

ALLEN CASS

Irvine, CA

LinkedIn: crossmedia411 • Web: allencass.com • Youtube: @AllenCass-Tech

SUMMARY

Experienced software engineer specializing in AI Agent Automation, Mixed Reality, and Computer Vision. Proven expertise in architecting intelligent systems combining Large Language Models (GPT, Claude, Gemini) with agentic frameworks (N8N, Google ADK, CrewAI, LangChain, Vertex AI). Extensive experience deploying XR and AI solutions, including HoloLens 2, Meta Quest, and ARKit/ARCore, and leveraging IoT platforms, including NVIDIA Jetson and microcontrollers, to deliver immersive applications.

EXPERIENCE

AI & XR Systems Developer – ThreadX (Contract Project)

Feb 2025 – Present | Remote

Architected and developed **ThreadX**, an AI-powered Learning Management and Training Platform integrating multimodal experiences across web and XR devices. The system combines an adaptive AI training assistant, dynamic RAG-based knowledge retrieval, and a synchronized XR training continuation system for immersive skill development.

- **Web Platform:** Built the core **Flutter-based LMS** allowing users to register, take interactive courses, track progress, and engage with content through a **Google Cloud-integrated AI chat interface (Gemini)**. Designed a **Firebase backend** (Firestore, Auth, Cloud Functions, Storage, Analytics) for real-time data handling and secure authentication.
- **AI & Agent System:** Implemented a **backend agentic RAG pipeline** enabling contextual question answering and personalized learning recommendations. Leveraged **vector-based retrieval with Qdrant** to power intelligent content interaction.
- **XR Platform:** Developed the **Unity-based Meta Quest application** extending LMS courses into immersive, step-by-step training simulations. Leveraged **XR Interaction Toolkit** to implement interactive course progression, object manipulation, and user feedback. Integrated **Firebase Auth and Firestore** to seamlessly continue user progress from web to XR.
- **AI Chat in XR:** Embedded an in-headset AI assistant allowing trainees to query training material in natural language, enhancing procedural understanding and engagement.
- Orchestrated **cross-platform data flow** between web and XR environments, ensuring training session persistence and real-time content updates.

Innovation Lead - AR/XR & AI Agent Developer III - Avatar Partners

Aug 2020 – Jan 2025 | Irvine, CA

- Pioneered and deployed an **AI Agent-driven system using N8N and PostgreSQL** to automate and enhance the V-22 aircraft maintenance workflow. This system intelligently triaged issues detected by a HoloLens 2 XR application (wiring, chafing, thermal damage), dynamically queried technical manuals via **RAG**, generated standardized defect reports, and updated maintenance logs in real-time, contributing to a 25% reduction in documentation overhead.
- Developed the core Phase I XR maintenance application for V-22 aircraft (HoloLens 2, MRTK) with computer vision (Azure Custom Vision) for real-time fault detection. Containerized the solution on NVIDIA Jetson Orin (Docker, REST APIs) and built the Unity3D stack, initially cutting procedure time by 20%.
- Led development of an **AI-augmented XR maintenance solution for F-18 aircraft**, integrating **N8N and LLM-powered agents** to automate data analysis and intelligent assistance. Agents processed xAPI data from an LRS, identified common error patterns in corrosion detection tasks, and proactively alerted supervisors to trainees needing support, resulting in a 15% increase in task accuracy and a 20% reduction in training time.
- Built the underlying AI-powered corrosion detection system (YOLOv8) and a Docker-based metrics tracking system (NVIDIA Jetson Orin) for the F-18 project, guiding maintainers through step-by-step procedures.
- Spearheaded the integration of a **Cognitive Assistant Agent leveraging Claude API and custom knowledge bases** into ARSurgeon, an XR medical training app for knee replacement (HoloLens 2, MRTK). The agent provided real-time, context-aware procedural guidance and dynamically answered trainee questions, enhancing knowledge retention and contributing to a 10-minute reduction in procedure training time per session.
- Developed a precision livestock monitoring system for the USDA (iPad LiDAR, YOLOv11 pose estimation), implementing real-time distance calculation. Built a Unity3D application integrating OpenCV, utilizing CVAT and synthetic data (Unity Perception) for efficient data labeling. Implemented an AI agent to automate the batch processing and preliminary analysis of collected pose data, streamlining reporting to researchers.

Mobile Cloud AR/XR/AI Software Engineer - Independent Contractor

Dec 2018 – Aug 2020 | Irvine, CA

- Developed an AI-powered digital advertising platform achieving 95% accuracy in real-time customer detection and demographic analysis. **Explored early-stage AI agent concepts for dynamic ad content tailoring based on detected demographics.**
- Built real-time audience analytics with MongoDB, handling 10,000+ interactions daily with sub-200ms response times, achieving 40% higher engagement and a 25% increase in conversion rates.
- Developed a cross-platform mobile app with Flutter for iOS and Android, enhancing engagement for lacrosse players and communities. Built key features like facility reservations, player tracking, event calendars, and real-time chat. Designed backend with Firebase.

Senior Software Engineer - Meridiun

June 2015 – Dec 2018 | Irvine, CA

- Developed a dual-purpose AR platform for Edwards Lifesciences (Unity3D, iOS/Android), enhancing heart valve training and marketing, reducing training time by 50% for 1,000+ medical professionals.
- Developed an AR onboarding app for Walmart's HR division, gamifying employee orientation on iOS. Successfully engaged 50,000+ employees.
- Developed a proof-of-concept AR app for the Ford Escape, enhancing the customer journey using spatial computing.

TECHNICAL EXPERTISE

Software Development and Workflow Tools

PyCharm • Visual Studio • Unity3D • Git/GitHub • Jira • Docker • Arduino

Programming Languages

C# • Python • Flutter/Dart

AI Frameworks & Platforms

- **Agentic AI & Orchestration:**
 - N8N • Google Agent Development Kit (ADK) • Vertex AI Studio • CrewAI • LangChain • LMStudio
- **Large Language Models / APIs:**
 - Anthropic: Claude • OpenAI: ChatGPT • Google AI Studio (Gemini) • Perplexity • Retrieval Augmented Generation (RAG)
- **Computer Vision:**
 - YOLO • Ultralytics • OpenCV • PyTorch • Azure Custom Vision • MediaPipe • Roboflow • Edge Impulse

Mixed Reality

- **Hardware:**
 - HoloLens 2 • Quest 3/2 • Apple Vision Pro • iOS • Android
- **Software:**
 - Unity3D • XR Interaction Toolkit • MRTK • ARKit/ARCore • ARFoundation • Vuforia

Internet of Things (IoT)

- **Hardware:**
 - NVIDIA Jetson Orin/Nano • Raspberry Pi • ESP32/ESP8266
- **Software:**
 - Arduino • Home Assistant • Node-RED • MQTT • ESPHome

EDUCATION & CERTIFICATIONS

- **Google / Kaggle:** Generative AI Certification
- **Microsoft Certified:** Azure AI Fundamentals
- **NVIDIA:** Getting Started With Deepstream For Video Analytics on Jetson Series
- **LinkedIn Learning:** API Gateway with HTTP, Lambda, DynamoDB, and iOS
- **LinkedIn Learning:** Machine Learning by Example
- **LinkedIn Learning:** Building Serverless Apps on AWS