

José Manuel Cortés Cerón

Software Engineer | AI Full-Stack Development
deepdevjose@itsoeh.edu.mx | +52 55 6011 5704 | Hidalgo, Mexico

github.com/deepdevjose | linkedin.com/in/deepdevjose | deepdev-portfolio.vercel.app

Professional Summary

Software engineer and AI researcher (Research Collaborator at XJTU). Builds intelligent systems at the intersection of AI, robotics telemetry, browser-native inference, and high-performance interactive computing. Proven ability to deliver quantifiable improvements: latency reduction, memory efficiency, and predictive accuracy. Published work on hybrid AI architectures for digital heritage.

Technical Skills

Languages: Python, TypeScript, JavaScript, C++, Java, SQL

Frontend: React, Next.js, Three.js/WebGL, TailwindCSS

AI/ML: PyTorch, ONNX Runtime Web, Transformers.js, LSTM, Diffusion (LoRA), XTTS

Backend/Cloud: Node.js, Docker, Firebase, Vercel, Kubernetes (basics)

IoT & Robotics: MQTT, ESP32, Node-RED, Sensor Fusion, Digital Twins

Experience

AI Research Collaborator – Xi'an Jiaotong-Liverpool University (XJTU), China

2025 – Present

Intelligent Systems Lab – Robotics Telemetry & Predictive Maintenance

- Architected **Arm Health AI**, an LSTM-based predictive maintenance platform achieving **85% Remaining Useful Life (RUL) accuracy**, reducing simulated downtime by **40%**.
- Built a real-time digital twin interface with Three.js, processing **10,000+ telemetry events/second** via MQTT from simulated robotic arms.
- Developed a WebGPU-accelerated vision-language runtime with **50ms inference latency** (3× faster than CPU) and **60% memory reduction** via kernel fusion and quantisation.
- Deployed scalable backend on Vercel/Firebase, supporting **100+ concurrent simulated robotic arms**.

Technology Workshop Instructor & Speaker

2023 – Present

ITSOEH & Conferences – AI, Software Engineering, Cybersecurity

- Delivered 15+ workshops to 200+ students, improving post-workshop assessment scores by **60%**.
- Keynote speaker at Expo Profesiográfica and CONAINTE (2024), reaching 500+ attendees.

Selected Engineering Contributions

• Arm Health AI

github.com/deepdevjose/Arm-SaaS

- LSTM-based RUL using multivariate telemetry (joint angles, torque, vibration) → **85% accuracy**.
- Real-time digital twin with Three.js → **10k+ events/sec** via MQTT.
- Deployed on Vercel/Firebase → supports **100+ concurrent simulated arms**.

Stack: Python, PyTorch, Next.js, Three.js, MQTT, Firebase

• Vision-Language Runtime

github.com/deepdevjose/Vision-Language-Runtime

- WebGPU acceleration for Moondream/LLaVA → **50ms latency** (3× faster than CPU).
- Kernel fusion and quantisation → **60% lower memory footprint**, no cloud dependency.
- Privacy-preserving inference runs entirely on consumer GPUs.

Stack: JavaScript, WebGPU, ONNX Runtime Web, Transformers.js

• Byron Dynamic Museum (Peer-reviewed article)

github.com/deepdevjose/Byron-s-virtual-museum

- Restored visual style with LoRA-fine-tuned Stable Diffusion XL.
- Reconstructed voice from **<10 minutes** of cleaned audio (XTTS v2).
- WebGL renderer with spatial lazy loading & occlusion culling → stable **60 FPS** on mid-range devices.

Stack: Three.js, WebGL, Python (Diffusers), XTTS, SDXL

Publication

Cortés Cerón, J.M. (2026). *Decoupled hybrid architecture for generative AI-assisted interactive cultural preservation systems*. Conference article (under review). [GitHub]

Education & Certifications

B.S. in Information and Communication Technologies – ITSOEH, Mexico (2023–Present)

Programming Technician – CETis 26, Mexico (2023)

Certifications: IoT (Cisco), Data Science Essentials (Cisco), Linux Server Administration

Languages & Awards

Languages: Spanish (native), English (professional proficiency), Chinese (elementary)

Award: Winner, Intercollegiate Programming Rally (ITSOEH, 2024)