

JESTAZ YAO

AI Product Designer · Systems & Decisions

"I design the decisions AI products make — not just the screens."

www.jestazyao.com · [LinkedIn](#) · yxhazdm@gmail.com · 347-622-6469 · New York, NY

AI Product Designer with 5+ years shipping end-to-end products across web, e-commerce, and LLM systems. I design **deterministic AI workflows** — multi-agent orchestration, intent taxonomies, policy engines, confidence gating, structured outputs — and translate them into measurable business outcomes. Most recent independent work: **Director**, a multi-agent supply-chain decision system where an LLM Director routes between deterministic sub-agents, benchmarked against keyword and multi-step routers across a 30-question eval harness.

SELECTED PROJECTS

Director — Supply Chain Decision System

April 2026

Independent · www.jestazyao.com/supplyai

- Designed a **multi-agent decision architecture** with four deterministic Python sub-agents (late-delivery risk, loophole detector, network visibility, KPI reporter) coordinated by an LLM Director that selects which agents to run per query.
- Built and benchmarked **three router strategies** (regex keyword, LLM tool-call, multi-step plan → reflect) against a 30-question eval harness — keyword F1 0.888, LLM 0.792, multi-step 0.831 — surfacing trade-offs between recall, precision, and cost.
- Shipped **citation drilling** UX where every finding can expand to show the underlying data rows on demand, eliminating black-box answers and giving users an audit trail.
- Implemented a **plan → run → reflect** agentic loop for multi-step queries: the model proposes a plan, executes sub-agents, then reflects on whether the answer is complete before responding.
- Built the entire product end-to-end: FastAPI backend, **streaming SSE responses**, live routing-graph visualization, drag-resizable activity panel, CSV upload with fuzzy column mapping, and Markdown / PDF export.
- **Result:** three benchmarked router strategies · F1 0.888 / Precision 0.792 / Recall 0.831 across 30 hand-labeled questions · ~2s p50 latency · live demo, case study, and LinkedIn carousel shipped.

The Translator Pattern — AI Support Automation System

March 2026

Independent · www.jestazyao.com/aipattern

- Designed a **4-gate decision pipeline** (intent → policy → confidence → clarification) that separates LLM reasoning from business rules and forces deterministic execution before any tool call fires.
- Built a structured **intent taxonomy** covering refund, recommendation, and upsell scenarios; reduced ambiguous routing in adversarial test runs to 0.
- Implemented a **rule-based policy engine** and confidence-threshold gating to block low-certainty actions and route them to fallback or escalation.
- Designed multi-turn clarification flows that systematically collect missing parameters before triggering decision logic.
- Defined **JSON-schema output contracts** for clean backend integration and predictable downstream handoff.
- **Result:** 100% unauthorized-refund block rate · 6/6 adversarial tests passed · 0 hallucinated “refund processed” outputs · ~900ms firing pulse.

Demon Rising — Conversational AI System in a Shipped Game

May 2026

Independent · www.jestazyao.com/demon-rising · orhpius.itch.io/demon-rising

- Designed a free-text negotiation system where players answer an in-game council in natural language; built a **5-way intent taxonomy** (agree / refuse / modify / counter / silence) that maps any open-ended input onto a bounded, actionable decision space.
- Architected a **fallback-first pipeline** — an LLM intent classifier backed by a deterministic keyword classifier — so every interaction resolves correctly even when the model is slow, offline, or returns malformed output.
- Implemented **bounded-latency control** (≤8s budget) with async polling and a keyword pre-pass that skips the model entirely on unambiguous input, eliminating model-blocked UI states.
- Designed **legible AI feedback**: an in-product “herald” surfaces the system’s intent classification back to the user in plain language, turning a hidden ML decision into a readable, trustable interaction.
- Built the full product end-to-end in Godot / GDScript — combat, systems, UI — with the LLM running locally on the player’s machine.
- **Result:** 0 dead-end states across the intent model · 100% of AI interactions resolve under model failure · ≤8s worst-case latency · playable build shipped on itch.io.

PROFESSIONAL EXPERIENCE

Lead Product Designer (Contract)

Sep 2024 – Mar 2026

Vetta Studio · New York, NY

- Led UX strategy and redesign across multiple client platforms in Figma — wireframes, high-fidelity UI, interactive prototypes, and reusable design-system components — pairing AI-assisted flows with conversion-driven product architecture.
- **Mechapro (high-ticket B2B e-commerce)**: designed product architecture, decision flows, and configurator UX — drove **\$38K+ revenue in 30 days at \$9.7K average order value**.
- **Beauty Affair**: shipped AI-assisted booking experience generating **90+ bookings/month and ~\$5K projected monthly revenue**.
- **Halo Wellness (DTC skincare)**: designed an AI product-recommendation quiz and onboarding flow in Figma; lifted email opt-in conversion in A/B testing.
- **Atlas Concierge (B2B travel SaaS)**: wireframed and prototyped a conversational booking dashboard with WCAG 2.1 AA-compliant components for keyboard and screen-reader users.
- Designed AI chatbot systems for product discovery, booking, and purchase flows — paired with paid media, landing pages, and lifecycle creative to compound funnel performance.

Lead UI/UX Designer · Mobile App Instructor

Aug 2025 – Jan 2026

Sage Test Prep · New York, NY

- Led redesign and launch of two production websites with AI-driven discovery and Q&A flows; mentored designers on UX quality and process.
- Built structured AI interaction patterns — intent capture, constraint handling, fallback escalation — to keep conversational UX scalable and predictable.
- Integrated AI chat with landing pages and ad funnels, reducing friction between marketing capture and product engagement.
- Delivered WCAG 2.1 AA-compliant, multi-language, cross-platform (iOS/Android) experiences shipped to production — including AI chat flows with screen-reader support, keyboard navigation, and explicit focus states designed in Figma.

Lead UI/UX Designer (Web3)

Mar 2023 – Sep 2024

Orbimatrix · New York, NY

- Led UX for a Web3 trivia platform with on-chain rewards and wallet connectivity; built reusable patterns for token incentives and player retention.
- Migrated team from Sketch to Figma mid-project; ran usability testing rounds and iterated on hierarchy, color, and flow — boosted engagement **15%**.
- Partnered with engineering to translate blockchain mechanics into visible, comprehensible UI states.

UI/UX Designer · Mobile & Design Systems

Oct 2021 – Mar 2023

Appetizer Mobile · New York, NY

- Led 0 → 1 product design across iOS/Android — admin panels, AI surfaces, blockchain apps — establishing scalable design systems and UX foundations.
- Defined and scaled Figma component libraries with auto-layout, variants, and documentation standards across complex, media-rich experiences — built to WCAG 2.1 AA contrast and keyboard-navigation specs.
- Bountify (token-based flows): aligned wallet connectivity, rewards, and real-time system states with mobile UX to lift completion and retention.
- Drove team adoption of AI-assisted design tooling to accelerate execution while holding craft.

E-commerce Designer

Jun 2021 – Oct 2021

ReikoWireless · New York, NY

- Wireframed and prototyped product pages and checkout flow in Figma to reduce abandonment and lift conversion; collaborated with marketing/dev on responsive rollout across desktop and mobile breakpoints.

SKILLS

AI Systems & LLM Design

Multi-agent orchestration · LLM Director / router patterns · Eval-driven design (F1, precision, recall benchmarks) · Plan → reflect agentic loops · Streaming SSE responses · Citation / grounding UX · Intent taxonomy · Structured JSON outputs · Deterministic policy engines · Confidence threshold gating · Fallback / graceful-degradation design · Latency budgeting · Local LLM integration · Tool-calling orchestration · Multi-turn state management · Prompt design & LLM behavior control · RAG architecture (foundational) · Embedding retrieval (foundational)

Tools: Anthropic API (Claude) · OpenAI API · OpenAI Playground · ChatGPT · FastAPI · Python · JSON Schema · Thunder Client · REST APIs · Render deploy

Product & UX Strategy

Information architecture · User flow & interaction design · Design systems & scalable components · Wireframing & rapid prototyping · Usability testing (qual + quant) · Heuristic evaluation · Stakeholder alignment · Accessibility (WCAG 2.1 AA · keyboard navigation · screen-reader support · color contrast validation)

Design & Build Tools

Figma · Figma Make · Sketch · ProtoPie · Adobe Suite · InVision · HTML / CSS · Wix Studio · Bubble.io · WordPress · Squarespace · Xcode · VS Code · Jira · Miro · Axure

Data & Growth

Google Analytics · Google Search Console · Google Ads · SQL · Excel

Generative

Midjourney · Runway ML

EDUCATION

Queens University

Bachelor of Fine Arts (BFA) · Queens, NY

Jun 2016 – Jun 2020

CERTIFICATIONS

LinkedIn Learning: UX Foundations (Usability Testing · Interaction Design) · Data Analytics for Business Professionals · UX Deep Dive: Remote Research

FutureLearn: Introduction to Creative AI · Introduction to Conversational Interfaces