

Generic User Framework: Topic Notes

Core Concept

A dual-AI architecture enabling natural interaction with web interfaces, mimicking how humans navigate browsers while maintaining a clear separation of concerns.

Key Components

Browser Interaction Specialist

- Understands UI patterns and interaction models
- Translates high-level intents into specific browser interactions
- Develops expertise in navigation patterns across interfaces
- Serves as intermediary between domain AI and browser

Domain Specialist

- Focuses on specific domain knowledge (healthcare, finance, etc.)
- Makes decisions about what needs to be accomplished
- Communicates with Browser Interaction Specialist via high-level intents
- Remains free from needing to understand browser interaction details

Training Methodology

- Interaction networks capturing paths through interfaces
- Logged data structures recording user actions, timing, and outcomes
- Hybrid approach combining aggregated patterns with raw interactions
- Learning models: sequence prediction, task decomposition, element recognition

Key Innovations

- Leveraging browser's own understanding of the page
- Separation of concerns mirroring human cognition
- Robustness to website changes through operating on rendered elements
- Scalable learning across different interface types

Current State

- Conceptual framework fully developed
- Training methodology defined
- Multiple papers created with proper citations
- Diagrams for various components

Open Questions

- Implementation approach for proof-of-concept
- Training data collection methodologies
- Evaluation metrics and validation approaches
- Integration with other AI systems

Related Concepts

- Memory Agent (for persistent context across sessions)
- Multi-agent communication protocols
- Software integrated circuit philosophy

References

- Academic papers on browser automation challenges
- Research on DOM-based interaction limitations
- Studies on AI approaches to visual interface understanding

Artifacts Created

- "The Generic User Approach: Reimagining AI-Browser Interaction" paper
- "Training AI to Navigate Interfaces as Humans Do" paper
- Various SVG diagrams illustrating components and processes