

Gavin Carothers

Sebastopol, CA | gavin.carothers.name | github.com/gcarothers | linkedin.com/in/gcarothers

SUMMARY

I've built an engineering and product organization from two people to global scale, and I'm just as interested in reshaping an existing one, whether that means new architecture, new practices, or a new direction entirely. My background is in data analytics, machine learning, data publishing, and text processing. In less dry tones, I have both Ink and Code in my blood.

KEYWORDS & TECHNOLOGIES

AI/ML (LLMs, multimodal document extraction, Gemini / Vertex AI, MCP (Model Context Protocol), Transformer/Encoder, Ray / Anyscale), Python (Pyramid, WSGI), SQL (Postgres, Cloud Spanner), Google Cloud (Cloud Run, GCS), Java (Restlet, JAX-RS, Jena), JavaScript (d3.js, ESNext, React, NextJS, TanStack), CSS (SASS), Semantic Web (RDF, SPARQL, OWL2), XML, XSLT, and XQuery

EXPERIENCE

Co-founder & CTO | Transpire | Apr 2026 – Present

Technical co-founder of an AI platform that turns supply-chain correspondence into a live, reconciled view of what was ordered, shipped, received, and billed. Built multimodal extraction on Gemini and Vertex AI across email, PDFs, spreadsheets, and phone photos of handwritten packing slips, with per-line confidence scores routing uncertain reads to human review. Designed an append-only observation log in which every claim carries its source document and a fingerprint of the model and prompt that produced it, making re-extraction idempotent and human corrections durable as evidence rather than overrides. Built the reconciliation engine that matches observations into canonical orders and adjudicates conflicting sources, running multi-tenant on Cloud Run and Spanner with customer-specific behavior expressed as declarative profiles rather than forked code.

Lex Machina (LexisNexis / RELX) | Oct 2011 – Apr 2026

Web Architect → Engineering Manager → Director of Engineering → Sr Director of Engineering

Joined as an early engineering hire when the team was two engineers racking physical servers, and became Engineering Manager within a year. As the company's engineering leader, served as its sole technical representative throughout the acquisition process, including data rooms and technical diligence with prospective acquirers, through the sale to LexisNexis. Promoted to Director following the acquisition and later Sr Director, growing engineering and product into a 30+ person global organization and leading its integration into RELX. Owned the design, implementation, and deployment of the legal analytics platform ([limited demo](#)), which grew to cover 37M+ court cases and 370M+ docket entries across Federal, State, and Specialty venues, with data processing pipelines in Ray and Python. Applied machine learning to legal data years before the current generative-AI wave, and later drove RELX's first customer-facing MCP applications and production AI systems. Served as W3C Advisory Committee Representative.

Semantic Solutions Developer | TopQuadrant | Jan 2011 – Oct 2011

Responsible for web features and core database functionality in the TopBraid Semantic Web development platform. W3C representative.

Senior Software Engineer | O'Reilly Media | Jul 2007 – Jan 2011

Developed single sign-on system in Python using early OpenID technology. Migrated product management/lifecycle system from Perl/SQL to RDF. Designed and implemented semantics-driven electronic media distribution system. Designed and acted as lead developer for semantics-driven template system for catalog pages. Released several open-source Python libraries, e.g. <https://github.com/norcalrdf/pysemantic>

XML/Ebook Consultant | The First Church of Christ, Scientist | Jul 2007 – May 2010

Developed TEI-based mark-up for historical and religious texts, designed conversion process from legacy SGML and tagsoup content. Provided tool chain for converting new TEI into ePub and mobi digital book content.

Software Engineer | Banta (R.R. Donnelley) | Sep 2006 – Jun 2007

Python-based website framework improvements to and modernization of HP's DesignJet site and the HP Graphics Art portal. Designed a range of web services for print-on-demand applications, including summary of benefits for WellCare (Medicare Part D) and other insurance providers. Worked on new marketing campaign tracking software spanning email, print, and web.

Senior Software Engineer | Ames On Demand | May 2005 – Sep 2006

Built XML toolchain based publishing systems producing custom textbooks for Pearson Educational Publishing. Created reporting tools for Fidelity's internal custom publishing solution, built using Cocoon, MS .NET, and multiple SQL databases. Designed and implemented a new XML-based custom publishing solution using eXist (XQuery engine / XML database), Saxon 8 (XSLT), Relax NG + Schematron, and Apache Cocoon.

Web Systems Administrator | CS Monitor Newspaper | Jun 2004 – Apr 2005

Responsible for building tools and major publishing system development for a daily news website serving 55,000–190,000 unique visitors a day. Environment was an in-house Content Management System based on Apache Lenya and Cocoon, with WebObjects and Struts used for UI on editorial-facing applications. Managed day-to-day systems administration of the web farm and application servers, and developed the network architecture for major publishing upgrades.

Project Manager & Embedded Software Engineer | OpenTuit, Inc | Sep 2003 – Mar 2004

Designed and implemented embedded Linux-based firmware for an XScale CPU ultrasonic digital pen. Developed desktop software for the pen using Cocoa and Trolltech Qt3. Managed the technical project schedule and was responsible for the technical section of venture capital presentations. (Shipped many bankruptcies and lawsuits later as <https://www.engadget.com/2011/09/12/wacom-inkling-digital-sketch-pen-hands-on-video/> seemingly with the exact same software 7 years later.)

BOARD SERVICE

President, Board of Directors | Summerfield Waldorf School & Farm | May 2025 – Present

Board leadership for an independent PreK–12 Waldorf school and working farm. Fiduciary and governance oversight spanning finance, HR, and legal. Led the search and selection process for the school's All School Director, and drove adoption of the board's governance framework defining executive authority and accountability.

Non-executive Director | Sebastopol Entrepreneurs Project | Jan 2012 – Feb 2017

Educational charity providing business advising, co-working, and educational programs in west Sonoma County.

PUBLICATIONS

Turtle: Terse RDF Triple Language — <http://www.w3.org/TR/turtle/>

TriG: RDF Dataset Language — <http://www.w3.org/TR/trig/>

N-Quads — <http://www.w3.org/TR/n-quads/>

N-Triples — <http://www.w3.org/TR/n-triples/>

Systems and methods for using non-textual information in analyzing patent matters — <https://www.google.com/patents/US20140180934>

Legal analytics based on party, judge, or law firm — <https://patents.google.com/patent/US10133755B2/>

Legal timeline analytics — <https://patents.google.com/patent/US11748832B2/>

Featured in Forbes ASAP, March 2002 — https://www.forbes.com/asap/2002/0325/034_print.html