

IN4MATX 133: User Interface Software

Lecture 18:
Augmenting Reality &
Conversation

Today's goals

By the end of today, you should be able to...

- Describe some of the history of AR and VR and conversational interfaces
- Differentiate augmented and virtual reality
- Explain key principles for designing a good AR experience
- Articulate some principles for designing conversational experiences

What is augmented reality?
What is virtual reality?
How do they differ?

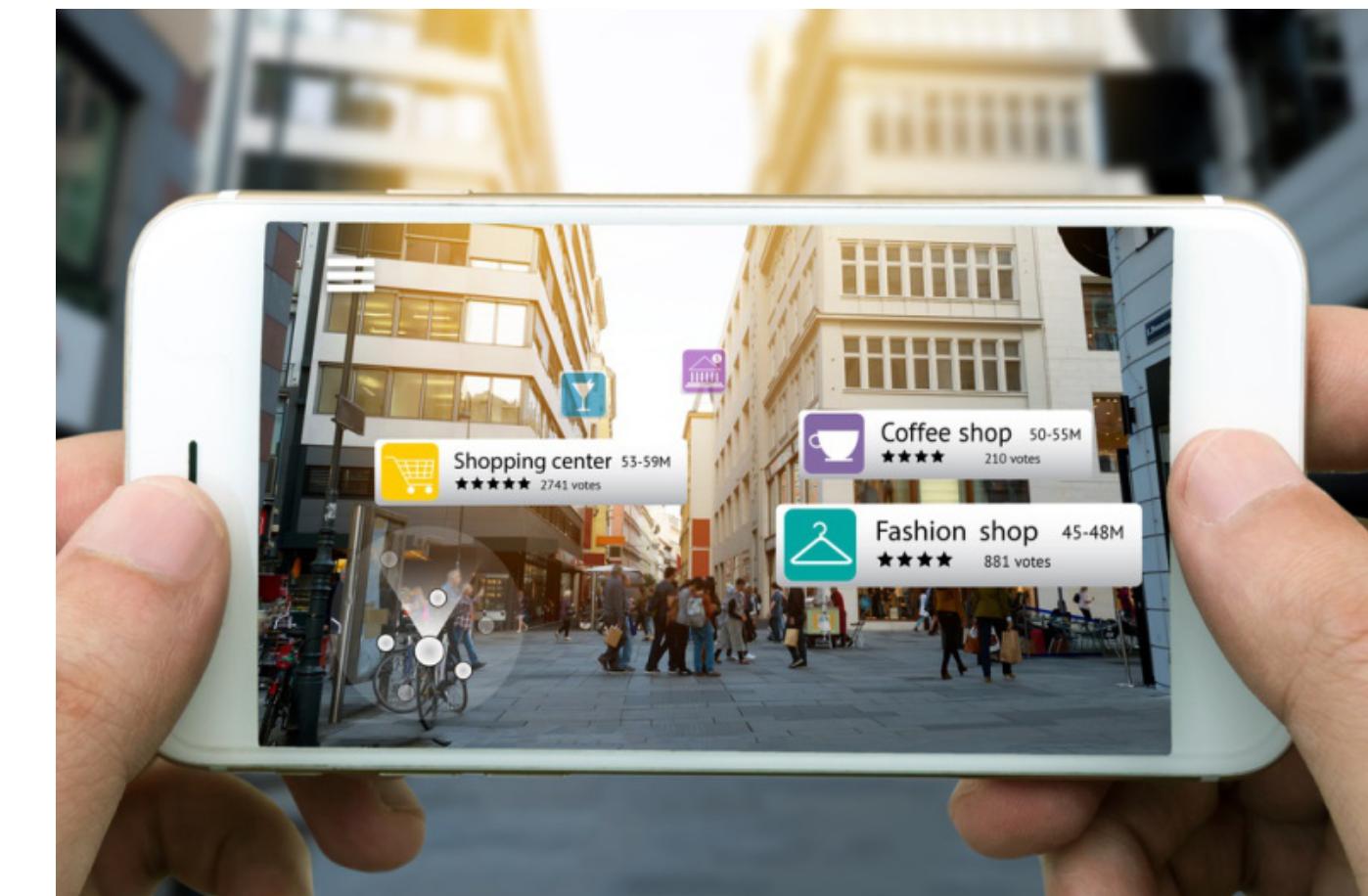
Augmented reality vs. Virtual reality

- Virtual reality aims to transport a person to an entirely new place by blocking out the real world as much as possible
 - Ideally a full sensory experience



Augmented reality vs. Virtual reality

- Augmented reality layers digital enhancements on top of a view of the real world
 - Digital content is interactive
 - Digital content is viewable in 3D (or projected 3D)



Mixed reality (hybrid reality)

- Physical objects in the real world can be interacted with in the digital world
 - Kind of a hybrid of AR and VR
- By comparison, VR seeks to immerse a person in a completely artificial environment
- While AR just overlays objects in the real world

Virtual, Augmented, and Mixed reality

- Virtual reality
 - HTC Vive, Oculus rift
- Augmented reality
 - Apple ARKit, SnapChat lenses, Pokemon Go
- Mixed reality
 - Microsoft HoloLens

CAVE (1992)



<https://www.youtube.com/watch?v=-Sf6bJjwSCE>

Carolina Cruz-Neira, Daniel J. Sandin, Thomas A. DeFanti, Robert V. Kenyon, and John C. Hart.
The CAVE: audio visual experience automatic virtual environment. *Communications of the ACM* 35, 6 (June 1992)

CAVE (1992)

- Viewing headset
(though it used projection walls)
- Sensors to detect a person's position and orientation in the space
- Audio feedback
- Motion controller



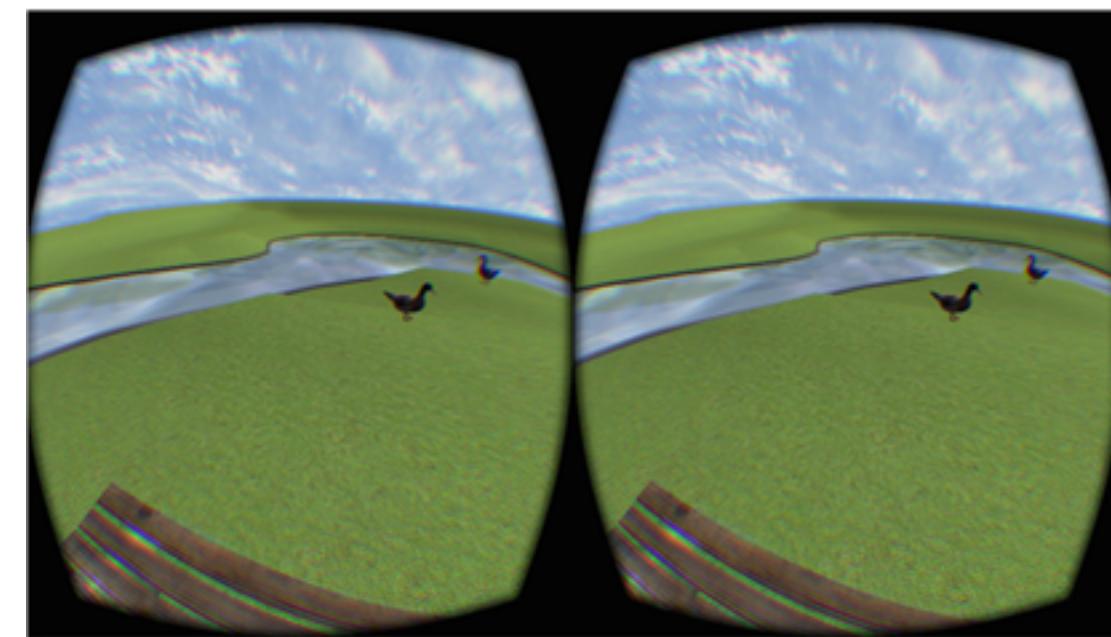
CAVE (1992)

- Necessitated a specialized setup
 - Large room
 - Four projection walls
 - Expensive wearable displays
- Technology could never be adopted by consumers
 - Who has that kind of space to devote to VR, never mind money?



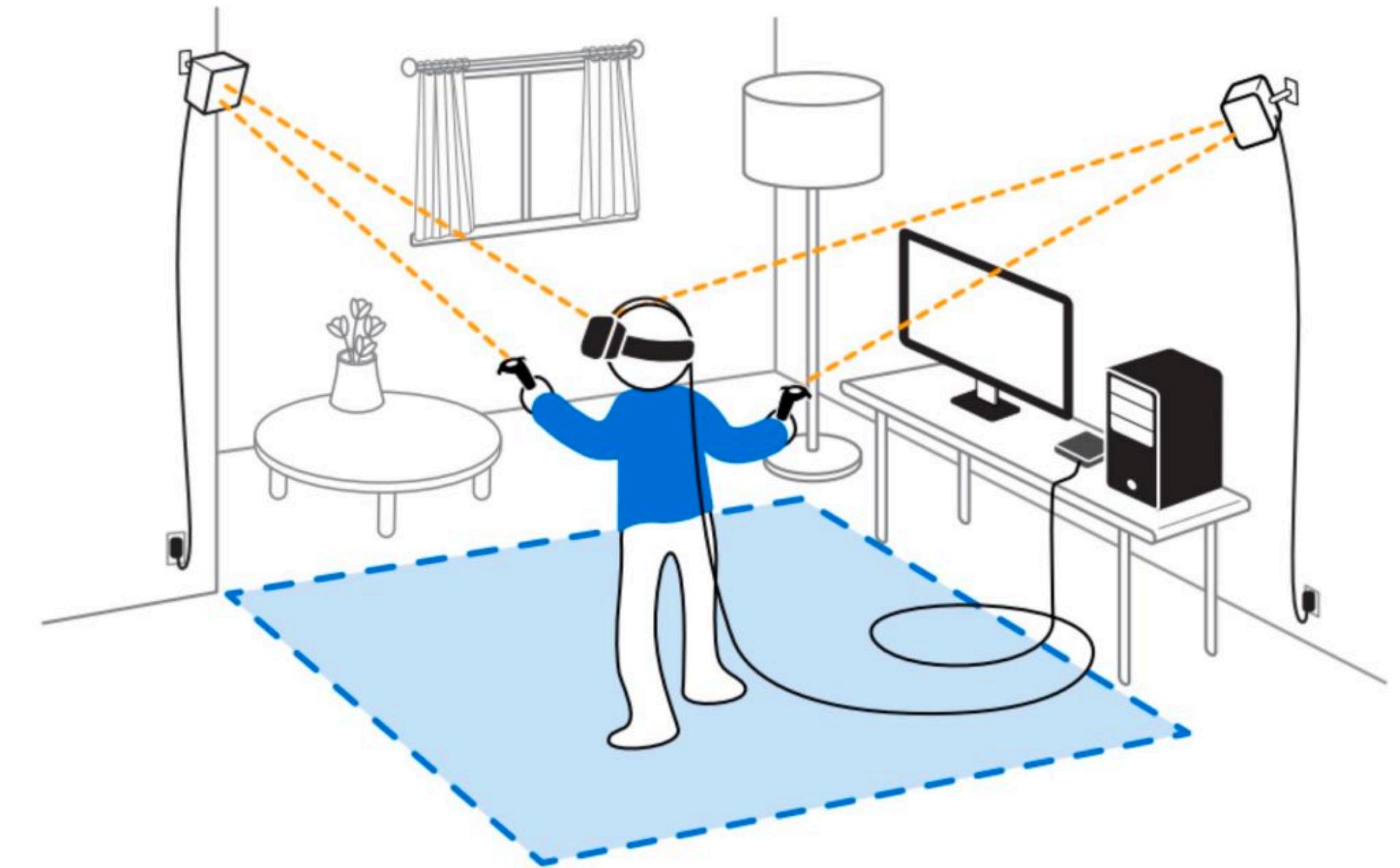
Google Cardboard (2014)

- Explicitly *not* a fully immersive experience
- Uses a phone to generate a screen for each eye
- Uses accelerometer and gyroscope for positioning



Oculus Rift & HTC Vive (2016)

- Clear out a play space
- Position sensors around the environment
- Motion controls
- An overlay headset



Microsoft HoloLens (2016)



Microsoft HoloLens (2016)

- Mixed reality system
 - Like AR, adds layers to the real world
 - Some physical objects can be interacted with
- Focused on commercial uses
 - Pricing: \$3000/unit
 - Though there are other videos which demonstrate entertainment applications



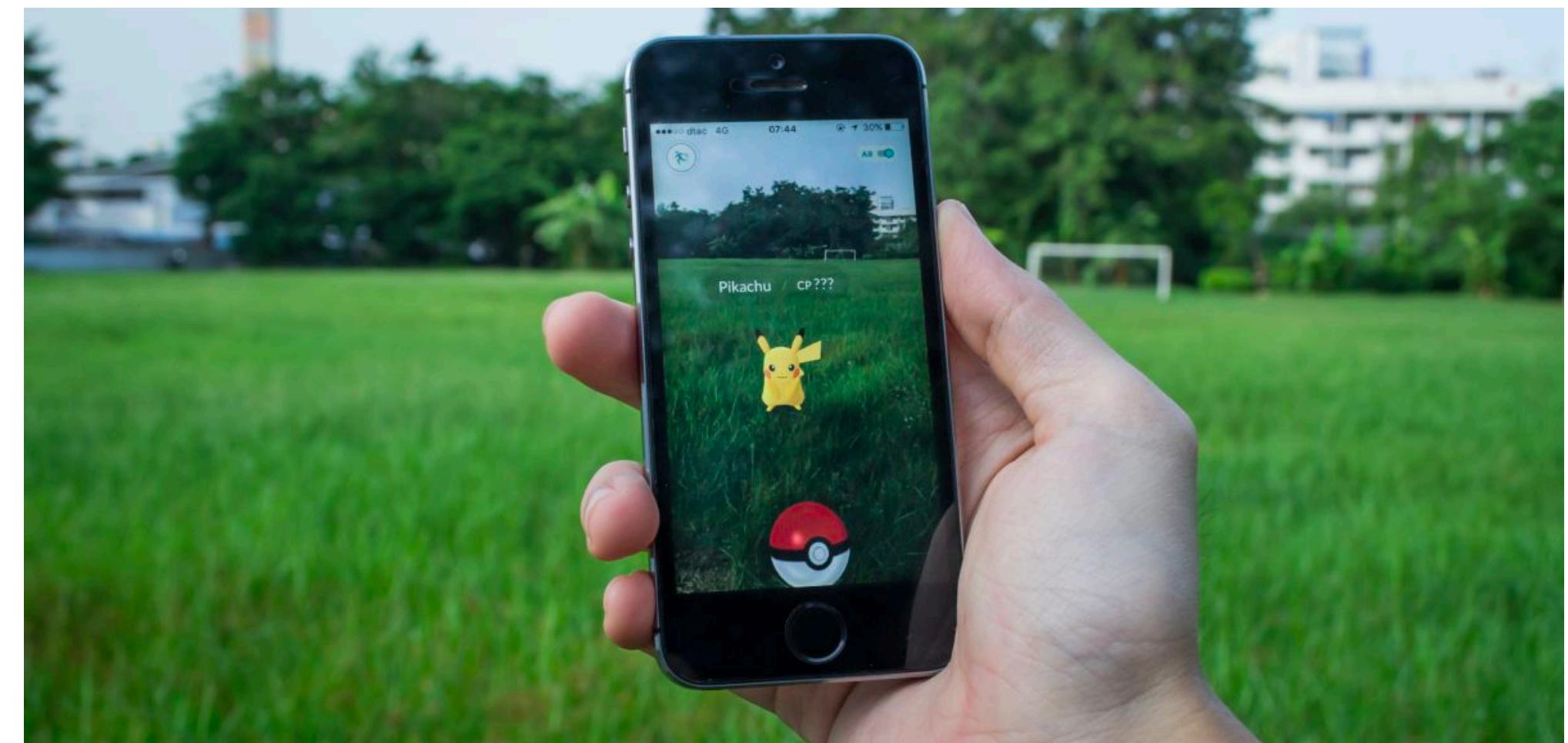
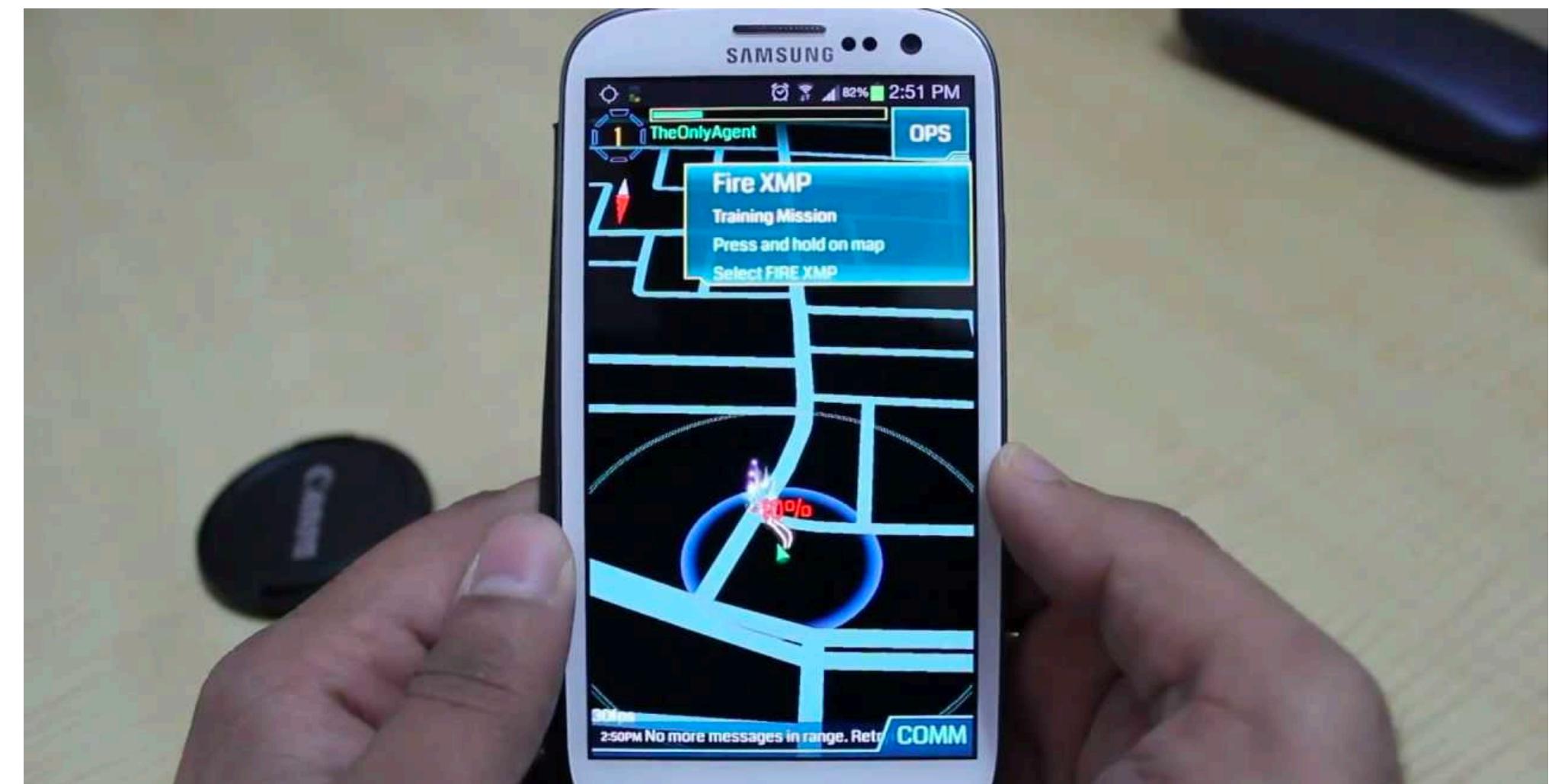
Snapchat (2011)

- A key component is augmenting the environment in fun ways
 - Mostly known for facial filters (I think? You tell me)
- A “camera company”, not a “social media company”
 - Core product is currently the social network



Ingress (2012) and Pokemon Go (2016)

- Popular games with millions of users (even today)
- AR aims to augment the real world
 - A layer over the camera is just one sense (vision)
 - Location-based PokeStops, Gyms, etc. add others
 - Location-based gameplay leads to social interactions between players



Question

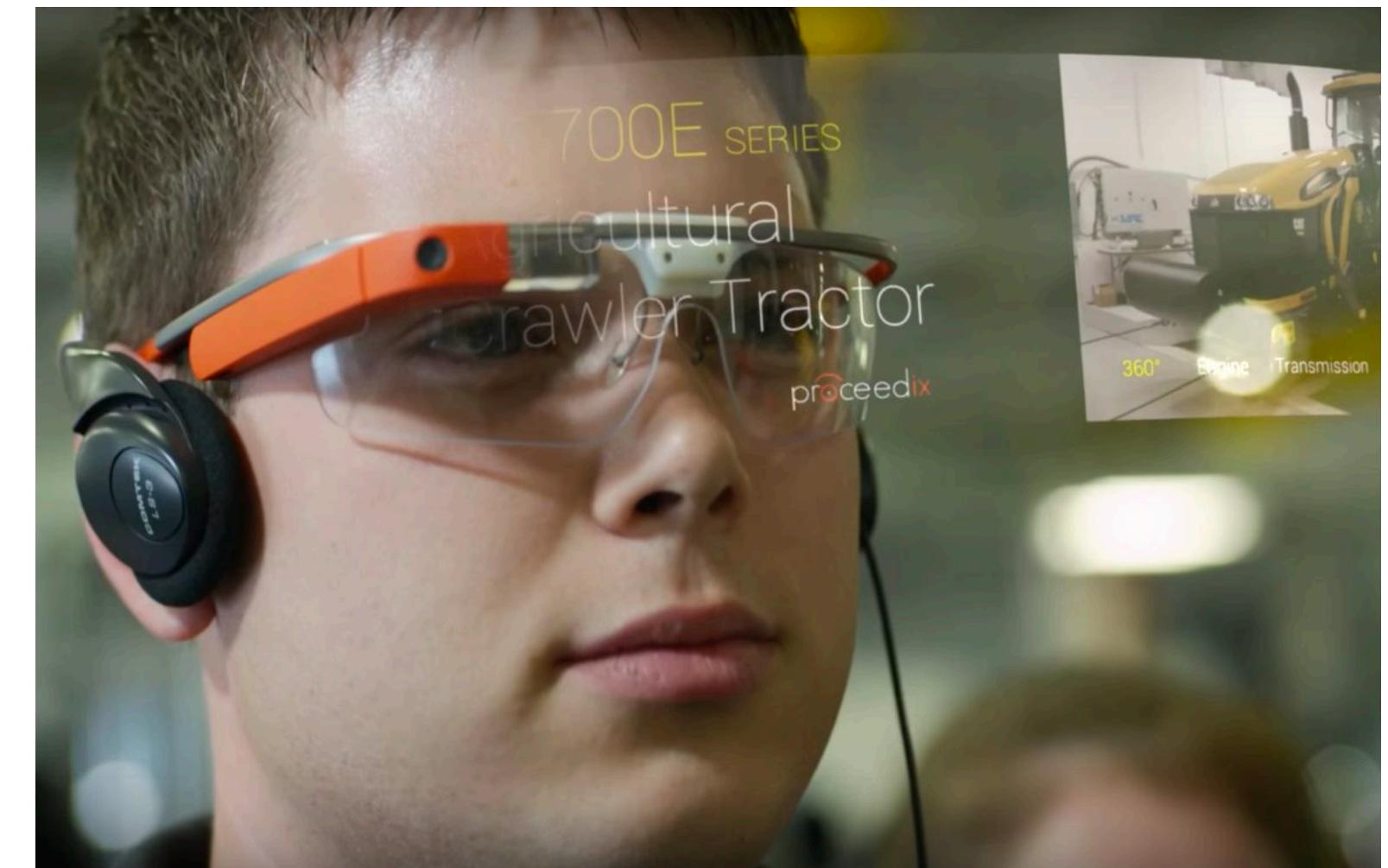


Augmented or virtual reality?

- A Both are AR
- B Cardboard is AR, Glass is VR
- C Both are VR
- D Cardboard is VR, Glass is AR
- E Both are neither AR nor VR



Google Cardboard



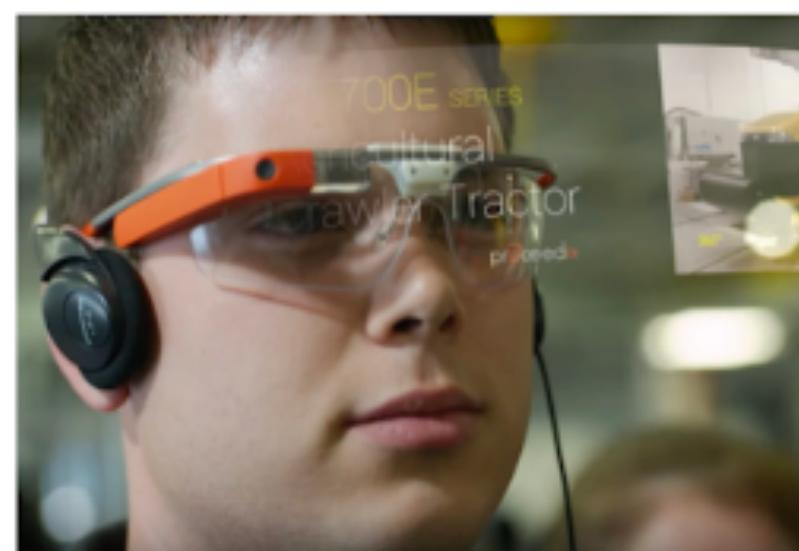
Google Glass

Augmented or virtual reality?

- A Both are AR
- B Cardboard is AR, Glass is VR
- C Both are VR
- D Cardboard is VR, Glass is AR
- E Both are neither AR nor VR



Google Cardboard



Google Glass

A

0%

B

0%

C

0%

D

0%

E

0%

Question

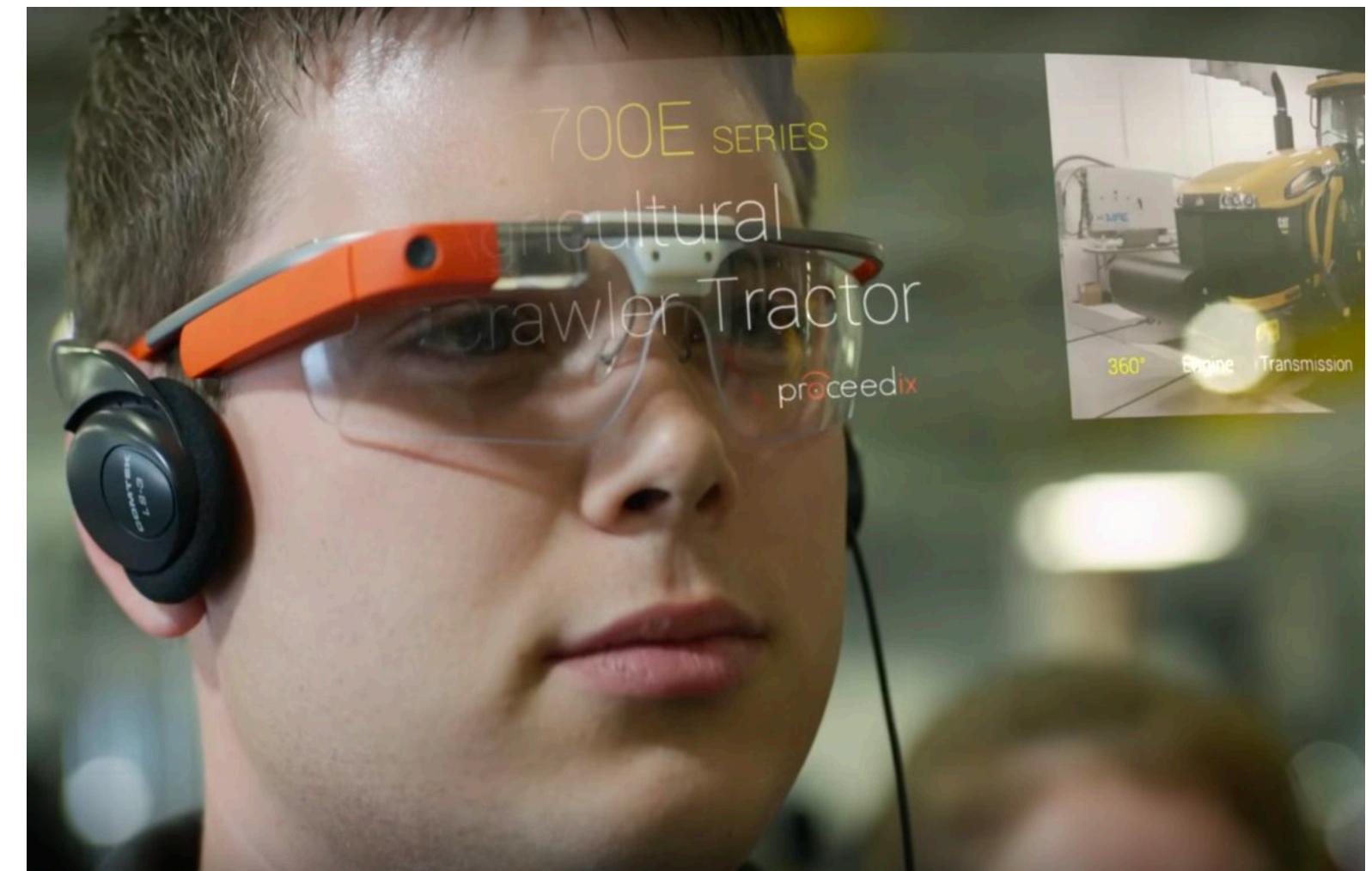


Augmented or virtual reality?

- A Both are AR
- B Cardboard is AR, Glass is VR
- C Both are VR
- D Cardboard is VR, Glass is AR
- E Both are neither AR nor VR



Google Cardboard



Google Glass

VR applications

- Games
- Flight simulators
- Sophisticated testing (car prototypes, etc.)
- Training (military, medical, industry, etc.)
- Recreation (nature hikes, exploring)

VR challenges remain consistent

- Creating a truly immersive experience is difficult
 - Need to replicate touch, sight, hearing... smell and taste?
- Most people do not have a physical space which can be taken over by a virtual environment
- Requires instrumenting the body, the environment, or both
 - Heavy, bulky, expensive, etc.
- Most practical uses are for enterprises rather than consumers

AR and Mixed reality avoid some problems

- The experience does not need to be fully immersive
- The physical space is literally part of the environment, and maybe even part of the experience
- Halfway-decent AR is pretty cheap computationally
 - It can run on your phone!
 - Other sensors like Gyroscope, etc. add to the experience
- Low cost makes AR practical for entertainment, etc.
 - Mixed reality is still very expensive

Implementing Augmented and Virtual Reality

VR Implementation

- Many common 3D development environments can be used
 - VR changes the rendering and the input
- Unity VR (and AR)
 - <https://learn.unity.com/>
- Unreal VR (and AR)
 - <https://www.unrealengine.com/en-US/vr>
- Google cardboard
 - <https://developers.google.com/cardboard>

AR Implementation

Hybrid Development

- WebXR
 - <https://immersiveweb.dev/>
- AFrame
 - <https://aframe.io/>

AR Implementation

Native Development

- Google ARCore
 - <https://developers.google.com/ar/>
- Apple ARKit
 - <https://developer.apple.com/augmented-reality/>

A few AR design recommendations

Label surfaces during set-up phase

- Highlight surfaces where objects can be placed
 - No visual cues can create confusion when a person goes to place an object
- Differentiate multiple surfaces
 - Update surfaces as a person moves their phone around



<https://developers.google.com/ar/design>

Provide guidance on interactions

- There gestures and interactions in AR are not standardized
 - Standards will come as the platform matures
- Use a combination of visual cues and text explanations to guide interactions
 - Dismiss hints when they have been performed successfully



<https://developers.google.com/ar/design>

Anticipate space limitations

- Do not require a user to walk backward or make quick sweeping motions
- Give a clear idea of the amount of space needed for an interaction
- Avoid putting large objects close to the user on the screen
 - They will instinctively back up



<https://developers.google.com/ar/design>

**Trying to make computers
conversational**

IBM Shoebox

- 1961
- About the size of a shoebox
- One of the early examples of natural language processing



https://www.ibm.com/ibm/history/exhibits/specialprod1/specialprod1_7.html

IBM Shoebox



<https://www.youtube.com/watch?v=rQco1sa9AwU>

Conversational assistants

- ELIZA, 1964
 - Used pattern matching to have a conversation
 - Not that “smart”, would just reflect back what a person said
 - But “smart enough” to potentially pass a Turing Test

```
Welcome to
      EEEEEEE  LL      IIII  ZZZZZZZZ  AAAAAA
      EE       LL      II     ZZ  AA   AA
      EEEEEEE  LL      II     ZZZ  AAAAAAAA
      EE       LL      II     ZZ  AA   AA
      EEEEEEE  LLLLLL  IIII  ZZZZZZZZ  AA   AA

Eliza is a mock Rogerian psychotherapist.
The original program was described by Joseph Weizenbaum in 1966.
This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?
YOU: Men are all alike.
ELIZA: What is the connection, do you suppose ?
YOU: They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?
YOU: Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here ?
YOU: He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU: It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy ?
YOU: █
```

<https://en.wikipedia.org/wiki/ELIZA>

Conversational assistants

- People expect to converse with computers differently from how they converse with people
 - When talking to computers, people tend to exhibit commands
 - When talking to people, people tend to be more polite

Appendix B: Example dialogue

Computer partner, sentence responses

System: Is anyone there?

User:

System: hello

User: hello

System: what do you want to know?

User: igor's degree

System: Igor's degree is a BA, in fine arts

User: igor's school

System: Igor's school was the Sorbonne

User: igor house

System: Igor's house is a commune

User: igor marry

Appendix C: Example dialogue

Human partner, short responses

Partner: Is anyone there?

=>

Partner: hello

=> howdy, are you ready

Partner: what do you want to know?

=> ok, i'm not too sure what you've got there,
but what i have here is a list of people and
i'm going to start by asking you:
what do you know about Igor?

Partner: what do you want to know?

=> what does Igor have a degree in?

Partner: BA, in fine arts

Conversational assistants

- Once Natural Language Processing (NLP) got good enough, conversational assistants started being ubiquitous
 - Siri (Apple) launched in 2011
 - Alexa (Amazon) launched in 2014
- Present on both on dedicated devices and part of apps
 - Chatbots are becoming widespread
 - Both in text and in speech



Conversational assistants

- ChatGPT
 - Driven by a Large Language Model, which should help conversations be even more fluid
 - But, interaction design is still largely the same
 - Talking to a chatbot

ChatGPT



Question



How do you feel about chatbots today?

- A I love using chatbots, and I prefer them over other interfaces
- B I'm pretty indifferent
- C I try to avoid using them when I can
- D
- E

How do you feel about chatbots today?

- A I love using chatbots, and I prefer them over other interfaces
- B I'm pretty indifferent
- C I try to avoid using them when I can
- D
- E

A

0%

B

0%

C

0%

D

0%

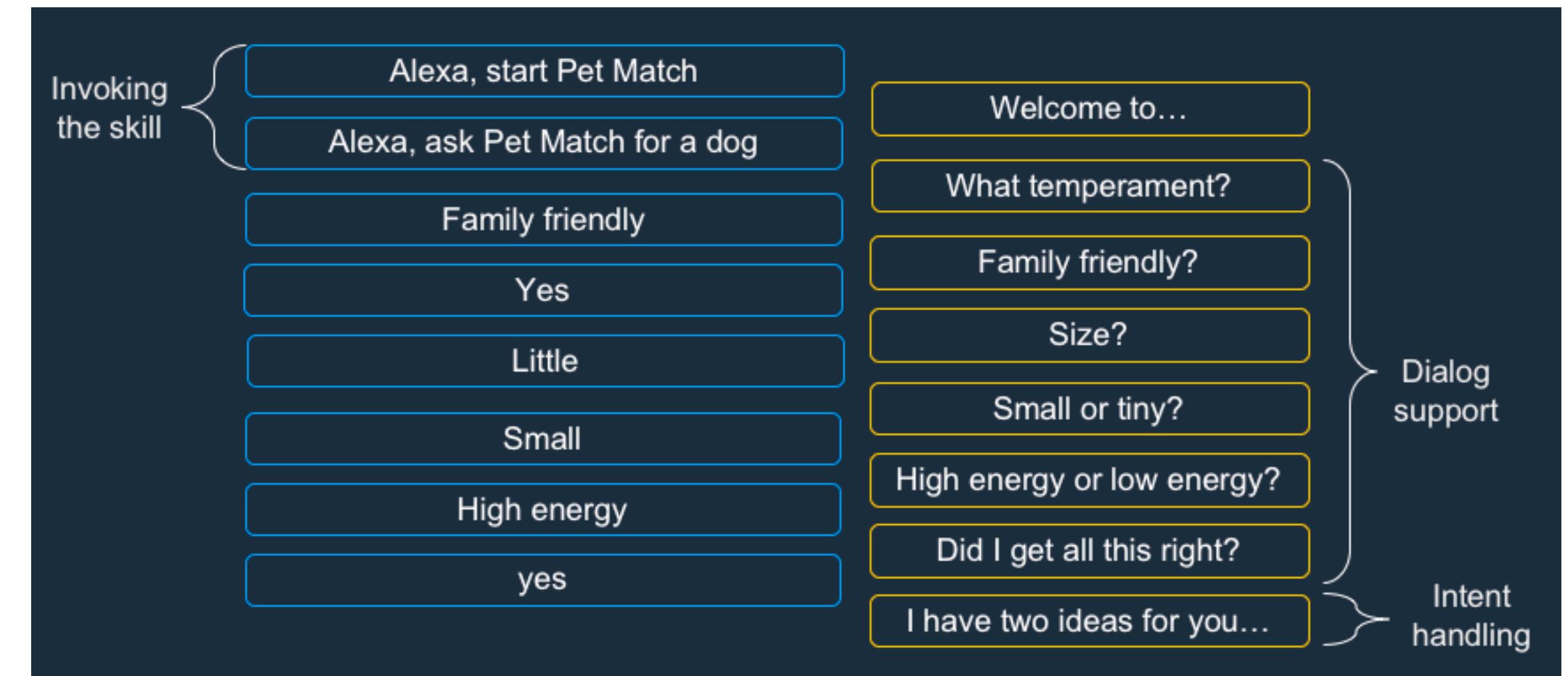
E

0%

Conversational design recommendations

Multi-turn dialogs

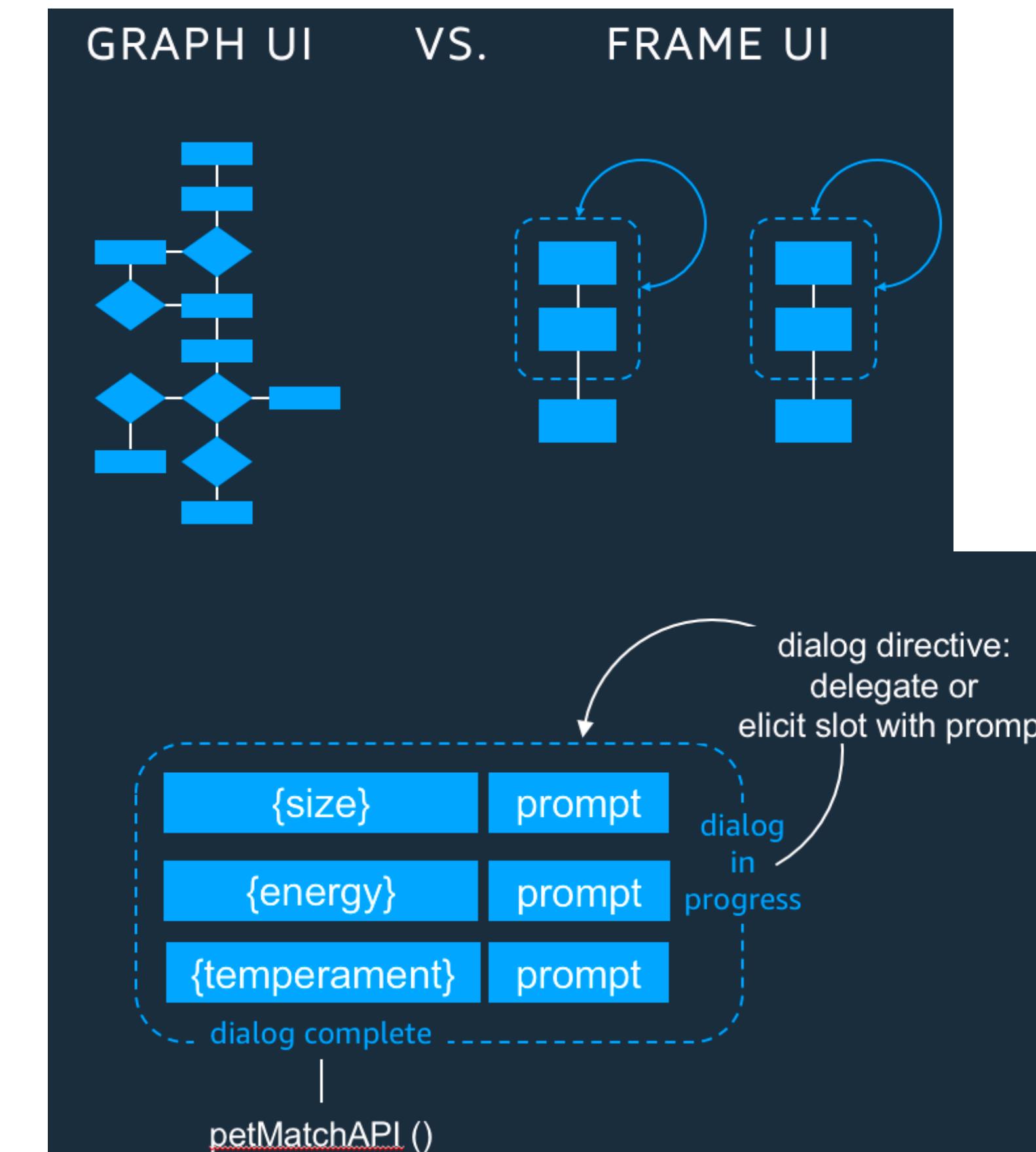
- When you have a conversation with someone, you typically ask them questions
- Voice assistants can do the same to fill up “slots” with all information
- On the web, these could be instantiated as labeled fields, but a person might be confused or turned off if the assistant just said the name of the field



<https://developer.amazon.com/en-US/blogs/alexa/post/57d0bb9c-19a6-4c51-bfa2-fc6753d14b68/4-principles-of-conversational-voice-design>

Optimize conversations

- Treat these question-answers as “frames” for updating information, rather than a graph of dialog options to traverse
 - A person might fill multiple slots at once
 - “I want a small, family-friendly dog”



<https://developer.amazon.com/en-US/blogs/alexa/post/57d0bb9c-19a6-4c51-bfa2-fc6753d14b68/4-principles-of-conversational-voice-design>

Diversify understanding

- Where you can, think through different ways a person might respond
 - Try to map those responses to your core concepts
 - If the dog should be low, medium, or high-energy
- Think of any descriptions you give as “training data” to the assistant’s model

{team}	Synonyms
high	high energy, energetic, play fetch, go hiking, that I can run with
med	plays tug of war, fun to play with, for my apartment
low	lazy, to cuddle, to watch amazon video with

<https://developer.amazon.com/en-US/blogs/alexa/post/57d0bb9c-19a6-4c51-bfa2-fc6753d14b68/4-principles-of-conversational-voice-design>

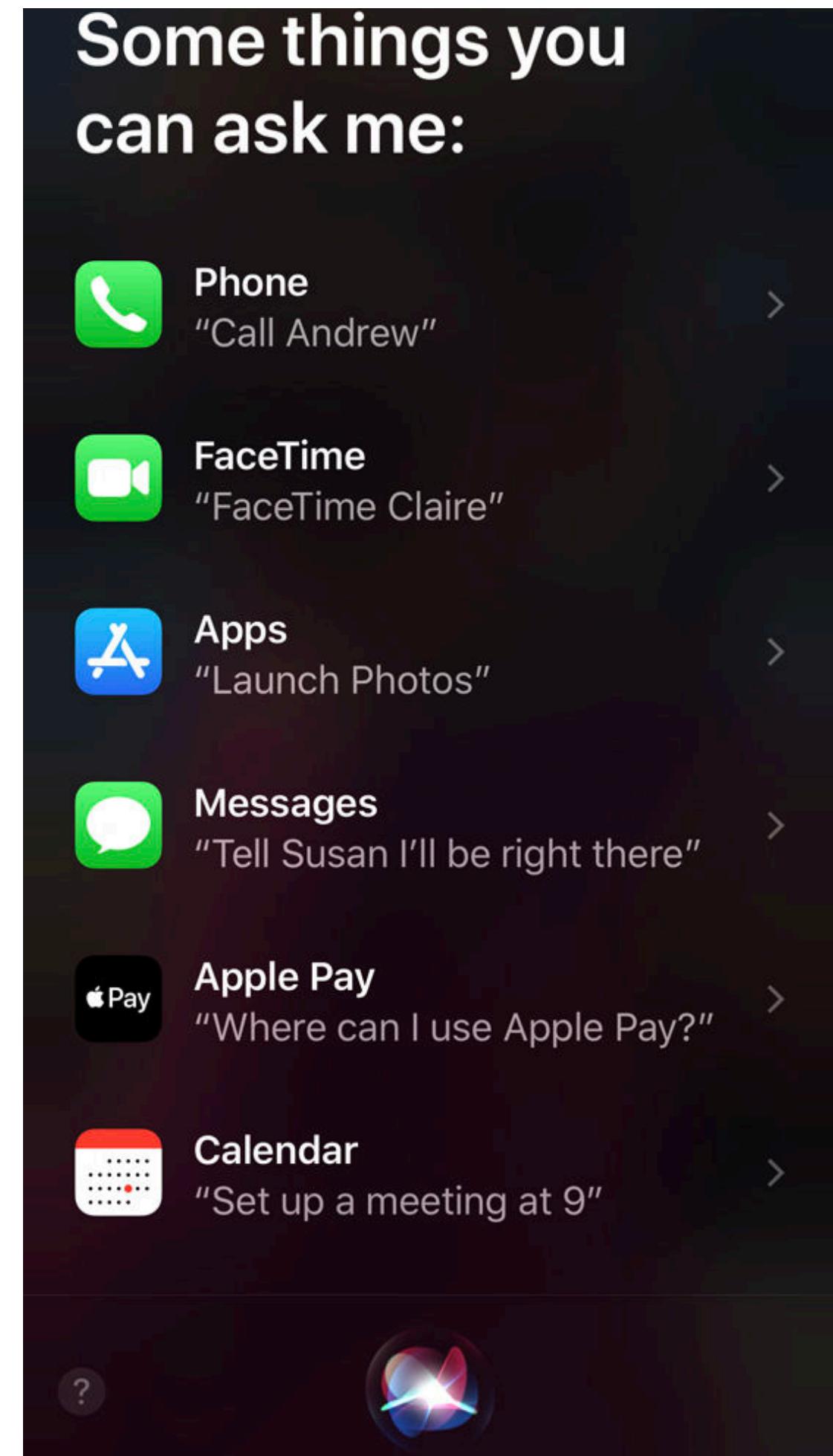
Include variance and memory

- Design your agent not to say exactly the same thing every time
 - Repetition sounds robotic
- If a person returns to a conversation, recall that as another person would
 - “Welcome back, Professor Epstein! Last we chatted, you seemed interested in a small, low-energy dog. Is that still the case?”
 - Can also sound corny, so be careful...
 - This is hard, even with LLMs

<https://developer.amazon.com/en-US/blogs/alexa/post/57d0bb9c-19a6-4c51-bfa2-fc6753d14b68/4-principles-of-conversational-voice-design>

Overall advice on conversations

- Discoverability is hard!
 - People can't "see" what conversational options (commands) are available to them
 - We can show people a list of options, but they're likely to forget
 - They're likely to not notice if new functionality is added

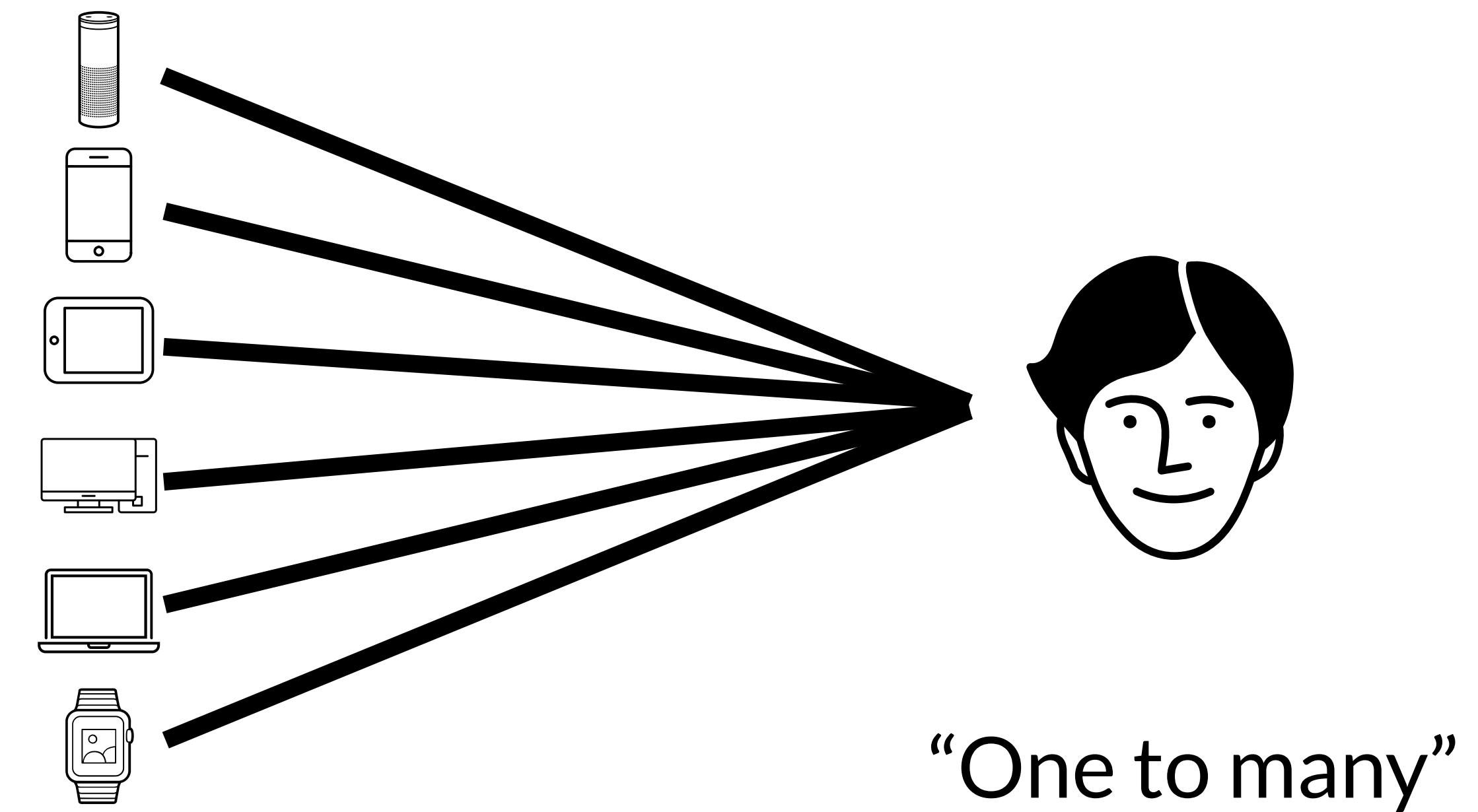


But what about LLMs?

- Many of the same challenges persist
 - Need training data to personalize an LLM to your use case
 - Hard to discover how the underlying model has changed
- Conversation is just a tricky format for interaction
 - People prefer to directly indicate their intention, like via pressing buttons

Reflections on beyond web and mobile

- Devices offer different interaction affordances and are used in different contexts
- Web and mobile are often sufficient
 - Not every app needs an Alexa skill, Watch app, or TV app
- As we've moved to the “one to many” Ubiquitous Computing era, we’re more often separating computation from interaction



Today's goals

By the end of today, you should be able to...

- Describe some of the history of AR and VR and conversational interfaces
- Differentiate augmented and virtual reality
- Explain key principles for designing a good AR experience
- Articulate some principles for designing conversational experiences

IN4MATX 133: User Interface Software

Lecture 18:
Augmenting Reality &
Conversation