AI-Vastra — Research Notes & Design Decisions

Project Title: AI-Vastra — A Multimodal AI Assistant for E-Commerce Trust and Fit
Team:Prathamesh Baviskar · Dheeraj Dagar · Ayushmaan Singh · Rishith · Rajesh Inapakurti
Course: DES646 — AI for Designers
Duration: 15 Oct – 09 Nov 2025
Goal: To design a multimodal AI prototype that enhances trust, accessibility, and scale perception for new online shoppers through speech, language, and vision AI integration.

Research Decisions

1. Project Framing
Objective:

To improve trust, accessibility, and fit visualization for first-time e-commerce shoppers through a multimodal AI interface that merges speech, language, and vision tools. The design problem aligns with the DES646 course theme of AI as a design material—using existing APIs to create inclusive, human-centered digital systems rather than novel algorithms.

Key Question:

How might AI multimodality reduce perceived risk and cognitive friction for low-literacy or non-English-speaking users engaging with online retail?

2. Conceptual Research
2.1 Trust & Transparency in AI-Mediated Shopping

- Research indicates trust increases when systems display clear verification and explainability cues.

- Decision: Introduce a static "Verified" badge instead of a dynamic trust score to prevent misinterpretation.

2.2 Accessibility via Multimodal Interaction

- Speech interfaces bridge non-literate and multilingual users.

- Decision: Speech-to-speech mode as default; text fallback retained.

- Design implication: Mic button, pulse animation, bilingual captions.

2.3 Scale and Fit Visualization

- 60% of e-commerce returns arise from misjudged size/fit.

- Decision: Combine two tools — Virtual Try-On (VTO) for apparel, AR Size Visualizer for non-apparel.

- API-based integration preferred for low latency and modular design.

3. Technical & Design Decisions
3.1 Framework Choice — Vanilla JS over React/Streamlit

- Simplicity and direct control over DOM for fast iteration. React was considered but rejected due to time constraints and the need for tight integration with HTML/CSS-based Meesho layout

3.2 Voice Layer — Azure Cognitive Services

- Provides robust multilingual speech support (Hindi, English, regional Indian languages) and easy client-side JavaScript integration.
- Tested alternatives (Google Speech, Whisper API) required higher latency or server deployment

- Ethical note: speech data is not stored.

3.3 LLM Layer — Gemini / Perplexity API

- Gemini offers structured, controllable summarization — ideal for review-based question answering.
- The assistant summarizes real product reviews rather than generating open-ended text, reducing hallucination risks.

3.4 Vision Layer — Clip Drop + Stream lit

- Clip Drop for object scaling; Stream lit for garment try-on.

- REST API integration with <3s latency target.

3.5 Language UX

- JS dictionary for translations.

- Bilingual English–Hindi support with mirrored text/audio output.

4. Evaluation Strategy
Latency ≤ 2.5 s

Trust comprehension ≥ 80%

Accessibility task completion ≥ 85%

Language switching success 95%

Perceived helpfulness ≥ 4 (Likert mean)

User walk-throughs confirmed 90% unaided task success.

5. Ethical and Design Reflections

- Every AI output includes a disclaimer.

- Voice assistant announces non-official nature of summaries.

- High-contrast, accessible bilingual UI.

- Lightweight, energy-efficient API integrations (no local training).

6. Synthesis: Research through Design

AI-Vastra's novelty lies in multimodal integration as design practice.

By combining speech, language, and vision AI, it reframes e-commerce trust as a designable, perceptual experience rather than a purely technical challenge.

It demonstrates AI as a medium for accessibility, confidence, and inclusivity in digital shopping.

AI-Vastra — Working Log & Iterations

Phase 1 — Foundation & Architecture (15–22 Oct 2025)

Objective: Establish baseline architecture and user flow.

- Defined three friction points: trust, accessibility, and size perception.

- Chose Vanilla JS/HTML/CSS for simplicity and direct control; avoided React for speed.

- Created modular system: web UI, Azure Speech, Gemini API, ClipDrop + Streamlit VTO.

- Delivered static homepage, product grid, and modal assistant design with responsive layout.

Phase 2 — Visualization Modules (23–28 Oct 2025)
Objective: Implement visual trust features (seeing is believing).

- Developed Streamlit-based Virtual Try-On (VTO) for apparel.

- Integrated ClipDrop API for AR Size Visualizer on non-apparel pages.

- Optimized API parameters to maintain clarity and low latency.

- Delivered functional visualization suite linked from all product pages.

Phase 3 — Multimodal Assistant Integration (29 Oct–2 Nov 2025)
Objective: Enable accessible, conversational interaction.

- Integrated Azure STT/TTS for English–Hindi voice support.

- Latency averaged 2.4s per query; added mic push-to-talk and listening animation.

- Connected Gemini API for LLM-based review summarization.

- Implemented bilingual dropdown and prompt localization.

- Delivered functional voice-text assistant with context-aware Q&A.

Phase 4 — Trust, Feedback & Usability (3–7 Nov 2025)
Objective: Refine UI and validate trust mechanisms.

- Added static "Verified" badge on product cards.

- Improved assistant modal with progress messages and animated mic feedback.

- Conducted usability testing (n=6, bilingual group): 90% task success rate.

- Added captioning for speech output and refined AR feedback flow.

Phase 5 — Final Integration & Documentation (8–9 Nov 2025)
Objective: Polish and consolidate prototype for submission.

- Code cleanup and responsive validation.

- Added ethical statement and disclaimer on AI outputs.

- Created README, screenshots, and demonstration flow.

- Finalized cohesive prototype integrating all speech, vision, and language modules.

Reflection

AI-Vastra demonstrates system integration as design research.

Its innovation lies in orchestrating existing AI modalities to enhance trust and inclusivity rather than algorithmic novelty.

The prototype reveals how AI, when framed as a design material, can enable confidence, comprehension, and participation in digital commerce for new user groups.

AI-Vastra System Workflow

User Starts

↓

User Interface
Lang Select

Voice Input

↓

Speech-To-Text
Azure Cognitive

↓

Query Process
LLM API

↓

Product Type?

Apparel
Virtual Try On

Non-Apparel
ClipDrop

Compile Response

↓

Text-to-Speech

↓

Output to User

Output to User