

# Multimodal Chatbots

Nancy Xu

Google's UI hasn't changed much in 20 years.

**Information interfaces are ripe for disruption.**

# The Internet looks like this ...

**Map of American New Restaurants:** Shows a map of Jackson, WY with several restaurant locations marked with red pins and numbered 1 through 10. A search bar at the top says "Redo Search In This Area".

**Milly Clothing Store:** Displays two dresses: a "LEOPARD SEQUINS ANNIE DRESS" and a "FLORAL CROCHET ADELINE DRESS".

**Tesla Inc. Market Summary:** Shows the stock price at 449.39 USD, up 7.24 (1.64%). It includes a line chart for the day's performance and a table of key statistics.

Open	453.13	Div yield	-
High	455.68	Prev close	442.15
Low	407.07	52-wk high	502.49
Mkt cap	418.75B	52-wk low	43.67
P/E ratio	1,156.25		

**Stay Wild Instagram Profile:** A profile for "staywildjacksonhole" with 965 posts, 30.2K followers, and 572 following. The bio describes it as a Jackson Hole, Wyoming community.

# Chatbots look like this ...

**Facebook Chat with Yelp:** Shows a conversation where a user asks for a 5-star restaurant in Jackson Hole. The chatbot responds with a helpful message and provides contact information for support.

**Stay Wild Text Message:** A text message from "staywildjacksonhole" asking "Can you help me?".

**Animal Facts Chatbot:** A chatbot interface where a user asks "Which animal do you like?". The chatbot provides a list of animals and their interesting facts, such as "I heard they go to college" for cows and "Horses go to Hayraed" for horses.

Virtual assistants are opening doors for **natural, dynamic** information-seeking experiences.



Maps



Graphs



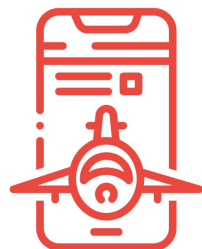
Products



Status

Our goal for integrating GUIs:

1. Improved **accuracy** + agent **control** of conversational flow.  
Completion rate
2. Enhanced **user satisfaction** with GUI experiences.
3. **Faster** time to task completion.



Actions



Lists



Filters



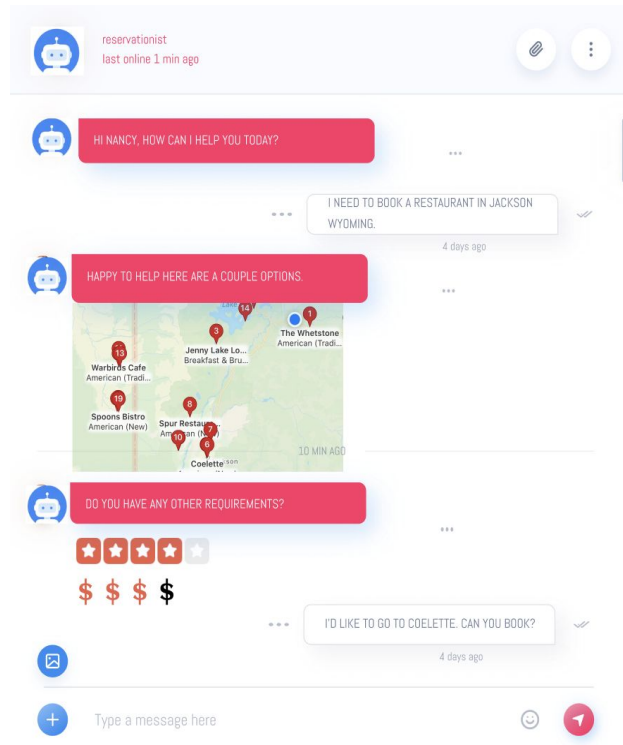
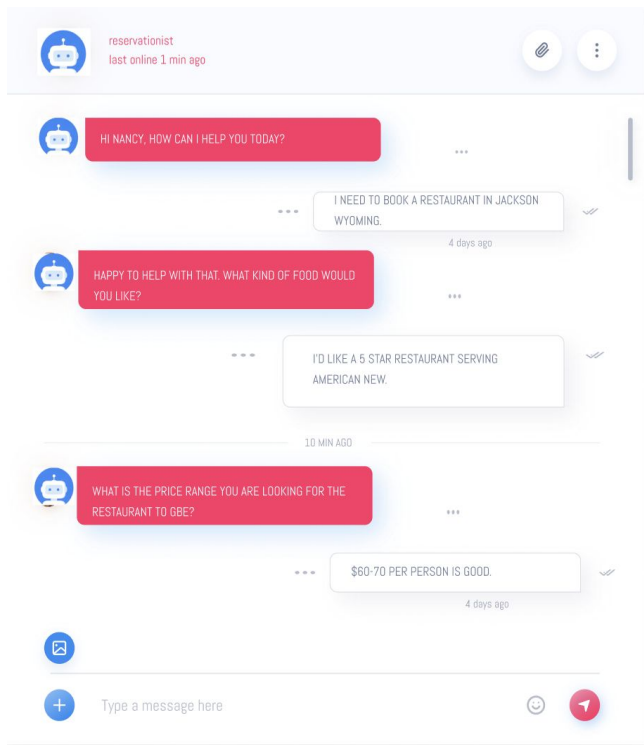
Images

For this project, we'll build a multimodal chatbot that utilizes GUIs in dialogue.

### Project Timeline:

- **Week 2-3:** Train + deploy a standalone text-based restaurant reservation chatbot running on Yelp API.
- **Week 4-5:** Build dialogue state machine (AI) to support GUI transitions.
- **Simultaneously Week 4-6:** Develop + integrate GUIs into the reservationist chatbot
- **Week 7-9:** Run user studies on time to completion, user satisfaction, completion rate / accuracy.

# Reimagine chatbots via a **Virtual Reservationist**

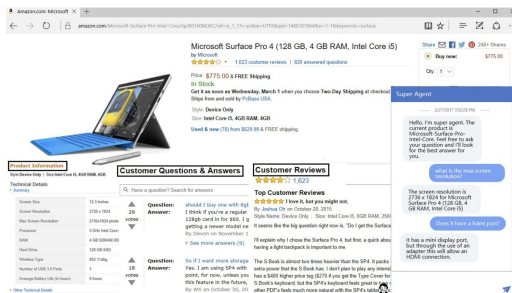


@com.yelp

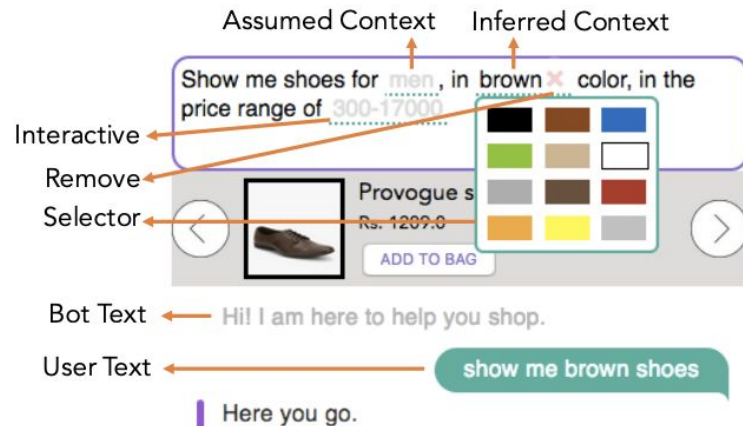
- Ratings
- Cuisine
- Price
- Name
- Geo
- Phone
- Site
- Review count

# yes, papers can be written too!

## SuperAgent: A Customer Service Chatbot for E-commerce Websites



## Convey: Exploring the Use of a Context View for Chatbots



## Brassau: Automatic Generation of Graphical User Interfaces for Virtual Assistants

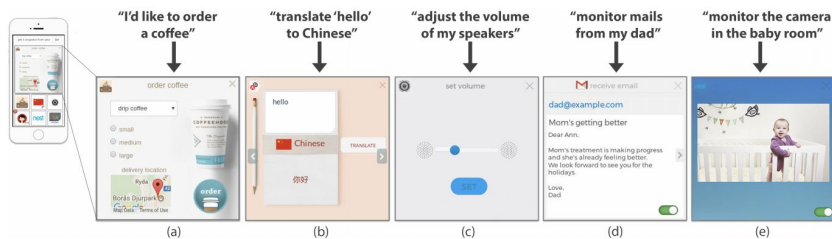


Figure 1: Brassau automatically generates interactive graphical user interfaces for virtual assistant commands in natural language.