Maxwell Collins

AI Systems Engineer

West Palm Beach, FL · Remote (US) · hello@maxwellacollins.com · maxwellacollins.com · linkedin.com/in/maxwell-a-collins · github.com/macollins27

Summary

I build AI systems that run real businesses. An event venue's whole front office runs on my system, with 1,222 bookings handled; after the rebuild the venue's revenue grew from about $55,000 to more than $200,000 a year. Its AI receptionist cut calls handed to staff from 59.8% to 9.5% and raised lead capture from 11.5% to 76.2% across 857 recorded calls. A six-product property suite was rebuilt in 64 days, and an AI-run campaign helped a neighbourhood win its zoning fight. I build by directing organizations of AI agents rather than writing code by hand. I am looking for a role as an AI systems engineer, in-house or forward-deployed.

Scale of work

- **15,999 AI sessions in the record**, 4,760 started by me and 11,239 run by agents my sessions dispatched, with 34,103 messages typed by me. Earlier sessions were lost to the tool's retention, so every count is a floor. Single sessions have dispatched up to 113 agents and run for as long as 22 days; 53 ran 47 hours or longer.
- **18,858 GitHub contributions** in one year, including 12,925 commits, 2,215 issues and 616 pull requests across 141 repositories.
- **94 projects with full written engineering records**, naming 381 top-level systems and 2,636 systems and mechanisms in all.
- **Peak days:** 218,186 lines built in one day, 162 commits in one day, and 43 projects advanced in one day.

How I work

- **Done is decided before the work starts.** Every job begins with what its user will open, do and see, checked against real data with cases of opposite shape, and each check fails before the work begins. The work is judged against the user's result, never against the checks.
- **Organizations of agents build it.** I map the process with the people who do it, write the business rules in plain English as the authority for the build, and direct organizations of AI coding agents to build it, tests first wherever behavior is testable.
- **No builder signs off its own work.** An agent that did not build the change reviews it, then uses the running system as its user would, with real data, trying to break it. I judge the work by the evidence the machine produces: tests, gates and recorded proof.
- **Seven hard limits, and nothing else.** Agents are stopped at only seven kinds of action: rewriting history, destroying data, making something public, spending money, deploying, committing a secret and skipping commit checks. Every other guard informs and lets the work through, and agents make and report the technical decisions.
- **Earlier method.** Domain-Rules-Driven Development: the business's rules outranked the code, wrong work was made impossible to compile rather than reviewed out, an independent examiner the builders could not edit decided acceptance, and every recurring AI failure became a permanent block.

Experience

Founder and Principal, AI Systems and Operations
Code Rescue / Mac-Ops (self-employed) · West Palm Beach, FL / Remote · Aug 2025 – Present

I build custom software, automation and AI systems for businesses; the business now operates as Mac-Ops, which sets up and runs the systems a business uses to operate: phones, bookings, quotes, paperwork, records and owner reporting.

- **Operating system for an event venue.** Runs the venue's whole front office, taking over work four staff had done: 1,222 bookings and six figures in card payments, and the venue's revenue grew from about $55,000 to more than $200,000 a year. Public booking with a five-step wizard ending in a digital contract and deposit, payments, deposits, inspections and an AI receptionist, across four portals on one core (public booking, customer, staff and inspector). The database refuses a double booking, every change to a booking or payment goes through one door, a reliability engine makes sure no email, calendar event or payment step is dropped, and a fresh AI session audits production every morning and reports in plain English. 128,927 lines of application code and 78,816 lines of tests across 41 screens, 70 service routes and 21 data tables, governed by a 1,055-line engineering constitution, 26 CI guard scripts and 21 decision records.
  - **The server's operating system.** Lets an AI agent run the venue's production server safely, with every customer payment across 1,224 bookings reconciled to the cent in a restore drill: one door of 18 audited commands, an 11-check gate whose exit code is the only "done", 140 planted-failure files that attack its own guards, deploy rights that follow fresh proof of rollback, and about 500 tests; 11,520 lines in 87 files, built in three days and 43 commits through nine rounds of outside review.
- **Voice-agent platform.** Its receptionist for the venue cut calls handed to staff from 59.8% to 9.5% and raised lead capture from 11.5% to 76.2% across 857 recorded calls; removing the end-call tool brought voluntary call drops to 0%. AI creates, launches, tests, measures and improves phone voice agents, running four production receptionists on real phone numbers as version-controlled configuration; every change is proved on a staging copy before a caller hears it, and every improvement must move a business measure. 6,761 lines with 2,083 lines of tests (134 test cases).
  - **Voice-AI quality platform.** Scores every call transcript with a cascaded cheaper-to-stronger AI judge, mines failure patterns across tenants and proposes prompt fixes gated by cost and safety; a five-service stack with an eleven-table schema and four work queues, built under 1,413 tests.
  - **AI agency platform.** Bills real clients: one booking core serves a public web form and an AI voice agent, ingests signed post-call webhooks into a multi-tenant dashboard isolated at three layers, and books into Google Calendar, behind an 80% coverage gate.
- **Six-product property suite.** Rebuilt in 64 days by AI agents I directed: owners, inspectors, service companies and property managers work on one shared record of each property. Built to an acquisition due-diligence standard from a 500,000-line legacy ecosystem, governed by 3,862 numbered rules, fresh-agent review, a 19-step gate and a browser-testing system that must prove it can see; a

forensic audit of the legacy code recovered features at risk of being dropped before the rebuild carried forward. 203,526 lines of code, 260,878 lines of tests, 13,380 test cases and 106 data tables; 1,217 commits, 1,207 of them in 64 days.

  - **A permanent identity for every US land parcel.** One record per plot that survives every split, merge and renumbering, on the RESO standard, with an uneditable ledger, provisional identities for new construction and portfolios for multi-parcel institutions; 17,520 lines of code, 6,970 lines of tests, 17 new tables and 28 reviewed layers in six days.
  - **Autonomous browser-testing system.** AI testers driving real browsers found 920 distinct defects in 25 runs, 427 critical or high; each run is trusted only after the testers find six planted defects, and every finding gets two independent rulings, on the code and on the specification. 15,350 lines, 67 scenario files and five agent roles; 33 permanent countermeasures guard every run.
  - **A software company of AI agents and its supervisor program.** Bounded roles, fresh reviewers and a read-only CTO agent, with a 7,983-line supervisor that checks every agent's report against disk and chains the results into a record the owner can re-verify; 382 agent dispatches and an unattended overnight build of nine commits in 81 minutes.
  - **Property data API, sign-on hub and ecosystem monorepo.** A property-condition backend that gives every building one permanent history buyers and insurers can verify and the owner controls, planned for 500 million properties, with an append-only signed observation ledger, a default-deny permission kernel and 6,469 tests against live PostgreSQL; a single sign-on gateway for twelve products; and a monorepo of 11,600 files whose fourteen-tool fact server makes a 149,000-line API machine-queryable.
  - **Owner-portal rebuilds and four-agent engineer.** A 153,000-line legacy property-management app rebuilt into a modern platform with a 75-model domain schema and a Claude assistant grounded in each owner's data; and a four-agent build-and-review tool backed by 830 unit tests.

- **Multi-tenant proposal SaaS for service businesses (Proposal Kernel).** Writes, prices, sends, tracks, gets signed and renews a service business's proposals, and its pricing reproduced a real signed service agreement line by line. Each business declares its assets, cited service rules and rate card; leads arrive by live sync from its field-service system or from documents Claude reads; the product derives every required service, prices it to the cent, has Claude write the narrative with no access to prices, routes approval, delivers a tracked web proposal and identical PDF, ranks follow-ups by engagement, takes e-signature and generates renewals. One declaration per fact produces the database, its security, audit trail, validation and screens; illegal states fail to compile; a second real customer is installed so hard-coding fails. 75,003 lines with 54,578 lines of tests (1,804 tests); 963 agents dispatched across 209 sessions.
  - **The declare-once kernel.** Business screens show every field of a record from one declaration of each fact: all 48 of the largest real record's, with 8 compile-rejection proofs and 15 demonstrated refusals.
  - **The independent examiner.** Scores a fresh copy of the product against a real signed contract on 88 clause-by-clause checks, from its own repository where the builders cannot write; 214 machine receipts, with its status page computed from them.
- **Construction bid analyzer and BidLedger.** On a real 653-page school-district bid it deleted 509 AI claims whose quotes were not on the page and turned the rest into 47 cited actions; a six-stage pipeline turns 600+ page bid sets into citation-verified risk reports and live-formula Excel takeoffs, recorded in a hash-chained audit log. Sold as BidLedger at $249 per bid, with Stripe billing, inbound-email intake and CRM sync.
- **AI technical due-diligence engine and Clear Report.** Compresses a 4–6 week codebase audit into hours with 6–10 specialist agents over a six-stage pipeline, every finding cited, sold as a CLI, library and SaaS from $79 per scan with nearly 2,000 tests; and a SaaS that turns environmental inspection reports into plain-language findings against their safe limits.
- **Control plane for a 1.78-million-line legacy estate.** AI agents work on an eight-application, twelve-year estate without being able to damage it or assert anything uncited: ten question-answering services over a 71,600-node code graph answer by exact file and line, five independent blocks stand between any agent and production, review by a different agent is enforced in the database, and a doctor proves each guard by attempting the forbidden act. The owner gets a 24-metric project page on which no figure is produced by an AI. 19,810 lines of application code, 20,593 lines of tests (1,060 test cases) and 141 commits.
  - **Receipts, the project oracle.** Answers a manager's questions about the estate with a file and line, or "can't tell": every AI claim must become a citation a model-free verifier confirms against frozen evidence or deletes; its core was built in two days under twelve fresh reviews. 12,956 lines of code and 7,084 lines of tests.
  - **The estate wireframe studio.** AI drew 23 buildable screen designs for the estate's four products in eleven days, every element traced to a verified reference or marked as proposed; 53,823 lines.
- **A company's operations, run in Notion.** Runs unattended beside the administrator who edits it daily, preserving every one of the administrator's 1,564 edits, with a nightly copy of the workspace. AI miners worked through millions of a long-established company's documents, its service system (554 customers, 5,145 work orders) and its historical database, producing 35 seats across five companies and 193 recurring duties, each tied to a primary file; a drill rebuilt all 311 rows exactly. 22,203 lines of Python and 422 tests.
  - **Email chief of staff and executive assistant.** A 24/7 agent that follows every matter across six inboxes and three businesses at $0 API spend, and a Microsoft 365 assistant that triages 500+ emails a day and books meetings from a one-word yes.
  - **Dashboard platform and broker CRM intelligence.** A multi-client dashboard design around one rule, that nothing a person is not meant to see reaches them (5,744 lines, five rounds of independent review); and an ETL that turned 1,446 free-text broker notes and 23,000 de-duplicated contacts into per-contact intelligence profiles.
- **Customer-acquisition engine (Front Desk).** Screens local businesses and builds each one that qualifies a complete personalised 22-page website for about 16 cents. No outreach message can exist unless six proofs hold: no invented fact, no opted-out recipient, no missing legal field, no stale approval, no double send, no malformed data. 19,203 lines of application code, 9,822 lines of tests (548 test cases) and a 14-state prospect lifecycle.
  - **The full Front Desk platform.** Wins a local service business with a working demo, then runs its phones, bookings, payments and contracts under nine machine-enforced invariants such as "the AI never holds the pen on money"; in four days the design became 26 domains, 352 procedures, 2,008 numbered rules and 149 data tables.
  - **Lead engine and outreach engine.** Delivers ranked leads every week, every reason linked to public evidence; and an engine that ranked 761 local-business leads from public reviews and runs daily sourcing and email sequences across 15 metro areas.
- **Evidence-led community campaign.** Helped a neighbourhood win its zoning fight: the zoning board reversed the approval the residents had appealed. Every published claim traced through a machine gate to a page of a public filing, an AI campaign manager was

held to what a neighbour can grasp in seconds, consent sat in an append-only record, and every sign-up flowed to the board's Notion workspace. 9,307 lines of application code, 2,228 lines of tests and 12 public pages.
    - **A repeatable campaign service.** Turned the campaign into a service: an AI campaign manager researched public records into a verified index of 147 active public fights from 757 source records and builds an organiser a fully cited campaign system; 43,858 lines with 17,750 lines of tests (598 test cases) and a 37-check gate.
- **Owner's representative inside every AI session (owner-rep).** Watches the coding agent for its owner: builds the objective from the owner's messages alone, flags work dropped or claimed without proof, writes the correcting message so one press sends it, and on request checks each claim against the real files and tests. 11,719 lines of application code, 32,814 lines of tests (759 test cases) and 791 commits.
- **AgentsView.** Every agent on my machine starts with instant search over what was said, decided and ruled out across 15,999 sessions. Built additions to an open-source session archive, carrying it 2,914 commits past the original (2,894 mine) with 51,683 lines of application code and 56,498 lines of tests; an always-on service with its own login and desktop app.
    - **corpus-brain.** A cited memory for every AI agent across 499,050 passages of transcripts and repositories; search time fell from about 16 seconds to 0.011 seconds, and nothing surfaces that cannot be cited.
- **The Forge.** Every project it stamps starts with an AI agent's shortcuts already blocked; later projects are stamped from it or carry its control kit, and its ledgers have recorded 371,817 refusals. 13 fail-closed controls proved by 312 must-block and must-allow examples, four starter packs; first commit to independent sign-off in one day.
- **The Operator Doctrine, vision-to-plan and governance plugins.** A published method for getting true work out of AI agents (CC BY 4.0; 151,750 words across its files, built from 145 recorded failure behaviors and 96 patterns); a plugin that turns a plain-English idea into an approved plan before code, which planned three products; and plugins that stop an agent writing before approval, shrinking the job or stopping without proof. Also a fifty-page governance manual that cut a booking platform from three or four AI sessions per feature to one.
- **Websites for my own businesses.** Sites that rank locally and turn visits into calls and bookings, with two AI phone agents answering, scoring 95 to 100 on Lighthouse; 280,184 lines across nine repositories.

Founder
Portstead Facility Services · South Florida · Nov 2026 – Present

- Founded a commercial cleaning contractor for property managers, facilities teams and institutional buyers, with two divisions (recurring janitorial, and dumpster, trash-room and waste-area sanitation); defined its services, standards and brand, and built its website and AI phone agent.

Founder and Healthy Home Inspector
Certified Environment · Tampa, FL · May 2025 – Present

- Founded and run a licensed healthy-home inspection company: home, healthy-home, 4-point and wind-mitigation inspections and mold, air and water testing, with laboratory coordination and science-backed reports. Runs every part of the business, from inspections to marketing.

Founder and Project Manager
MAC Structures · Tampa, FL · Jan 2025 – May 2025

- Founded and ran a home renovation company covering estimating, scheduling, subcontractors and customer follow-up: restored 15 flood-damaged homes, built 5 home additions and about 6 decks covering 4,000+ sq ft.

Sales Specialist
Lakeland Genesis (Lakeland Automall) · Lakeland, FL · Jul 2023 – Jan 2025

- Top salesman in my first month and for my first four months, while new to auto sales; top 10% of the team with the highest closing ratio (about 60%) and highest customer satisfaction (98%); $7.2M+ in annual sales, about 14 vehicles a month.

Environmental and Construction Project Manager; Real Estate Manager
An environmental consulting firm and a construction company · Long Island, NY · May 2019 – Jul 2023

- Managed a portfolio of 20+ residential and commercial construction projects valued at more than $10M, led construction totaling more than 100,000 sq ft as foreman, and directed a crew of 20+ on a warehouse build.
- Ran compliance inspections, remediation and underground storage tank work of up to 30,000 gallons under New York State DEC guidelines.
- Managed a 40-property portfolio (multifamily, warehouse, retail and storage) and more than 20 renovation projects above $1M combined, led the move to property-management software, ran the group's event venue and served as hiring manager.

Product Manager and Systems Designer
A property technology company · Long Island, NY, and remote · Jun 2016 – 2023, part time

- Designed a connected four-product suite of property software for owners, inspectors, service companies and managers, wrote the specifications and roadmap, managed a development team of eight under Scrum, and onboarded more than 200 enterprise clients in construction and real estate. Rebuilt by AI agents I directed as the six-product property suite above.

Education and licenses

B.A., Economics with Financial Applications, Southern Methodist University, 2019 · Florida Licensed Home Inspector, HI18109 · InterNACHI Certified Professional Inspector · OSHA 30-Hour Construction Safety

Skills

- **AI systems:** Claude and Claude Code, Codex, organizations of AI agents, MCP servers, evaluation and test cases for AI output, citation-gated retrieval, voice agents (ElevenLabs, Twilio)
- **Software:** TypeScript, Next.js, React, Python, PostgreSQL, Docker, Stripe, Cloudflare, CI gates, end-to-end browser testing
- **Business systems:** Notion, Google Workspace, Microsoft 365, workflow automation, APIs and webhooks, email delivery
- **Practices:** process mapping, specifications, business rules, monitoring and alerts, audit trails, training non-technical staff