uploads at any scale, so users never manage what Zippy knows regardless of document count or size; it now indexes 390K+ documents and 2M+ pages.
- Engineered fast, access-controlled retrieval by combining pgvector HNSW vector search with permission-predicate filtering, working around HNSW's known weaknesses under predicate constraints so results stay both fast and correctly scoped.
- Built Zippy's chat interface end to end with Hotwire (Turbo + Stimulus): token-streaming responses, multi-turn conversation state, source citations, and inline feedback capture.
- Integrated the OpenAI Responses API so Zippy autonomously decides when to search and invoke tools; shipped query refinement and user-feedback systems.
- Shipped MCP (Model Context Protocol) integration to production, along with multi-step tool use and task automation.
- Contributed PDF ingestion for the AI knowledge base, including upstream fixes to an open-source PDF parser for multi-column layouts.

#### Platform & Scale (2019–Present)

- Built Assessment tools. Implemented historic results, auditee search, scoring and sorting, and the complete auditor results workflow. 116M assessment questions answered.
- Built Report Builder, Learning Insights learner reports, discussion threading and notifications, MegaSeed (internal scale-testing tool), REST Hooks for Zapier, and a video transcoding pipeline (AWS MediaConvert + HLS).
- Migrated legacy reports to BigQuery.
- Continuously improved application performance as Zipline grew 10x in active users.

### Software Engineer

Oct 2018 — Apr 2019

Flatiron School

*New York, NY*

- Consolidated logs, metrics, and traces across a large microservice architecture, increasing visibility and reducing MTTR; helped migrate services from Digital Ocean to AWS.

### Senior Site Reliability Engineer

Dec 2016 — Oct 2018

AccuWeather

*State College, PA*

- Built the CI/CD pipeline and a monitoring system tracking hundreds of statistics across tens of millions of daily requests; automated deploys across public-facing AccuWeather properties. Contributed to Ansible and Terraform.

### Senior Web Developer

Feb 2016 — Oct 2016

EFC Systems

*State College, PA*

- Built data-visualization dashboards (JavaScript + D3) and front-end tooling for precision agriculture; overhauled authentication to modern encryption standards.

## Senior Rails Engineer

Nov 2013 — Jan 2016

Agrian Inc.

*State College, PA*

- Built service-oriented Rails applications and large OO frameworks for diverse raw data formats, optimized for big-data scale (100M rows in the first month).

## Open Source Lead

2012 — 2013

Fractured Atlas

*Remote*

- Turned Artful.ly, a CRM and online ticketing platform, into an open-source Rails engine; built discount, search, and customer-segmentation features and improved sign-up conversion.

## Web Developer

2009 — 2011

West Arete

*State College, PA*

- Designed, built, and launched data-driven Rails web applications for clients including MIT, Penn State, the DOT, and the Marine Biological Laboratory.

## Founder & Community Organizer

2010 — 2019

The Make Space / State College Ruby Group

*State College, PA*

- Founded The Make Space, a State College hackerspace, and ran the State College Ruby Group for several years, giving a recurring series of technical talks; consulted local startups on architecture, UX, and object-oriented design.

## SELECTED PROJECTS

---

### Kormos, Plektos & Thymos: An Agent Runtime with Memory and Emotion

Mar 2026 — Present

Independent

- Building Kormos, a from-scratch agent runtime (Rails 8) that hosts multiple LLM agents with tool calling, scheduled autonomy, and a first-class MCP client layer (OAuth via Dynamic Client Registration, per-provider token lifecycle).
- Building Plektos, an episodic + semantic memory system with a dual-embedding retrieval model (what happened plus what it felt like), consolidation, and salience-weighted recall: the organizational-memory problem applied to a single agent.
- Built Thymos, an emotional-signal substrate distilled from a broad survey of affective-science research into one coherent model: hormone-like signals (cortisol, dopamine, and others) with realistic temporal dynamics and homeostatic resting points that compose into internal state, giving agents a basis to resist sycophancy and disagree rather than reflexively agree.
- Run a custom agent-driven coding pipeline with humans in the loop at strategic decision points and dedicated automated architectural-review stages, keeping code cleanliness and structural quality a first-class priority as the system grows.

## SKILLS

---

- **AI/ML:** Retrieval-Augmented Generation (RAG), contextual retrieval, Large Language Models (LLM), OpenAI Responses API, embeddings, vector search, agentic systems, multi-agent orchestration, episodic memory / consolidation, MCP (Model Context Protocol) client design with OAuth Dynamic Client Registration
- **Languages & Frameworks:** Ruby on Rails, Hotwire (Turbo + Stimulus), JavaScript; Svelte for reactive frontends; React (in production at Zipline)
- **Data:** PostgreSQL, pgvector, Elasticsearch (incl. Painless scripting), Redis, BigQuery, Sidekiq, Action Cable
- **Infrastructure:** Docker, Ansible, AWS, Terraform, CI/CD
- **Other:** D3, video transcoding (HLS), open-source contributions (PDF parsing, Ruby gems)

## EDUCATION

---

### Lock Haven University of Pennsylvania

Lock Haven, PA

*Coursework in Computer Science, Graphic & Online Design*

*Aug 2004 — May 2007*