

Praveen Sehgal

Applied AI Engineer | Forward Deployed Engineer | Solution Architect

Portfolio and case studies: <https://sehgalp.ai>

Washington, DC | 703-477-4001 | sehgalp@me.com | [linkedin.com/in/praveensehgal](https://www.linkedin.com/in/praveensehgal)

U.S. Citizen | Public Trust (active) | TS/SCI Eligible

PROFESSIONAL SUMMARY

Applied AI engineer and solution architect with a 25+ year career architecting, building and deploying enterprise systems for the world's most demanding organizations - NASA, the Department of Energy, FAA, NOAA, the World Bank, JPMorgan Chase and Johnson & Johnson. Started as a developer writing PL/SQL and Java, and grew into an architect who embeds directly with customers and owns delivery end to end. Over the last three years I have woven AI into everything I build: Large Language Model (LLM) pipelines inside air-gapped government environments, RAG-powered SaaS products, multi-agent automation platforms and models fine-tuned on domain data. I work the way forward-deployed engineering works - understand the customer's hardest problem, prototype against their real data, integrate with the systems already in place, and own the path into production. I hold an active Public Trust clearance and am eligible for TS/SCI. Detailed case studies for the systems below are at <https://sehgalp.ai>.

AI PRODUCT AND PLATFORM ARCHITECTURE

Founder and Solution Architect | Kakoi (kakoi.ai) - Sovereign AI Coding Environment

08/2026 - Present | Remote

Commercial AI SaaS product: a self-contained, air-gapped-by-default AI-assisted development environment for regulated engineering teams. Per-seat subscription targeting federal and SOC 2-regulated clients.

- Designed and am building a commercial AI SaaS from scratch - a sovereign AI coding environment delivered as a multi-container Docker Compose stack: agentic coding harness, local LLM inference (Qwen3-Coder), GitLab CE, Oracle Database 26ai Free, ORDS/APEX, documentation RAG, credential broker and immutable audit store.
- Architected the valved egress model: zero data leaves the enclave by default; a single policy gateway redacts payloads, requires per-call developer consent, and immutably audits every authorized outbound request - designed for federal and SOC 2 buyers who cannot tolerate data exfiltration.
- Built the agentic harness with first-class Oracle and APEX capability: SQLcl tool access, APEX export and import workflows, RAG over product documentation, and grammar-constrained PL/SQL generation that prevents syntactically invalid output from reaching the database.
- Engineered the context layer rather than the prompt layer: deterministic context assembly under fail-closed budgets, provenance-tracked retrieval, typed intent classification ahead of the model turn, and immutable graph pinning so agents reason against a single verified view of the codebase instead of drifting mid-session.
- Enforced deny-by-default tool policy and end-to-end auditability, with grounded change planning validated against live system state before any material operation executes.
- Commercial model: per-seat subscription, self-hosted, targeting federal systems integrators and enterprise Oracle shops.

AI Solution Architect - Air-Gapped LLM Pipeline | Sandia National Laboratories / DOE-NNSA

02/2026 - 07/2026 | Albuquerque, NM (Remote) | Program: Financial Systems Modernization

Full-stack AI pipeline engineering in a zero-internet, air-gapped government environment - the exact deployment scenario federal AI deployment work handles daily.

- Designed and shipped a fully offline six-stage agentic pipeline (parse, schema mapping, overview, parallel page and inventory generation, assembly) that reverse-engineers legacy Oracle E-Business Suite OAF code and produces production-ready APEX migration design documents - zero cloud APIs, all local inference via Ollama and Apple MLX on a Mac Studio M3 Ultra with 256 GB unified memory.
- Fine-tuned 26B-30B parameter open-weight models (Gemma 4 26B-A4B MoE, Qwen3-Coder-30B) with PEFT/LoRA - rank-16 adapters on attention projections, approximately 15M trainable parameters - under both HuggingFace Transformers (MPS bf16) and Apple MLX. Achieved a 67% evaluation-loss reduction.
- Expanded scarce labeled training data 7x by generating synthetic instruction pairs from curated gold examples, resolving the binding constraint on model quality; managed overfitting through checkpoint selection and validation-loss monitoring.
- Built an automated evaluation harness with a weighted multi-dimension benchmark (structure, label accuracy, sample data fidelity, inventory completeness) and a 90% pass threshold, preventing regressed adapters from ever entering the production pipeline.
- Engineered prompt templates with slot-based structured output and Jinja2 assembly to obtain reliable, deterministic generation from local models at production scale.
- Navigated federal AI governance and model-approval compliance at a national laboratory; re-architected the pipeline from a prohibited model family onto the approved Gemma 4 family (including EmbeddingGemma for embeddings) with zero capability loss.
- Delivered 5,000+ lines of implementation-ready page-design documentation via parallel LLM generation with human-in-the-loop verification against a requirements traceability matrix, cutting legacy-code documentation effort from weeks to days.
- Packaged the entire toolchain - inference, fine-tuning, synthetic data generation and automated evaluation - for air-gap transfer with reproducible offline installs, running entirely on on-premises hardware with zero external API calls.
- Alongside the AI work, re-platformed a production Oracle EBS OAF non-PO payment application (CHREQ) to Oracle APEX at full functional parity: seven-status approval workflow, four-role security model mapped to EBS financial responsibilities, distribution-line-driven amounts and legacy-parity audit history across 14 screens, backed by a 28-item requirements traceability matrix.
- Built an autonomous 53-check data-verification harness for production handover and resolved all first-run failures the same day.

AI Architect and Lead Developer - SaaS Product | Narabi (narabi.io), Tsubaki Nakashima

07/2026 - Present | Remote

| *Greenfield AI SaaS - concept to pilot in under 8 weeks.*

- Architecting a multi-tenant AI SaaS product replacing a 700,000-formula Excel planning workbook with real-time, AI-assisted production scheduling for a global precision-manufacturing client.
- Embedded a Claude API conversational assistant over live production data using RAG on Oracle 23ai AI Vector Search - planners query schedules, material constraints and work-order status in natural language rather than through reports.
- Built a containerized ERP integration agent (Python, Docker) syncing QAD ERP via Progress ODBC on 5-15 minute incremental polls across inventory, work orders, sales orders and forecasts.
- Reduced the client-side footprint to outbound HTTPS plus a read-only ERP account, turning enterprise security review from a multi-week project into a single conversation - a deliberate architectural choice to remove the most common enterprise deployment blocker.
- Delivering an Excel-grade planning experience: Next.js frontend with AG Grid Enterprise driving a 270-column editable planning matrix, Entra ID SSO, and per-customer isolated instances on Azure.
- Designed the PL/SQL netting and allocation engine as retained product IP, with a weekly enhancement-shipping cadence and a documented handoff model so the client owns operations.

AI Architect - Sales and Operations Planning Platform | Mitooshi (mitooshi.io)

08/2026 - Present | Remote

| *Second SaaS product on a shared data platform - forecasting and capacity planning.*

- Built statistical demand modeling with size and gauge aggregation, finite-capacity modeling, and a rolling twelve-month raw-material forecast, backtested to 4.81% MAPE at full coverage before planners were asked to trust it.
- Architected as a separate schema with read-only grants against the planning platform, so the forecasting system can never write to the system it forecasts from.
- Enforced correctness through automated reconciliation gates - tonnage balancing to 0.00 t and monetary values to six decimal places across the cell set - required to pass on every build rather than checked by inspection.
- Distilled 75 requirements from the customer's own planning spreadsheet rather than from interviews, surfacing the operational rules nobody thinks to state out loud.

AI Architect - Multi-Agent Automation Platform | Global Manufacturing Group (NN Inc.)

Ongoing | Remote

| *Multi-agent orchestration and document intelligence across a multinational manufacturing operation.*

- Designed a multi-agent orchestration system with policy-based routing, tool invocation and lifecycle observability for manufacturing domain workflows.
- Engineered document-intelligence agents combining LLM-powered retrieval (RAG), OCR extraction and agentic reasoning to automate invoice and purchase-order processing across 7 global ERP systems.
- Placed human-in-the-loop validation specifically on the low-confidence extraction tail - where document automation actually fails - rather than routing every document through review.
- Built AI provider abstraction layers enabling integration across Claude, GPT-4 and Gemini, with evaluation harnesses measuring accuracy, latency and cost so models can be swapped and A/B tested without application changes.
- Implemented agentic supply-chain workflows applying predictive models to demand forecasting, inventory optimization and anomaly detection.
- Deployed containerized AI agents on Oracle Cloud with an event-driven architecture, enabling real-time processing for manufacturing operations across North America, Europe and Asia.

FORWARD-DEPLOYED PRACTICE

| *How I work with customers, not only with systems - the half of delivery that determines whether anything survives.*

- **Discovery from operations, not workshops.** Requirements taken from what people actually do - reading the live spreadsheet, the production system and the exceptions. At two manufacturing clients this meant distilling 75 requirements directly out of the customer's own planning workbook.
- **Traceability from requirement to delivered screen.** Requirements traceability matrices maintained through delivery (28 items across 14 screens on the most recent federal program), so acceptance is a checklist rather than an argument.
- **UAT as a feedback loop.** Built and ran a reviewer-feedback portal managing UAT backlogs across two concurrent federal modernization projects, rather than treating UAT as a sign-off ceremony at the end.
- **Verification automated before handover.** Built a 53-check autonomous data-verification harness for production handover; all first-run failures resolved same day.
- **Security and governance as design inputs.** Navigated federal AI model-approval processes that constrained which models could be used at all, and designed integration surfaces (outbound HTTPS plus read-only accounts) specifically to shorten enterprise security review.
- **Handover so the customer owns it.** Documented handoff models, 30,000+ lines of architecture guides, API specifications, QA test plans and deployment runbooks on a single federal program.
- **Technical leadership.** Led teams of up to 9 engineers and consultants; managed change control, code review and production support organizations at enterprise scale.

FEDERAL SOLUTION ARCHITECTURE

Lead Solution Architect - ERP Modernization | U.S. Department of Energy / Accenture Federal Services

09/2022 - 02/2026 | Washington, DC | Program: BPMI, Naval Nuclear Propulsion Program

| *Full lifecycle solution architecture for a mission-critical federal program serving the U.S. Navy.*

- Designed and led delivery of a 77-page APEX application managing military supply chain operations - sales orders, transfer orders and shipment tracking. Owned architecture, technical scoping, development and stakeholder engagement for Bechtel Plant Machinery (BPMI) under the Naval Nuclear Propulsion Program.
- Architected 15+ REST API integration packages with bidirectional synchronization to Oracle Cloud ERP (FSCM, HCM, CRM) using a staging-table pattern with error handling, retry logic and reconciliation across DEV, UAT and PROD environments.
- Built a custom approval workflow engine replacing Oracle BPM - sequential and parallel approvals, vacation delegation, group-based routing, automatic escalation and a complete audit trail.
- Built the federal financial audit (FFA) cost allocation calculation engine handling PO, TO and SO allocations across fiscal years with Oracle Projects integration.
- Resolved a critical production defect affecting 11,700+ records by root-causing accounting date versus expenditure date logic in federal procurement views, restoring accurate cost allocation.
- Produced 30,000+ lines of technical documentation: architecture guides, API specifications, QA test plans and deployment runbooks.

Solution Architect | NASA Jet Propulsion Laboratory

02/2014 - 03/2015 | Pasadena, CA

- Designed and built the Institutional Business Forecast (IBF), a 20-year budget forecasting system integrated with Oracle Financials across AP, AR and GL modules; served as technical lead and lead developer.
- Built live SOAP-based API integrations with UPS and FedEx for real-time price quotes and label printing, plus integration with proprietary warehousing and receiving software at JPL.
- Designed a custom single sign-on solution integrating with JPL infrastructure, and developed custom APEX plugins to satisfy strict institutional requirements.

Solution Architect - Cybersecurity Data Platform | Federal Aviation Administration

04/2020 - 10/2023 | Washington, DC

- Architected the ATOM system, a nationwide cybersecurity data collection platform covering 500+ users across 160+ FAA systems, giving system owners and auditors a consolidated operational view.
- Implemented custom PIV card authentication via SAML2 to meet federal identity requirements.
- Built a full lifecycle workflow engine handling 2,000+ data attributes, an ad-hoc reporting framework and a custom notification framework.

Solution Architect - Cloud Re-Architecture | National Oceanic and Atmospheric Administration

09/2020 - 08/2021 | Maryland

- Led the complete re-architecture of all five NOAA Fisheries lending programs onto Oracle Cloud, replacing a system more than ten years old; technical lead for both cloud infrastructure and applications.
- Delivered 8 APEX applications, a custom messaging platform, and a document management framework indexing millions of PDFs.
- Implemented custom OAuth2 authorization across the program suite.

Solution Architect - Migration Standards | The World Bank, Treasury Division

03/2020 - 04/2020 | Washington, DC

- Set infrastructure standards and design patterns for 70+ Oracle Forms to APEX migrations across the Treasury division.
- Designed custom themes and four authentication schemes - Azure, OpenID, OAuth2 and SAML - plus ORDS SODA collections for JSON document storage.

Solution Architect | U.S. Department of Energy - Mission Support and Test Services

04/2017 - 11/2019 | Las Vegas, NV

- Delivered P-Card reconciliation as a custom Oracle APEX extension to Oracle E-Business Suite, with ongoing enhancement and support.
- Executed the APEX platform upgrade from 5.2 to 18.2 across the environment.

Designed custom themes, templates and authentication frameworks deployed agency-wide at multiple federal clients, covering PIV, SAML, OAuth2 and OpenID.

COMMERCIAL SOLUTION ARCHITECTURE

Solution Architect - Pharmacovigilance Platform | Johnson & Johnson

08/2021 - 09/2022 | New Brunswick, NJ (Remote)

| *Enterprise solution architecture at one of the world's largest healthcare companies.*

- Redesigned the SMART pharmacovigilance application - a system acting as a magnifying glass over thousands of potential adverse events for drugs under surveillance, identifying statistical signals of possible relationships between adverse events and drugs.
- Optimized ingestion pipelines processing approximately 2 GB of new safety data nightly for a large multinational user base, standardizing diverse source formats for statistical signal detection.
- Built custom SAML authentication, custom PDF templates for regulatory reports, and custom Excel workbooks for downstream analytics.

Senior Solution Architect | JPMorgan Chase (retained via Oracle Corporation)

2010 - 2014 | Phoenix, AZ and Columbus, OH

| *Enterprise solution architecture at scale - 700+ applications, millions of daily users.*

- Technical lead for Chase's Mortgage Banking line of business, responsible for 78 APEX applications serving millions of customers; managed a team of 8 production support resources and oversaw all change management, code reviews and enhancements.
- Designed and coded 5 new applications for the bank over an 18-month period while managing the broader architecture team, and led the performance testing and coding standards teams.
- In an earlier engagement, optimized and tuned APEX infrastructure at JPMC - one of the largest APEX deployments globally, with 700+ applications on a single database instance serving millions of daily users. Designed a new SDLC and restructured the deployment process.

Solution Architect - Global E-Commerce Platform | Estee Lauder, Online Division

11/2019 - 03/2020 | New York, NY (Remote)

- Complete rewrite of the entire online platform for all 30+ Estee Lauder brands across 80+ countries, driving 200+ e-commerce sites on Oracle microservices.
- Architected and coded microservices supporting the web storefronts, and built APEX applications to generate, maintain and manage web service migrations.
- Dockerized Tomcat and ORDS for deployment on Kubernetes and Docker Swarm.

Solution Architect - Budget and Performance Systems | Los Angeles Metro, Office of Management and Budget

03/2015 - 04/2017 | Los Angeles, CA

- Built annual budget forecasting, quarterly budget amendment and budget analysis applications as custom Oracle APEX extensions to Oracle E-Business Suite.
- Delivered annual and quarterly performance-metric collection and reporting, including an executive dashboard for the CEO.

Senior Architect (Part-Time) | The Heritage Foundation

01/2023 - 05/2023 | Washington, DC

- Gathered requirements and designed and delivered four applications for a professional networking and resource-sharing platform as part of the initial team.

CIO and Co-Founder | Getamover.com

2002 - 2007 | Washington, DC

- Grew a venture-backed startup from zero to 200 employees as CIO and built the entire technical platform from scratch; the company was acquired in 2007.

EARLIER NOTABLE ENGAGEMENTS

- **Navy Marine Corps Intranet (NMCI), 2001** - Led a team of 9 building the Security Information Warehouse and Security Data Management applications, processing 10 TB of data monthly from 14 different security products across thousands of instances. Project manager and technical architect on a 30 TB warehouse; used Oracle Express, Oracle Warehouse Builder, Oracle Discoverer, Oracle Portal and HTML-DB, the precursor to APEX.
- **Lockheed Martin / U.S. Air Force, 2007-2008** - Developed the replacement for isqlplus, a legacy COTS tool, using APEX. Fulfilled 200+ requirements and provided simultaneous web access to 30+ Oracle databases.
- **U.S. Department of Justice, U.S. Sentencing Commission, 2010-2012** - Led a team of 6 rebuilding the case processing system handling 60,000+ cases and millions of case files, migrating from FoxPro to Oracle with an APEX front end.
- **Food and Drug Administration / Booz Allen, 2012** - Rebuilt the Tobacco Inspection Management System (TIMS) for the Center for Tobacco Products, processing thousands of inspections, using APEX 4.0 with extensive

JavaScript and jQuery.

- **North Carolina Department of Public Instruction, 2008-2010** - Technical lead; built the Direct Certification and Verification system certifying students statewide for free and reduced lunch programs against Department of Health and Human Services data, integrating 18 source systems. Owned database and application architecture, ETL and front-end delivery.
- **Fairfax County Public Schools, 2009-2010** - Maintenance and enhancement of a 500+ table Oracle data warehouse reporting on student disciplinary and safety data.
- **Boart Longyear, 2009** - Built a project management portal managing 4,000+ oil and gas drilling projects worldwide, interfacing with Oracle E-Business Suite; established development, test and production environments.
- **Additional engagements, 1998-2008** - U.S. Department of the Treasury, United States Postal Service, Massachusetts Department of Education, Baltimore Gas and Electric, Nextel, AT&T / Lucent, State of Georgia, United States Naval Academy and The World Bank.

TECHNICAL SKILLS

AI/ML and LLMs: Large Language Model (LLM) application architecture, RAG pipelines, retrieval and context engineering, embeddings and vector search, LLM fine-tuning (LoRA/PEFT), model evaluation and benchmarking, synthetic training data generation, multi-agent orchestration, tool calling, Model Context Protocol (MCP), provider abstraction and A/B testing, structured and grammar-constrained output, human-in-the-loop workflows, OpenAI/GPT-4, Claude, Gemini, Gemma, Qwen3-Coder, EmbeddingGemma, HuggingFace Transformers, PEFT, Ollama, Apple MLX, Oracle 23ai AI Vector Search, HNSW indexing, knowledge graphs, OCR and document intelligence, predictive analytics.

Solution Architecture: Enterprise architecture, federal systems integration, ERP modernization, legacy migration, cloud migration, API design, event-driven architecture, SDLC design, technical documentation, stakeholder management, requirements traceability, UAT management, production handover.

Cloud and Infrastructure: Oracle Cloud Infrastructure, Microsoft Azure, AWS, Docker, Docker Compose, Kubernetes, Docker Swarm, Terraform, Ansible, Jenkins, CI/CD, Git, GitLab, air-gapped deployment packaging, reproducible offline installs, observability and audit logging.

Development: Python (asyncio), JavaScript, TypeScript, Next.js, React, PL/SQL, SQL, Shell scripting, REST APIs via ORDS, SOAP web services, JSON, XML, CSS3, Tailwind, Chart.js, AG Grid Enterprise, jQuery, Jinja2.

Data: Oracle Database 23ai / 19c / 18c / 12c / 11gR2, Oracle E-Business Suite R12.2 (Financials, SCM, Projects, HCM, CRM, GL, AR, AP, PO, PA, FND, AOL), QAD ERP, Progress ODBC, data warehousing, star and snowflake schema design, Oracle OLAP, Oracle Data Integrator, OBIEE, ETL and ELT frameworks, master data management, BI Publisher, ORDS SODA collections.

Security and Compliance: Public Trust (active), TS/SCI eligible, air-gap architecture, zero-egress and valved-egress design, federal AI governance and model-approval compliance, SAML2, OAuth2, OpenID Connect, PIV/CAC authentication, Entra ID SSO, role-based access control, SOC 2-oriented and FedRAMP-adjacent patterns, immutable audit trails.

Product: SaaS architecture, multi-tenancy, per-seat licensing, one-command deployment, greenfield product development, pricing and packaging, startup leadership (CIO through acquisition).

CERTIFICATIONS, AWARDS AND PROFESSIONAL DEVELOPMENT

- Advanced AI for Enterprises - Massachusetts Institute of Technology, 2024
- Certified Blockchain Developer - Massachusetts Institute of Technology, 2018
- Certified Tableau Administrator - Tableau, 2017
- Certified Oracle Application Express (APEX) Developer - Oracle Corporation, 2015
- Oracle Top 5 Consultant (award) - Oracle Corporation, 2000
- Oracle Data Warehouse Wizard (award) - Oracle Corporation, 2000

- Certified Cognos Developer - Cognos/IBM, 1997
- Oracle Bootcamp - Oracle Corporation, 1995

CLEARANCE AND ELIGIBILITY

U.S. Citizen. Active Public Trust clearance. Eligible for TS/SCI. Delivery experience inside Department of Energy and NNSA national laboratory environments, the Naval Nuclear Propulsion Program, FAA, NOAA, NASA JPL, the Department of Justice and the Department of the Treasury.