Robert R. MacKay

Chief Technology Officer · Software Architect · AI-Native Platforms

Salt Lake City, UT • rrmackay@gmail.com • 801-787-0504 • linkedin.com/in/robert-r-mackay

SUMMARY

Hands-on technology leader and software architect with **25+ years** turning emerging technology into shipped, production platforms across full delivery lifecycles — TDD, Agile/Scrum/Kanban, and modern GitOps deployment. Currently CTO of an AI-native marketing-intelligence platform built on agentic AI and semantic retrieval; previously drove an enterprise SaaS cloud migration to **80–90% customer adoption** and delivered real-time trading, fintech, public-safety, and enterprise integration systems. Equally comfortable setting technical strategy and writing the code that proves it.

CORE EXPERTISE

AI & DATA

Agentic AI & LLM orchestration (CrewAI, Anthropic, xAI server-side agent tools), streaming tool-use loops, semantic retrieval (ChromaDB, local embeddings), ML/MDM evaluation (DataRobot, SageMaker, TAMR)

ARCHITECTURE

AI-native & microservices platforms, multi-tenant SaaS, event-driven systems, API & integration design

CLOUD & DEVOPS

Kubernetes, Docker, Terraform, Helm, AWS, Azure, GitOps CI/CD, schema-migration pipelines, Telepresence

LANGUAGES

TypeScript · Next.js/React · Node.js · Go · C# · Java · Python · PHP

DATA & SECURITY

PostgreSQL, Drizzle ORM, Oracle, Redis, MariaDB · OAuth/OIDC, Keycloak, Azure AD/SSO, DISA

LEADERSHIP

Technical strategy, multi-team & onshore/offshore delivery, Agile/Scrum/TDD/OneFlow, mentoring

EXPERIENCE

Chief Technology Officer — Expona AI

Jul 2023 – Present

Architected and shipped a **20-service, 30-repository** AI-native marketing-intelligence platform — a Next.js/TypeScript web app, Python CrewAI generation services, and a dedicated TypeScript agent runtime — running on Kubernetes across dev, stage, and prod, and operated end-to-end by a **3-engineer team**. Users create a project from nothing more than a company name or URL, and the platform fills it in through direct LLM processing of loaded content and dynamic selection across specialized agent crews.

- Built **20+ CrewAI agent crews** that generate buyer personas, competitive intelligence, and campaign content on demand, dynamically selecting the right crew for each request.
- Delivered semantic retrieval through a ChromaDB-backed Knowledge Service so agents reason against embedded, workspace-scoped context rather than raw prompts.
- Designed a provenance & “composition-receipt” layer so every AI-generated claim is traceable to its sources — with field-level provenance and anti-clobber locks that protect human edits from automated regeneration.
- Enforced workspace-scoped multi-tenant isolation on every read and write, keeping each customer’s data cleanly partitioned across the platform.
- Stood up GitOps CI/CD with a schema-migration pipeline (Postgres branch previews, destructive-change lint gates, gated dev→prod promotion) and Redis pub/sub for multi-pod real-time updates.
- Shipped a vendor-agnostic publishing layer (social + email) behind a single provider interface — verified end-to-end from chat to LinkedIn and to the inbox.
- Built in-product support on GitHub as the system of record — tickets open as issues with priority, category and status encoded as labels — with Telegram push alerts gated to high and urgent priority so the channel stays a pager rather than a firehose, and designed to fail open: a notification outage or missing configuration degrades to a logged warning and never blocks a ticket from being filed.
- Established developer-first practices — codemods, Telepresence-based live debugging, and branch/tag deployments with TLS across every environment.

Senior Software Consultant — Utah State Board of Education

Aug 2023 – Nov 2024 · concurrent

- Built C# financial-reporting systems supporting statewide school-district and charter-school compliance, with Azure DevOps pipelines for build, release, process management, and repository governance.

Senior Software Architect — CHG Healthcare

Jul 2021 – Oct 2022

Focused on data-quality initiatives, running a series of proof-of-value projects to vet vendor offerings in the master-data-management domain and mapping where ML fits versus where purpose-built MDM tooling wins.

- Led MDM/ML proof-of-value evaluations across DataRobot, AWS SageMaker, and TAMR — ultimately selecting TAMR — clarifying which data-quality problems each tool is genuinely suited to.
- Acted as team architect across three Scrum teams and architected a Kafka-based data-hub project.
- Authored “zero-to-hero” workshop documentation onboarding new developers onto the platform and systems to build new projects.

Senior Software Architect — Earnest

Apr 2020 – Mar 2021

Product-vertical architect on the monolith-to-microservices redesign at a fintech lender spun out of a large student-loan processor bringing new technology to market.

- Re-architected the online loan application into independent pages and data submissions, so the flow could be reordered with minimal effort — which the UX team used to A/B-test scenarios and lift conversion rates.

Senior Software Architect — Assure Services

Jul 2019 – Jan 2020

Joined into an environment with no coherent infrastructure strategy and unpredictable releases; defined and delivered a modern Kubernetes-based platform and development process end to end.

- Replaced ad-hoc AWS EC2/Fargate releases with an EKS + Terraform + Helm platform providing isolated feature, stage, and production environments; automated deployment via GitLab runners in the dev cluster.
- Instituted a OneFlow development process tying JIRA, GitLab, and cloud test systems into a seamless flow, turning unpredictable releases into repeatable ones.
- Implemented Azure AD OIDC SSO for the SaaS application and set PaaS/SaaS strategy with a path to Azure-hosted Kubernetes.
- Managed a 3–6 person onshore/offshore team and built a career-development plan around Pluralsight training and relevant certifications.

Senior Software Engineer — Omadi, Inc.

Jul 2018 – Jul 2019

Full-stack role building road-incident-management systems across web, mobile, and high-performance backend services.

- Built complex multiform mobile applications in React, React-Native, and Titanium, with end-to-end test automation via Appium and Selenium WebDriver (Node and Java) targeting device farms.
- Developed backend services ranging from gRPC in Go and Java for high-performance inter-cluster communication to Node and PHP REST services for generic access — all deployed on AWS Kubernetes with Jenkins, GitFlow, and Bitbucket pipelines.

Senior Software Engineer — J.P. Morgan

Mar 2017 – Jul 2018

- Built a high-speed FX trading platform (Spring Async) streaming real-time market data alongside on-demand quotes from multiple providers into existing desktop trading systems via a web-based stream display.
- Automated the pipeline with Maven, Gradle, Jenkins, Helm, and Chef targeting Kubernetes clusters in a private cloud; secured with Keycloak/OAuth and hardened network architecture.

Senior Software Architect — Ellucian

2007 – Jan 2017

Owned product analysis and future-vision definition for a next-generation higher-education platform, from strategy and prototypes through delivery, evangelism, and cloud transformation across geographically dispersed teams.

- Set the technical vision and evangelized it to the existing customer base, reaching **80–90% adoption** of the new platform over several years.
- Led the conversion of the enterprise software to cloud compatibility on AWS services, then moved into a cloud-architect role designing and building new cloud-native services.

Chief Technology Officer & Co-Founder — AlertFM

2005 – 2007

- Delivered a wireless public-safety alerting platform for the Dept. of Homeland Security and Mississippi Emergency Management — engineered to run independent of traditional communications and power to reach first responders and the public during crises.
- As founder, owned the product end-to-end: contracts, office and team creation, hiring and management, architecture, and customer acceptance by local, state, and federal contract managers.

SELECTED PROJECT — PUBLIC SOURCE

GoReactChat — github.com/notbobutah/GoReactChat

2026 · Go, gRPC, Next.js

A gRPC/Connect streaming chat service in Go with a React client, running on Kubernetes behind TLS. Built as verifiable evidence rather than a portfolio piece: the code is public, the deployment is live, and the service answers questions about this résumé using it as its grounding corpus.

- One **connect-go** handler serves gRPC, gRPC-Web and Connect from a single wire contract, with server-streaming turns over h2c and TLS terminated at the ingress.
- Hand-written streaming tool-use loop against the Anthropic Go SDK — including a recall guard rail that discards a model's "let me check" preamble once a retrieval tool returns, so the reader sees the answer rather than the narration.
- Retrieval-augmented grounding over locally generated embeddings, with per-document-kind balancing so one long document cannot crowd out the rest of a corpus.
- A second agent watches Go, gRPC and Protobuf releases with its tool loop executing **server-side on xAI**: a single Responses API request declares the tools and the model runs roughly fifteen web searches, reads and iterates there. No local agent loop and no additional service — the application subscribes and is pushed the result over a streaming RPC, because a scan takes about a minute.
- Because that agent bills per tool call rather than per token, spend is bounded by frequency: it scans only while a client is subscribed, at most once per interval, one scan at a time however many are watching, and the last result is persisted so a restart restores instead of rescanning.
- Total spend capped service-wide and persisted in Postgres, with per-request and per-conversation ceilings — the substitute for authentication on a deliberately public, sign-in-free deployment.
- Kubernetes manifests committed in **deploy/**, so the infrastructure is reviewable rather than described: hardened pods (non-root, read-only root filesystem, dropped capabilities, resource limits, health probes), an nginx ingress splitting one host across both services with cert-manager TLS, and a workflow that applies them and smoke-tests through the ingress.

EARLIER CAREER

Contract Security Architect, SPAWAR 2004–05 — HR/career-management systems on PeopleSoft integration broker across distributed systems; cross-domain security architecture and documentation to DISA approval. · **Development Team Lead**, SirsiDynix 2002–03 — defined the engineering process and led a J2EE next-generation library-automation platform (JBoss, Oracle, Struts, XML/XSL) on a 70-engineer program. · **Team Lead**, Verticore 2001–02 — object-oriented plant-management suite (VPS) improving operating efficiency in oil & gas refineries; UML-driven Rational process on Java/XML/XSLT and the Oracle stack. · **Director, Integration Solutions**, Ventro 1999–2001 — distributed data-integration solution (web services, JMS, redeployable connectors) for healthcare and life sciences; awarded U.S. patent for embedded-transaction-history data transport.

PATENT

Hub & Spoke Architecture and Methods for Electronic Commerce
US PCT/US2001/003087 — B2B commerce integration via standardized XML/EDI and ERP interfaces.

EDUCATION

Utah Valley University
Associate of Science — Electronics, Electrical Engineering & Mathematics