Building a Socialbot: Lessons Learned from 10M Conversations

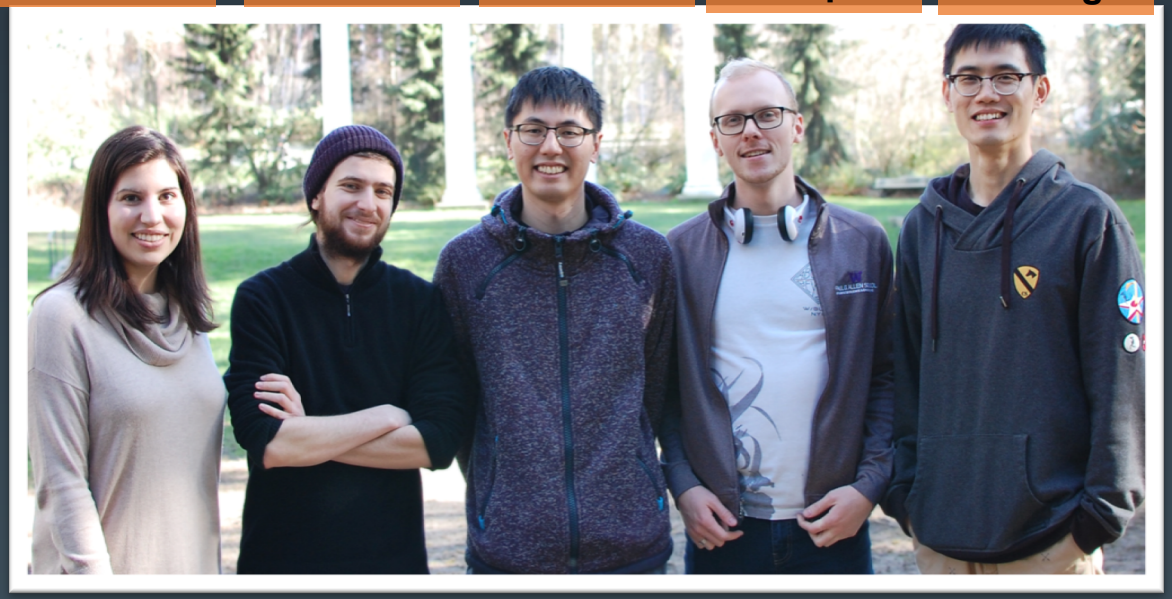
Mari Ostendorf
University of Washington

ELECTRICAL & COMPUTER ENGINEERING

Acknowledgments

UW conversational AI system: Sounding Board

Elizabeth Clark | Ari Holtzman | Hao Fang | Maarten Sap | Hao Cheng



Advisors: M. Ostendorf, Y. Choi, N. Smith

Amazon Alexa Prize

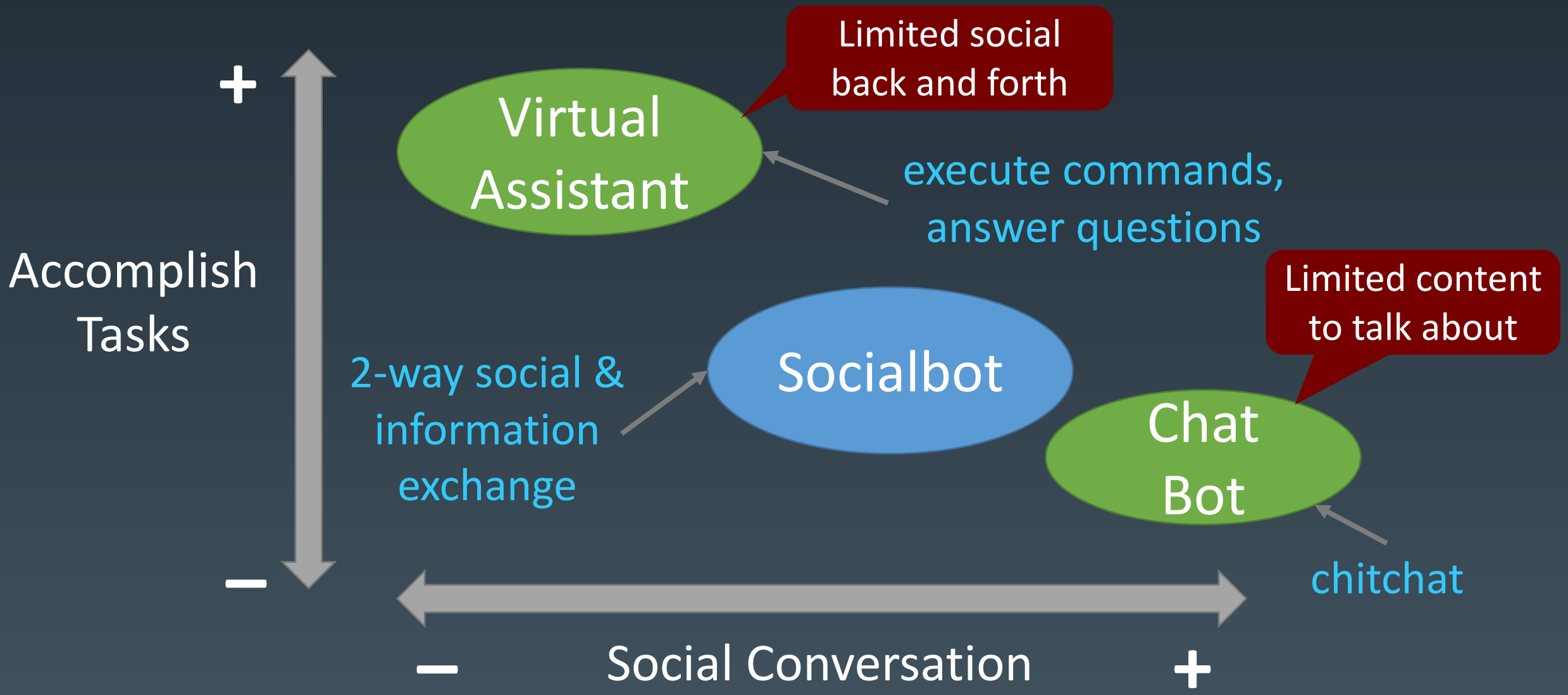
Teams of university students try to build a socialbot that

- converses coherently and engagingly with people
- on popular topics and current events.

Roadmap

- Context: conversational artificial intelligence
- Sounding Board system
- Characteristics of real users & leveraging them
- Parting words

Types of Conversational AI Systems



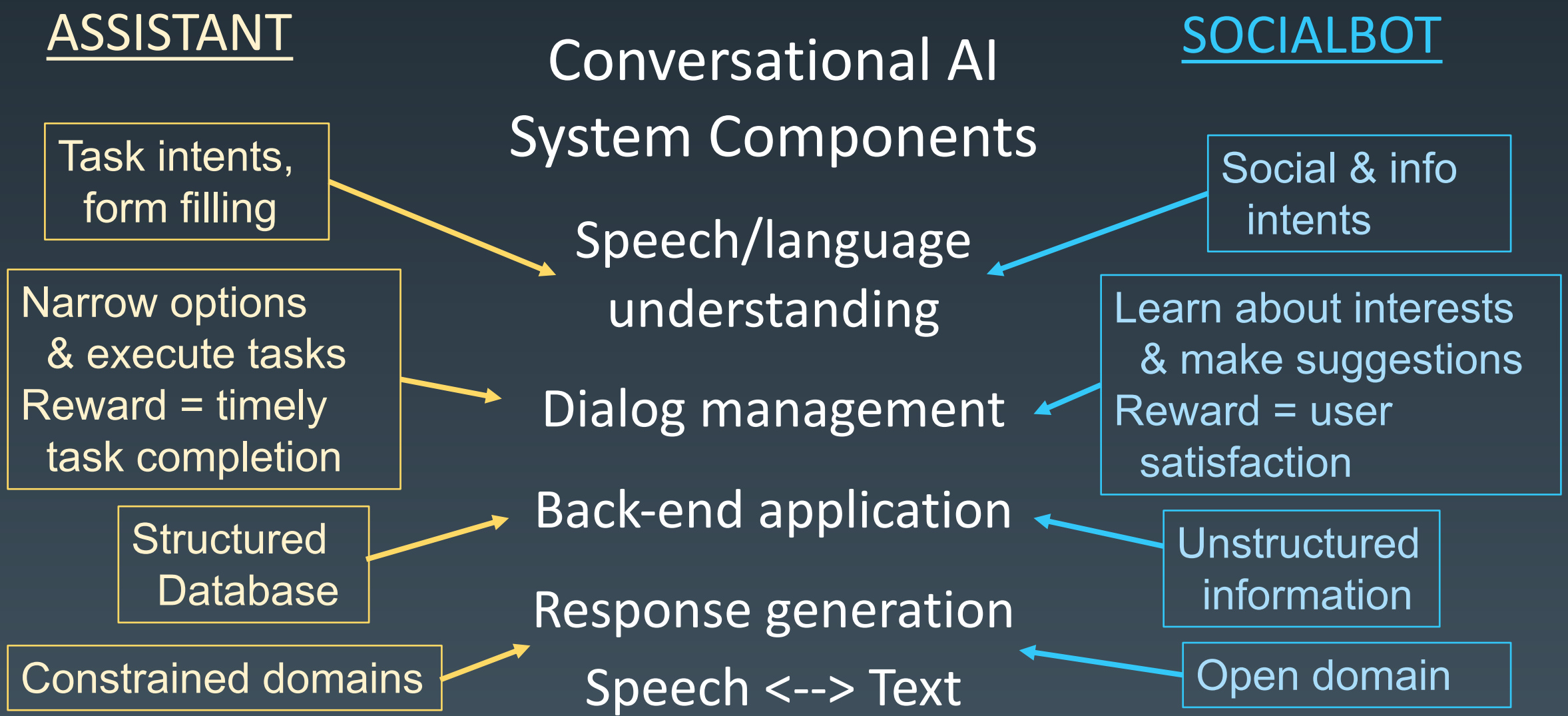
Sounding Board:
A Conversational Gateway to Online Content

Sounding Board

Turn	User	Sounding Board
1	Let's chat	Hi! This is an Alexa Prize socialbot. How's your day?
2	I'm five.	Glad to hear it! We could talk about robots, batman, or superman. Say "next", to chat about other things.
3	Superman.	It looks like you wanna chat about superman. I was high up in the cloud when I realized: If people don't recognize Clark Kent as Superman because of his glasses, does his eye doctor know his true identity?
4	I guess so.	Did you know that Henry Cavill almost missed the call for the role of Superman cause he was playing "World of Warcraft"
5	Really, I didn't know that.	Weird, right? Speaking of superman, Did you see: Superman? It was released in 1997. It's a comedy.
6	Yes, it was hilarious.	I'm glad you feel this is hilarious. Which part do you like best about this movie?
7	The part when he met lewis leah.	Interesting. meccartin, and raffi co-directed this film. The movie has a 6.3 out of 10 on IMDB. which seems pretty good! do you like the movie's director?
8	Yeah, he's cool.	Cool! By the way, do you wanna know something interesting about raffi?
9	Okay, go ahead.	Super! My friend told me that Raffi is a worldwide children's rights activist and has fought against exploitative advertising to children.

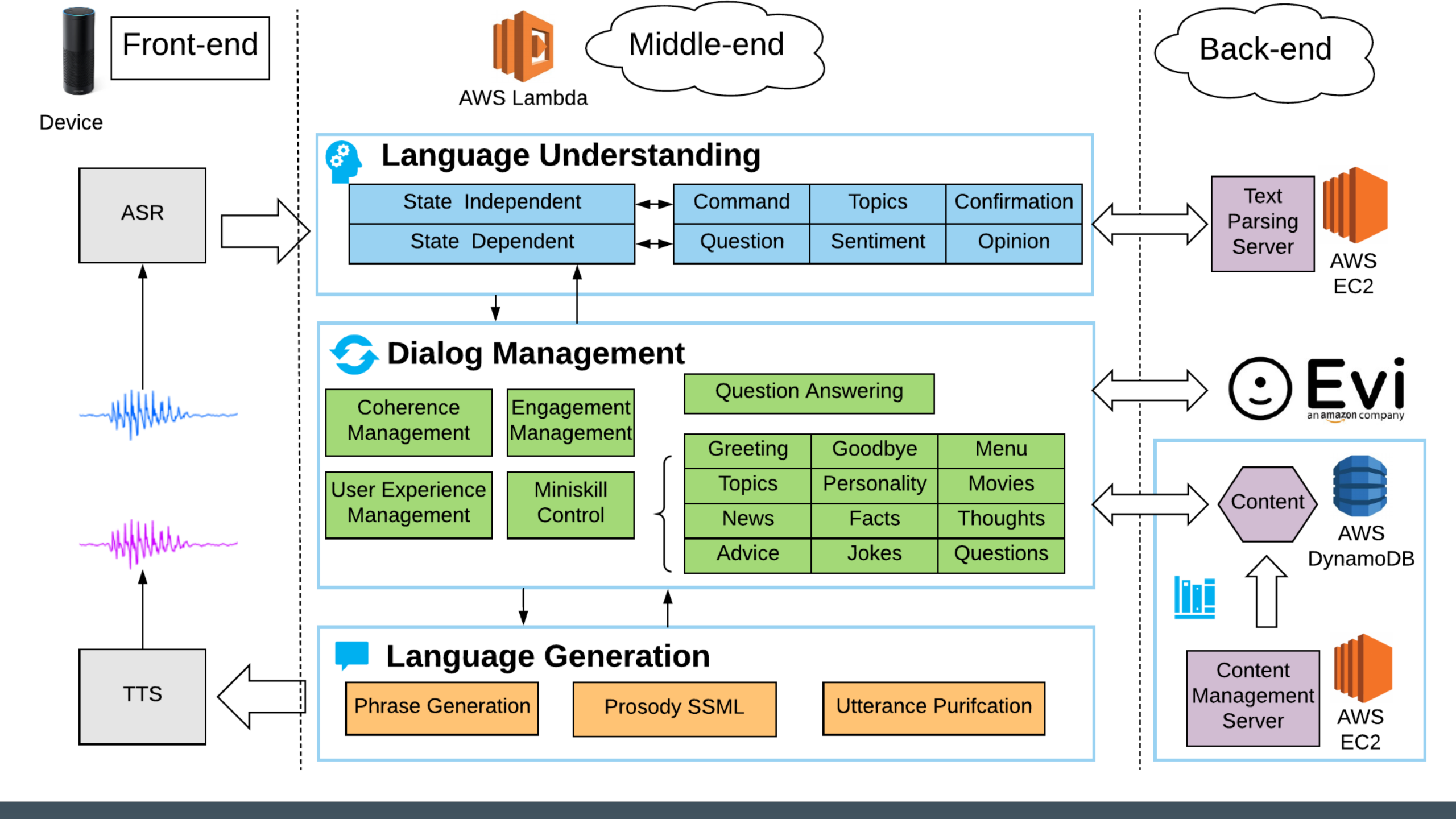
Sounding Board – System Overview

Differences in Conversational AI Paradigms



Design Philosophy

- Content-driven
  - Daily content mining, large & dynamic content collection
  - Knowledge graph
  - Dialog manager promotes popular content, diverse sources
- User-centric
  - Language understanding detects user sentiment
  - Try to learn user personality & interests for content selection
  - Language generation with grounding


Front-end
Device
AWS Lambda
Middle-end
Back-end
ASR
Language Understanding
State Independent
State Dependent
Command
Topics
Confirmation
Question
Sentiment
Opinion
Text Parsing Server
AWS EC2
Dialog Management
Coherence Management
Engagement Management
User Experience Management
Miniskill Control
Question Answering
Greeting
Goodbye
Menu
Topics
Personality
Movies
News
Facts
Thoughts
Advice
Jokes
Questions
Evi
an amazon company
Content
AWS DynamoDB
Content Management Server
AWS EC2
TTS
Language Generation
Phrase Generation
Prosody SSML
Utterance Purifcation

Hierarchical Dialog Management

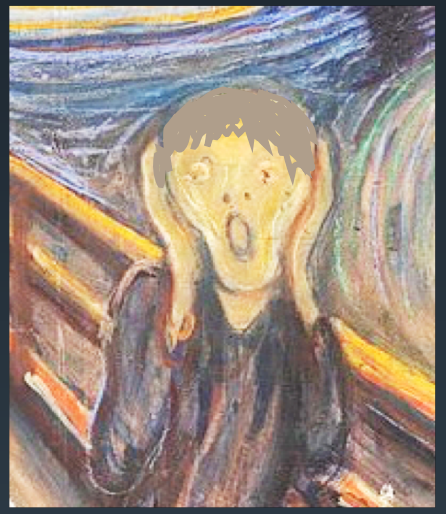
- Master (Global)
  - Rank topics, miniskills, content
  - Consider: topic coherence, user engagement, content availability
- Miniskills (Local)
  - greeting / goodbye / menu
  - probe user personality
  - discuss a news article / movie
  - tell a fact / thought / advice / joke

Evaluation – AlexaPrize Protocol

- Users (judges) rate conversations on a 1-5 scale
- Duration of conversation (tie breaker)

	Rated by	Rating	Duration
Finals (mid Nov 2017)	Judges	3.17	10:22
Post-finals (Nov 24 – Dec 24)	Alexa owners	3.65	4:30

So what's a nice EE like me doing in a field like this?

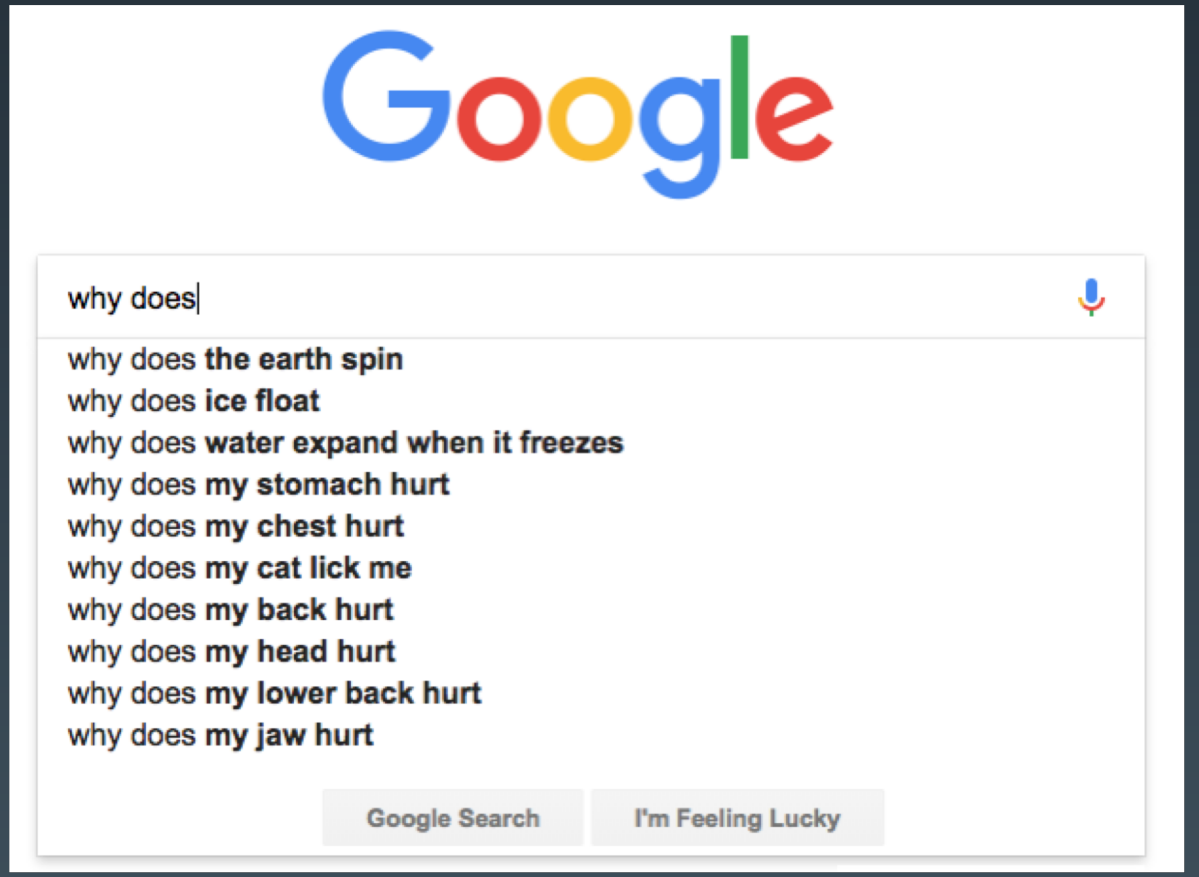


- I'm an expert in speech signal processing and I can't use the audio signal! (for privacy reasons)
- Actually, a word sequence is a signal
  - Project discrete word representations to continuous vector space
  - Integrate language, image, audio in multi-modal processing
  - Leverage signal processing methods in neural models:
    rank reduction, sparsity, basis vectors, adaptive models…
- I get to work with data from real users

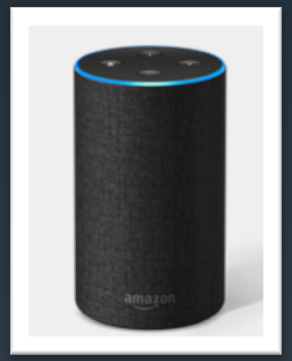
Characteristics of Real Users
(from 10M Sounding Board Conversations)

Importance of User Modeling

- User interests vary
- For text-based search, personalized query completion improves over popular queries
- In conversational search, user language can augment topic history



Content Preferences Vary



Did you know that Malaysian vampires are tiny monsters that burrow into people's heads and force them to talk about cats?

Amused

Oh you are so funny.

Oh my god that's funny.

Not amused

That's not true.

Oh gods are you have to hear this.

What the heck.

Not really listening?

Cool.

Wow that's interesting.

That's awesome.

Cat lover

Cats are my favorite animals.

Let's talk about cats.

That's creepy.

Interaction Styles

Talkative User

- So you're saying mean is a type 2 diabetic art is junk food
- No I'd rather go back to talking about vampires
- I love vampires the something that I really loved do you know and rice the author
- it is weird i don't understand
- Yes I would
- i don't know that's an interesting question and is it really true that garlic keeps vampires the wedding and what i
- what are they have their long fingernails for
- i think that that's probably true but i think it vampires are evil and they don't care about sustaining things for human be-...

Terse User

- Cool
- One excuses enough
- No thanks
- No
- No
- Sure
- Yeah
- Yeah
- No I didn't know that
- No
- Yeah
- No

- No
- Yes
- Cool
- Yeah that's cool
- No I didn't
- No
- Yes
- No I didn't know that
- No
- Yes
- Cool
- ...

Different User Goals

- Information seeking

Tell me more about that.

What is that?

Why?

i don't know that's an interesting question and is it really true that garlic keeps vampires the wedding and what i

- Opinion sharing

i think that that's probably true but i think it vampires are evil and they don't care about sustaining things for human be-

- Getting to know each other

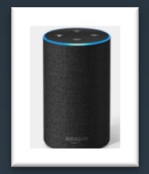
What’s your favorite video game?

My favorite video game is minecraft

Adversarial User Examples

- Taboo topics

Alexa, are you naked?



I'm a chatbot.

- Offensive language

@#%*#*!!.

I heard you mention unicorn. I'll do some research on that. How about we talk about some news about artificial intelligence?

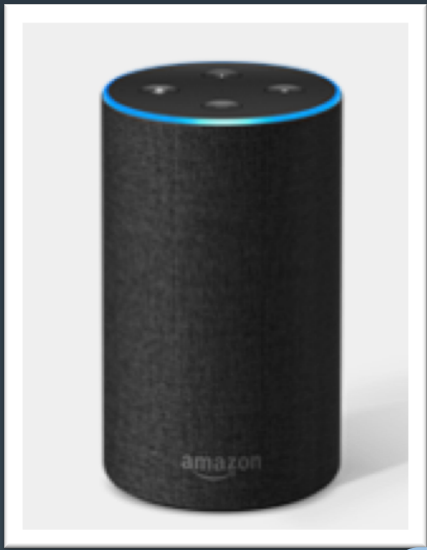
- Venting

You suck.

Unfortunately, I'm not sure I understood what you said.

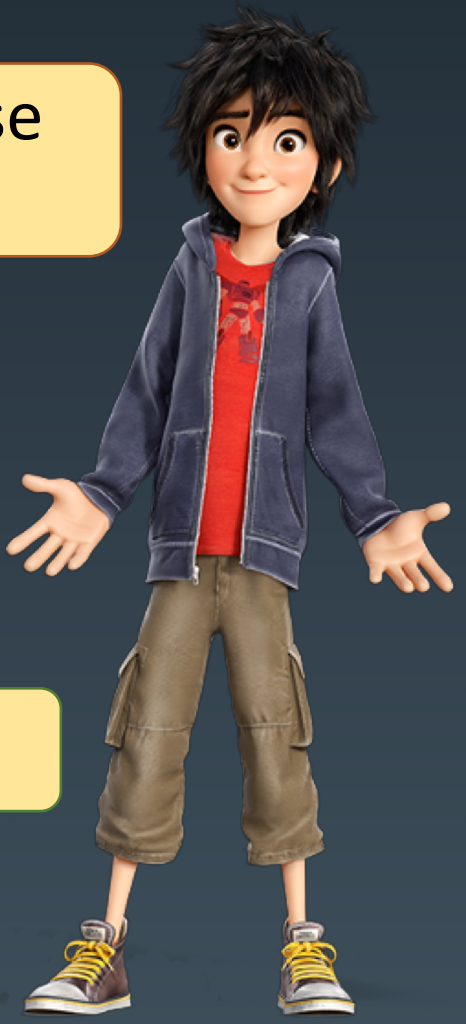
Challenges with children ….

yeah can i get my butt hampshire suspense are there was a cough sure stop



No problem. Let me think. How about we chat about…

Let’s talk about Santa Claus!



You know what I realized the other day? Santa Claus is the most elaborate lie ever told.

https://www.verywellmind.com/the-big-five-personality-dimensions-2795422

User Personality Types

- Include personality quiz questions in the conversation
- Personality correlates with user ratings
  - Extroverted, agreeable, open -> higher ratings ($p << .001$)
- Topics brought up by users
  - Introverts (AI, food, cats), extroverts (news, fashion)
  - Open & imaginative (AI, time travel, aliens)
  - Low conscientiousness (pokemon, video games, minecraft)

Implications

Spoken Language Understanding

Age & dialect impact ASR;
verbosity impacts NLP

Dialogue Management

Content ranking (sources, topic, entities)
Error handling, follow-up strategy

Content Management

Flag rather than filter controversial content
Multi-dimensional content index (e.g. ratings of user types)

Language Generation

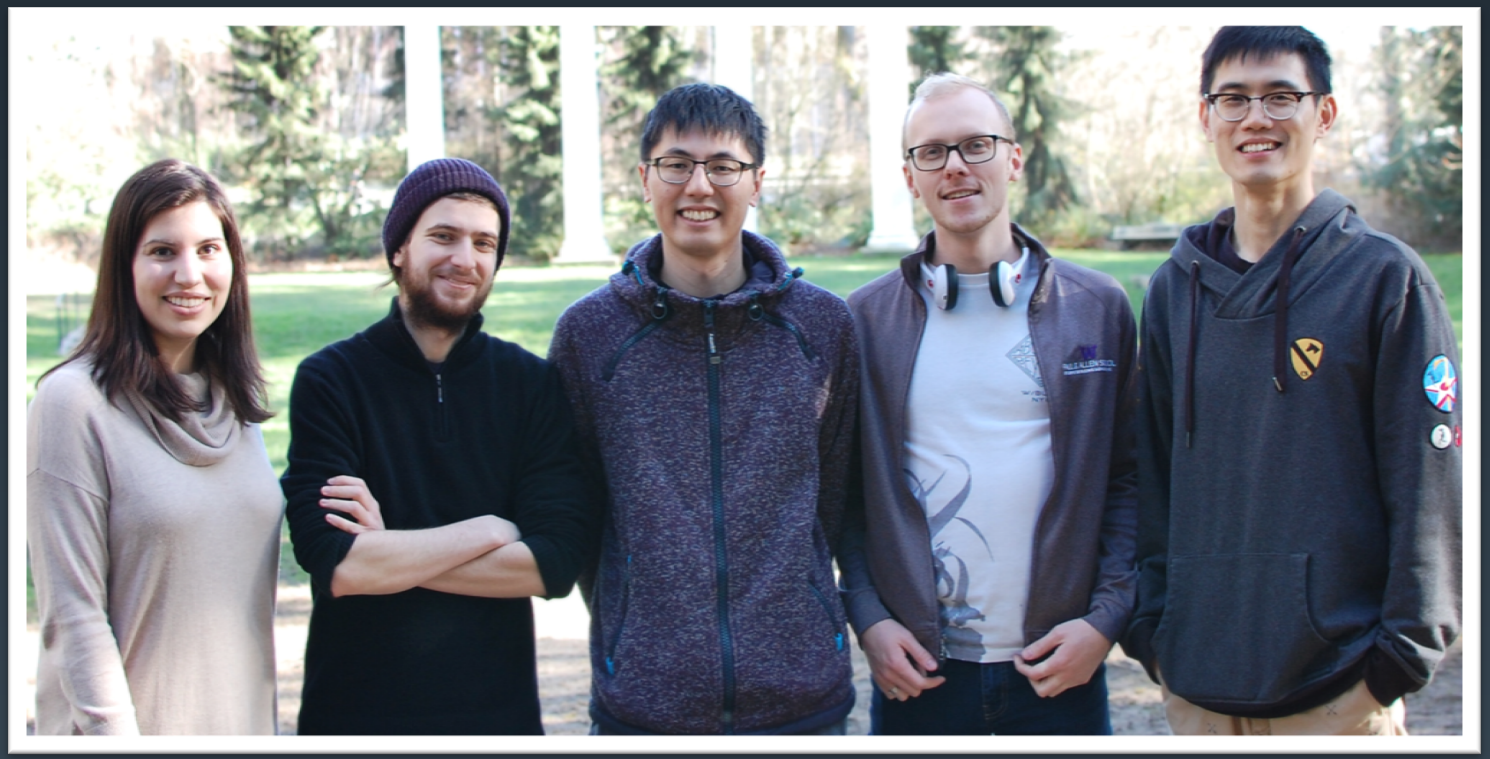
Politeness, repetition rephrasing

Socialbot User Modeling

- Problem: Content recommendation, leveraging
  - Engagement with content earlier in the dialog
  - Age/personality factors reflected in language use
- Challenge: *User-topic engagement data is sparse*
- Solution: Dynamic user state
  - Map user language & topic content to vector spaces
  - Recurrent update of user state vector; input = vectors of user language, accepted & rejected topics
  - User state + content vector → content ranker

Sounding Board – Parting Words

- A socialbot can serve as a conversational gateway to the web, but there are many other applications
- Critical system capabilities
  - Track user intent & engagement, learn user interests
  - Manage dynamic, unstructured content
- ML Challenge: Dialog policy learning given
  - A sparse, subjective reward signal
  - Evolving system capabilities & user behavior



Thank You