

Introduction to Conversational Artificial Intelligence

Conversational Artificial
Intelligence -- EE596/LING580

Hao Cheng
University of Washington



Agenda

- Course Overview and Logistics
- Conversational AI Overview
- UW Sounding Board -- Amazon Alexa 2017 Winner

Course Goals

- Understand basic concepts in conversational AI
- Learn recent advances in conversational AI
- **Build a social chatbot**
- Expected Background
 - Courses: Natural Language Processing / Artificial Intelligence / Machine Learning
 - Programming Language: Python
 - System: Linux / MacOS / Windows Subsystem for Linux
 - Tools: Git

About the course

- Course website: https://hao-cheng.github.io/ee596_spr2019/
 - Syllabus
 - Labs descriptions
 - Resources
- Team registration/report submission/course discussion: Canvas
- Labs: **BRING YOUR LAPTOP**
- Project: up to 4 members each team

Course & Office Hours

- Lectures
 - Tuesday 4:30pm – 6:20pm
 - EEB 026
- Labs
 - Thursday 4:30pm – 6:20pm
 - EEB 026
- Hao Cheng (Instructor)
 - Tuesday 3:00 – 4:00pm
 - EEB 431
 - chenghao@uw.edu
- Trang Tran (TA)
 - Thursday 3:00 – 4:20pm
 - SIEG 126A
 - ttmt001@uw.edu

Class Format

- Lectures on Tuesday
 - Lectures cover basic concepts and recent advances in conversational AI
 - We will have 2 or 3 guest lectures in this quarter
 - Paper discussion and presentation led by students (weeks 8 & 9)
- Labs on Thursday
 - Each lab has 2 sessions (2 weeks)
 - In-class lab checkoff in either session (no checkoff after class)
 - Lab reports due on the Tuesday after the 2nd lab session
- Final week (June 6/June 7)
 - Project presentation, system demo, and poster session
 - We may invite external visitors

Syllabus: https://hao-cheng.github.io/ee596_spr2019/syllabus.html

Grading Policy

- Labs: 40% (4 x 10%)
 - Lab checkoff: 5%
 - Lab report: 5%
- Paper presentation: 10%
- Project proposal: 10%
- Project demo and presentation: 20%
 - Every team member needs to participate in the presentation
 - Systems will be evaluated by judges (peer students & external panel)
- Project report: 20%
- **No late submission!**

April 9, 2019	Team registration
April 16, 2019	Lab 1 report
April 23, 2019	Project proposal
April 30, 2019	Lab 2 report
May 14, 2019	Lab 3 report
May 28, 2019	Lab 4 report
June 14, 2019	Project report

Collaboration Policy

- Students should work as a team on labs, presentations, and projects
 - Each team should submit a **single** lab/project report together (we will set up the Canvas for all assignments)
 - The report should include information about the structure of the collaboration (i.e. who did what)
- Use of an open-source software must be identified and credited in comments in your codes

Lab & Project Platform

- Primary programming language: Python3
 - Python 2.7 will not be maintained after 2020
- Device: Echo dot
 - Each team will get 1-2 devices.
 - You need to return the device on June 6, 2019 (last class).
- Amazon Alexa Skill Kit (<https://developer.amazon.com/alex-skills-kit>)
- [Optional] Amazon AWS Lambda (<https://aws.amazon.com/lambda/>)
- Lab 1 will help you go through these steps!

Team Registration (due on April 9, 2019)

- Submit on Canvas using a 1-page DOC or PDF
- Team name
 - The team name will also be the name of your bot
 - <https://developer.amazon.com/docs/custom-skills/choose-the-invocation-name-for-a-custom-skill.html>
- Team members (up to 4)
 - every team should have one LING student & one ENGINEERING student
 - 5 LING students currently
- Team vision (1-2 paragraphs)
- You must join a team for this course

Project Proposal (due on April 23, 2019)

- Motivation
 - Explain and motivate the system you are going to build
 - Discuss the target users
 - Examples of user cases
- System overview
 - System architecture
 - High-level description of individual modules
- Potential data, model, evaluation
 - Optional for the proposal, but required for the final report
- Timeline
- Description of Team Expertise

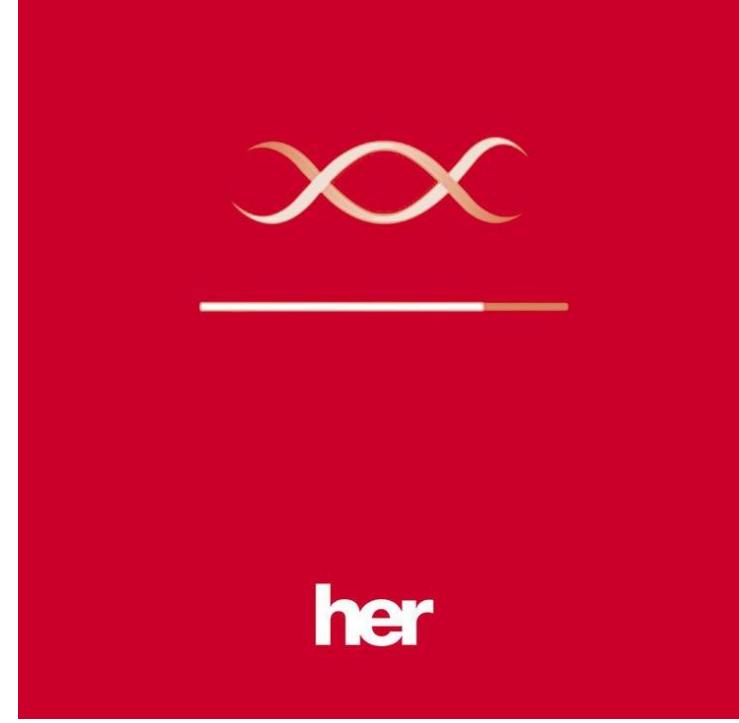
Project Report (due on June 14, 2019)

- Example outline
 - Introduction
 - System description
 - Data
 - Model
 - Evaluation
 - Discussion
- More details to follow towards the end of the course

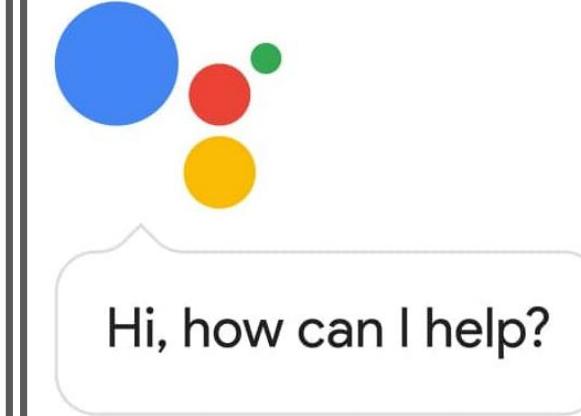
Questions?

Next: Student Self-Introduction

Conversational AI



Sci-Fi Movies

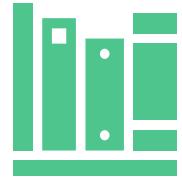


Daily Life

Types of Conversational AI



**Task
Definition**
task-oriented
non-task-oriented

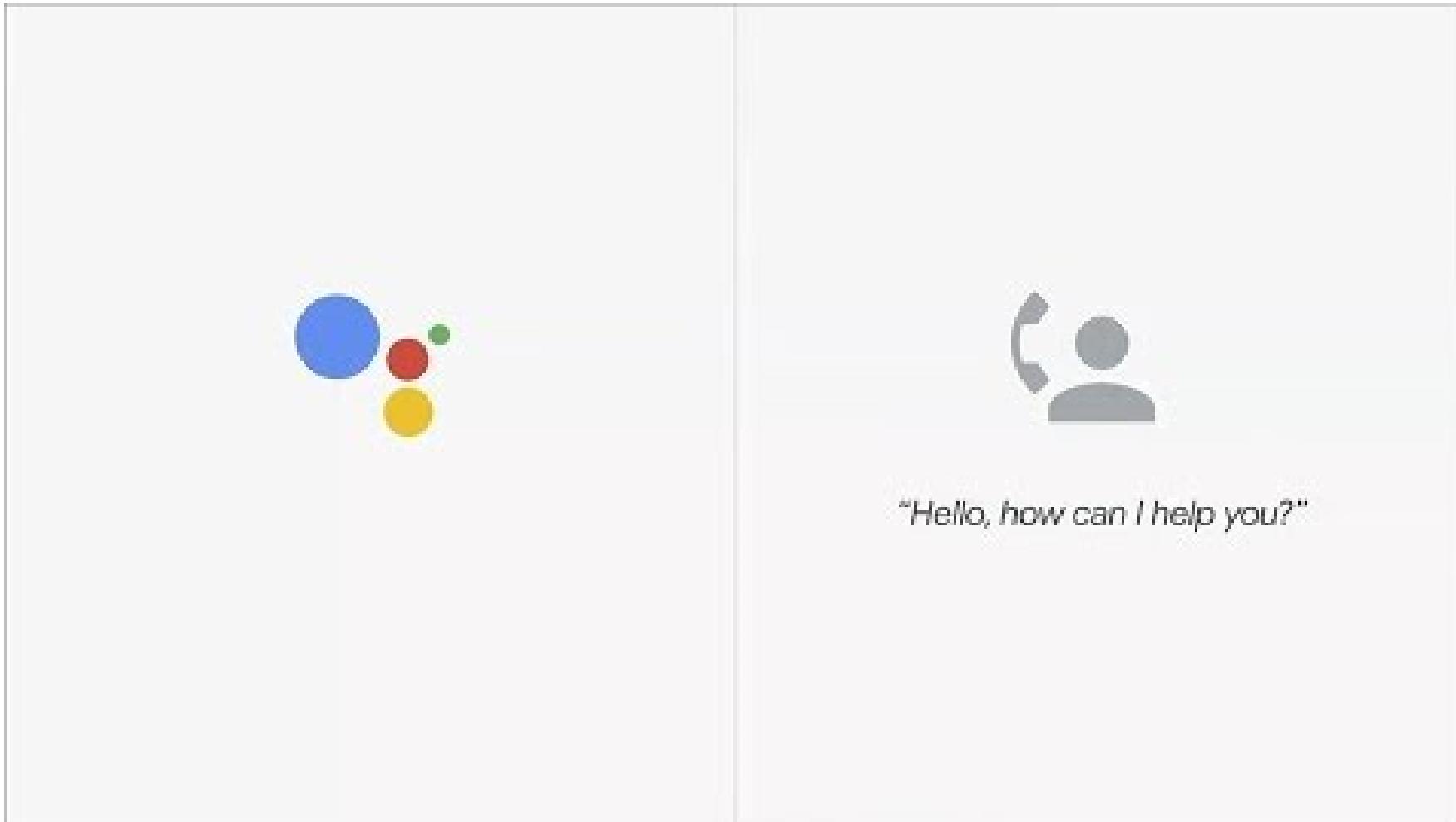


**Domain
Coverage**
single-domain
multi-domain
open-domain



**Dialog
Initiative**
system-initiative
user-initiative
mixed-initiative

Task-Oriented System (Google Duplex)



Chatbot Example (Cleverbot)

User: Hello.

Cleverbot: I mean besides that?

Cleverbot: How are you?

User: Enjoying sunshine.

User: I'm good how about you.

Cleverbot: No, there's autumn, so it's raining all the time.

Cleverbot: Good,!

User: What.

User: {pass}

Cleverbot: I love you!

Cleverbot: What are you up to?

User: What do you mean.

User: Talking with you.

Cleverbot: Are you a human?

User: I guess so.

Education Robot (Language Acquisition)



Break (10 min)

Next: More on Conversational AI

Try some online chatbots if you want.

https://hao-cheng.github.io/ee596_spr2019/resources.html

Task-Oriented vs Non-Task-Oriented Systems

- **Task-oriented systems** interact with users to accomplish specific tasks
 - Simple and well-defined tasks
 - e.g., hotel or flight booking, restaurant reservation
 - Complex tasks involving sophisticated planning and reasoning
 - e.g., holiday planning, contract negotiation
 - Sometimes called “goal-oriented systems”
- **Non-task-oriented systems** engage users in conversations that do not necessarily involve a task to be accomplished
 - Usually referred to as “chatbots”, or “non-goal-oriented systems”
- Naturally, a conversation can involve a mixture of task-oriented and non-task-oriented interactions with smooth transitions.

In practice, there may be some overlap between these two types of systems.

Definition of Domain

- For task-oriented dialog systems, the domain is essentially equivalent to the task (e.g., restaurant search, flight booking).
 - A database containing all available information about the task.
 - A domain-specific ontology, i.e., structured representation of the database.
 - Recently, there is great interest in moving from single-domain systems that operate on a small ontology to multi-domain or open-domain dialog systems that can converse about anything in a large knowledge base.

Definition of Domain

- For chatbots, the domain usually means the topic of the conversation (e.g., sports, movies) or is reflected in the dataset that they are trained on (e.g., the Ubuntu dialog corpus, the Twitter corpus).
 - Due to the nature of chit-chat conversations, such chatbots are usually viewed as open-domain systems.
 - Most existing chatbots either do not rely on **web-scale** knowledge or use off-the-shelf APIs for question answering which may have access to an external knowledge base.

Dialog Initiative

- **System-Initiative**

- Early task-oriented systems (e.g., the AT&T How May I Help You system) mostly use the system-initiative design.
- A form-filling process completely controlled by the system.
- The user simply responds to provide requested information.
- In this design, the system needs to manage how the conversation should move forward, but the management is relatively easy since the user actions are constrained and only allowed at certain points in the dialog.

Dialog Initiative

- **User-Initiative**
 - Many conversational assistants uses the user-initiative design.
 - Relies on the user to control the conversation flow.
 - The user directs the system to provide information or complete a task.
 - The system passively responds to user requests but does not proactively push the conversation forward.
 - Dialog control is mainly related to task clarification and/or error correction.
 - Poses challenges to language understanding to cover a wide variety of user utterances.

Dialog Initiative

- Mixed-Initiative
 - Some task-oriented systems (e.g., the AT&T Communicator system) start to use the mixed-initiative design.
 - The initiative shifts back and forth between the system and user.
 - A more natural way of communication.
 - Brings challenges to both language understanding and dialog control.

Socialbot (The Alexa Prize Challenge)

- A socialbot is required to hold a coherent and engaging conversation with human users on current events and popular topics
 - Non-task-oriented, open-domain, mixture initiative dialog system
- Challenges:
 - Voice-based interaction:
 - Understanding the continuous human speech
 - Generating natural speech response
 - Coverage of wide range of topics:
 - Extracting interesting content
 - Representing the content for dialog control

Technologies Behind Conversational AI

- Speech:
 - Automatic Speech Recognition (ASR): Human speech to text
 - Speech Synthesis: Text to human-sound speech
- Text:
 - Natural Language Understanding (NLU): Unstructured text to structured information
 - Intent classification, sentiment analysis, ...
 - Natural Language Generation (NLG): Structured information to natural language
- Dialog and Discourse:
 - Dialog Management (DM): Control the dialog flow
- Natural Language Processing (NLP) is a sub-field of AI concerning human-computer interactions involving natural language



Sounding Board

A User-centric and Content-driven Social Chatbot

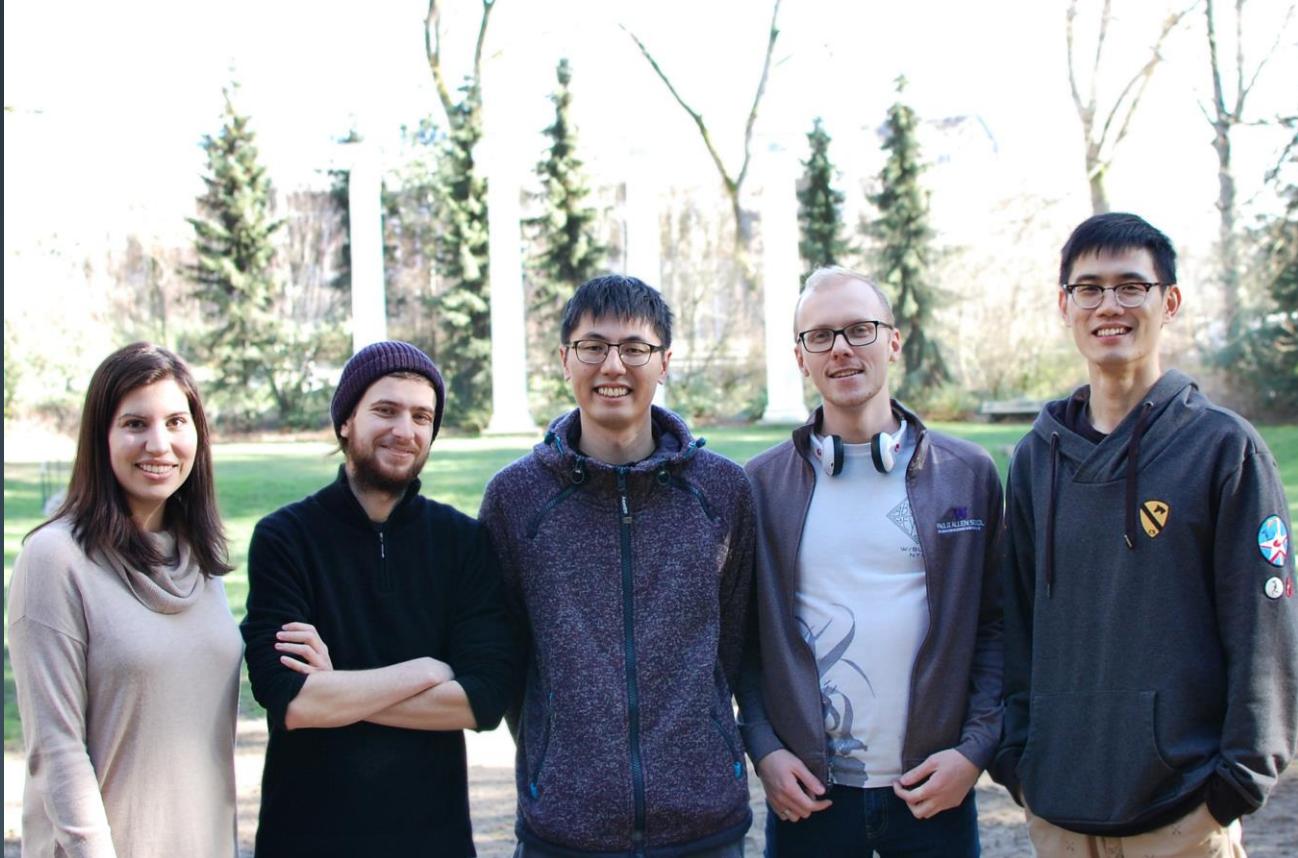
2017 Amazon Alexa Prize Winner

University of Washington

The Sounding Board Team

Students (left-to-right)

Elizabeth Clark, Ariel Holtzman, Hao Fang, Maarten Sap, Hao Cheng



Faculty Advisors



Mari Ostendorf



Yejin Choi



Noah Smith

Agenda

- Amazon Alexa Prize
- Our Approach
- Evaluation
- Future Work

Agenda

- Amazon Alexa Prize
- Our Approach
- Evaluation
- Future Work

Amazon Alexa Prize

Advance conversational Artificial Intelligence (AI)



“converse coherently and engagingly with humans on popular topics and current events”

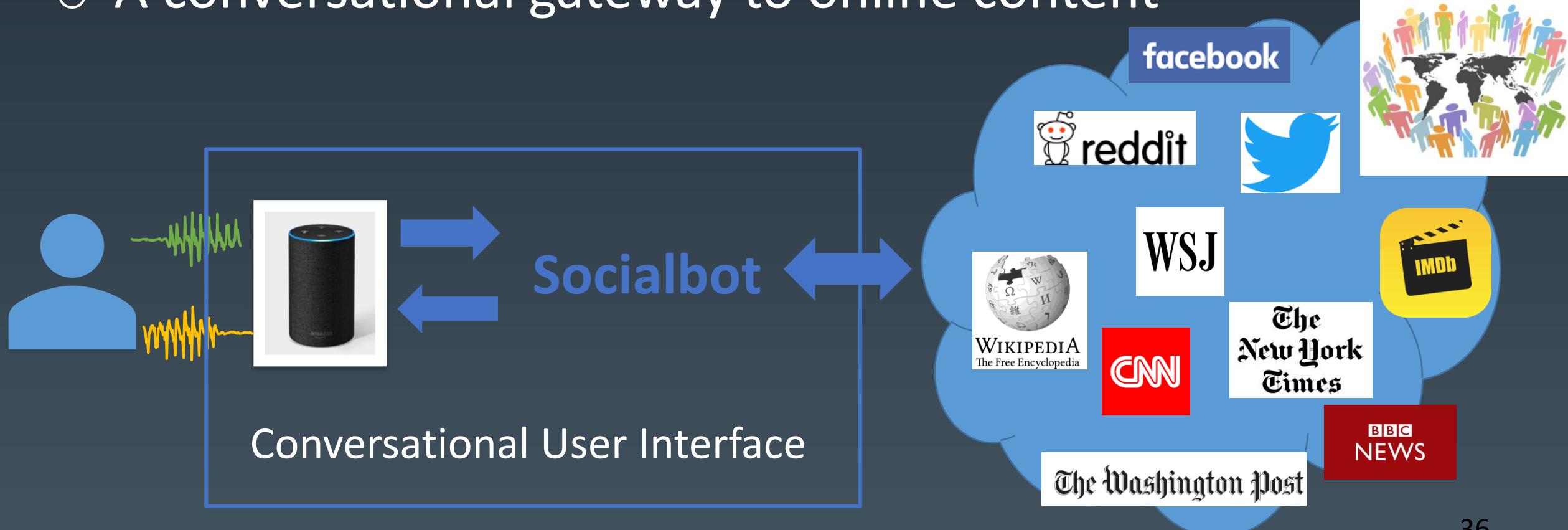
- Competition for university students to develop a socialbot
- 2017 Competition attracts 100+ teams applied, 16 teams participated, 3 made it to the finals
 - University of Washington
 - Czech Technical University
 - Heriot-Watt University
- Anyone with an Alexa-enable device can talk to it by saying “*let's chat*”
 - Most recent 3 months (post-finals): 500K users, 10M interactions

Agenda

- Amazon Alexa Prize
- Our Approach
- Evaluation
- Future Work

Socialbot Applications

- Entertainment, education, healthcare, companionship, ...
- A conversational gateway to online content



Design Objectives

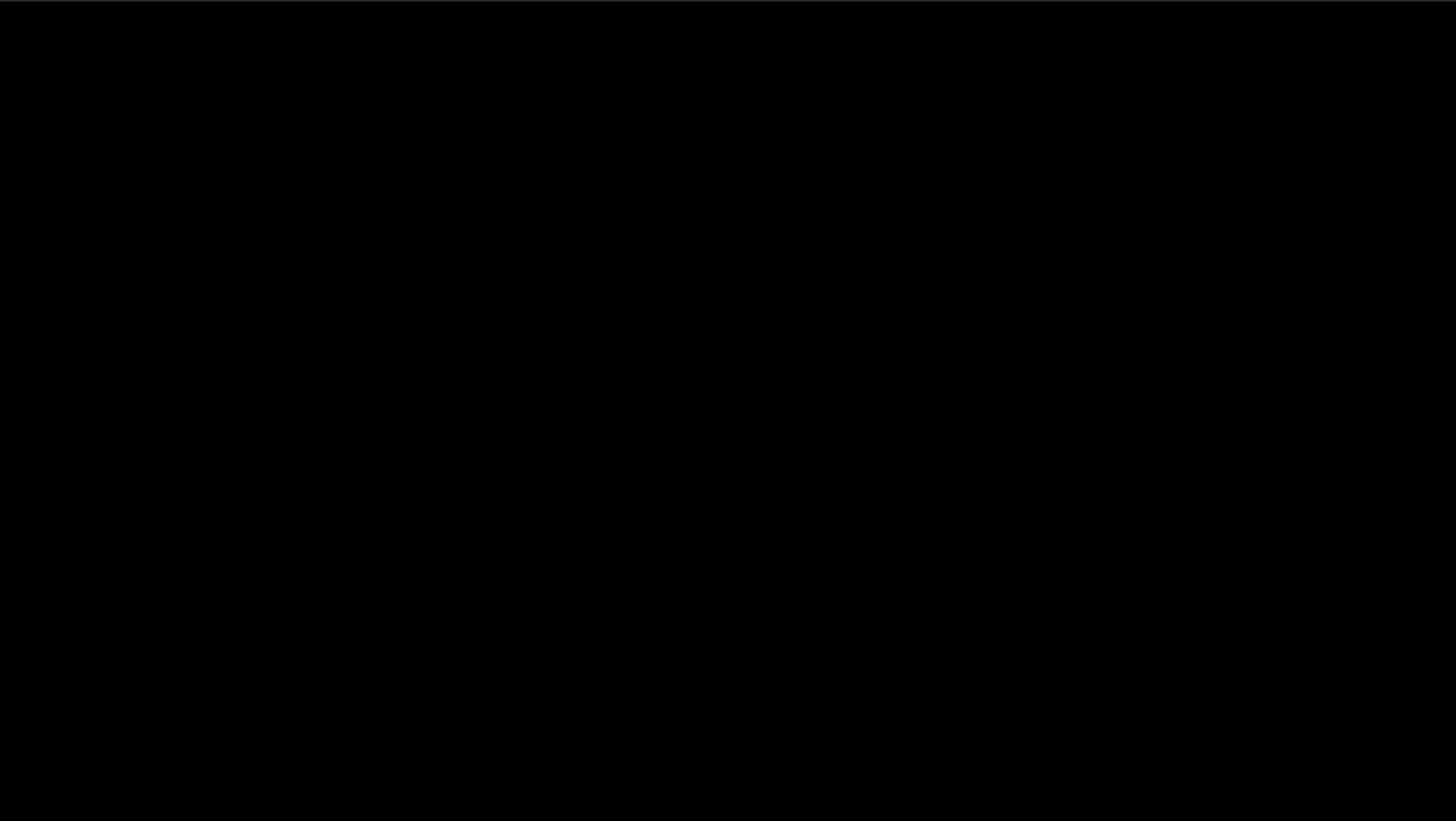
User-Centric

Content-Driven

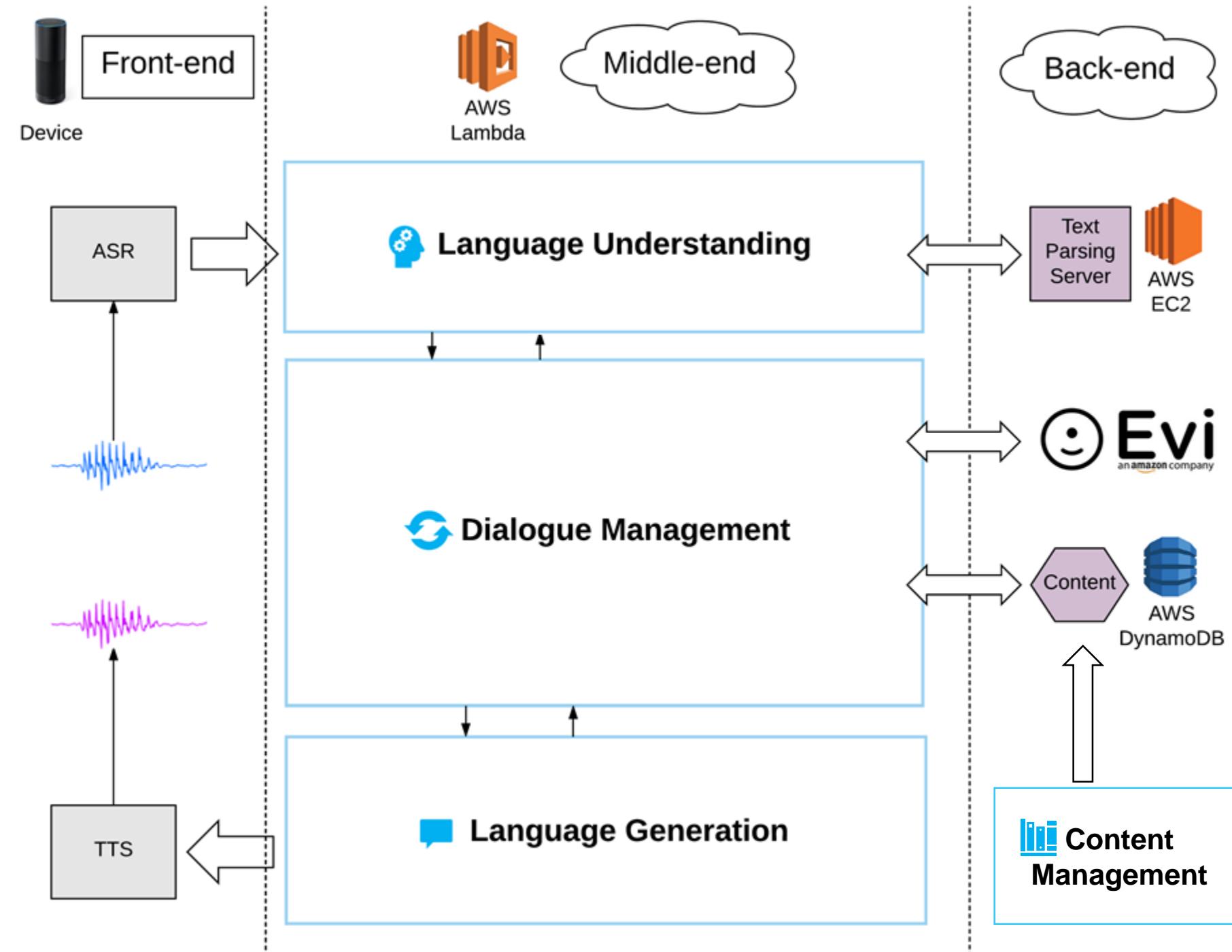
- Users can control the dialog flow and switch topics at any time
- Bot responses are adapted to acknowledge user reactions

- Content cover the wide range of user interests
- Dialog strategies to lead or contribute to the dialog flow

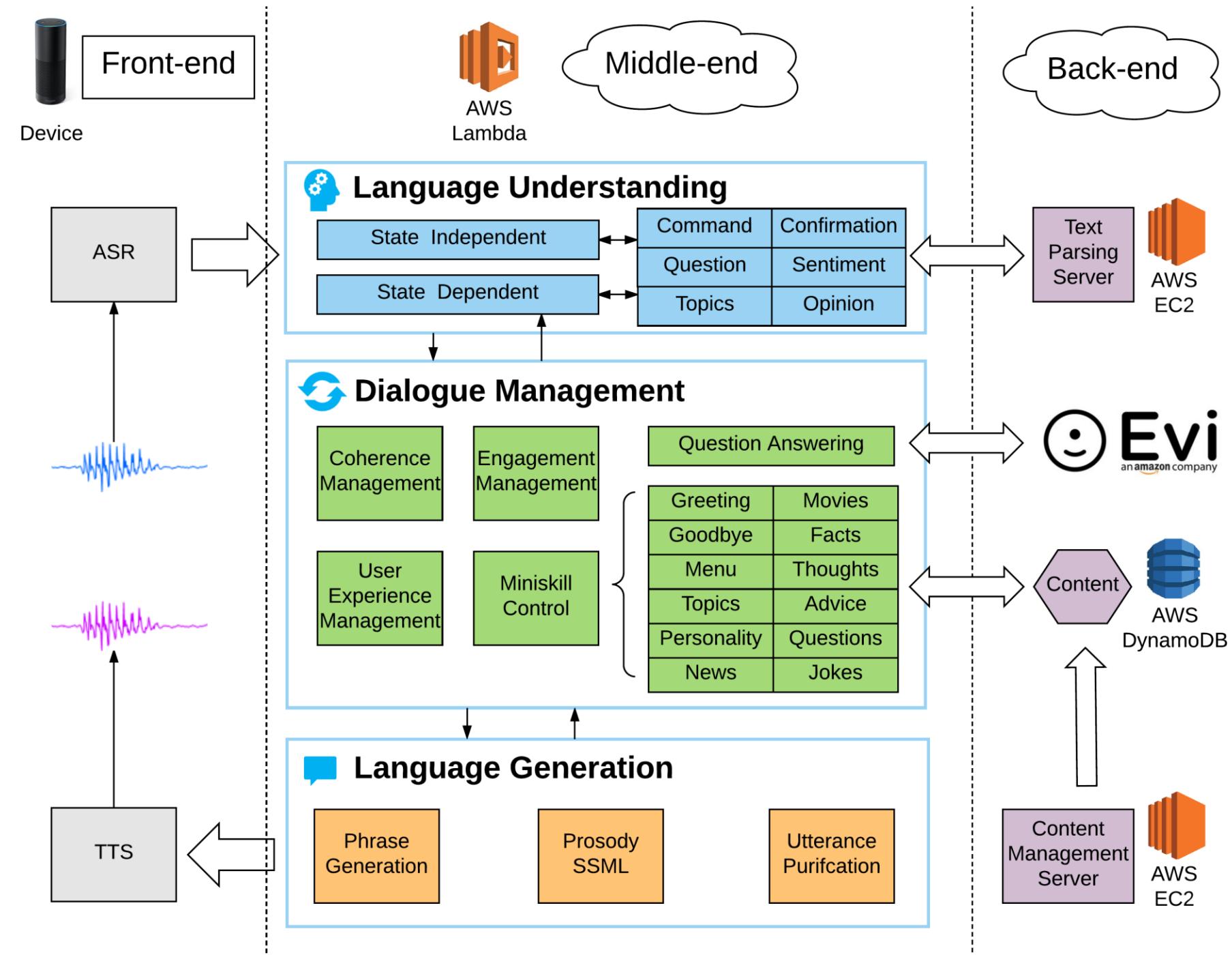
2017 Alexa Prize Finals



System Diagram



System Diagram



Turn	User	Sounding Board
1	Let's chat	Hi! This is an Alexa Prize socialbot. How's your day?
2	I'm <u>five</u> .	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.

ASR error

Negotiation

Grounding

Thoughts

Movie

Fact

Agenda

- Amazon Alexa Prize
- Our Approach
- Evaluation
- Future Work

Evaluation

... for system development and Amazon analysis

- Interaction Ratings from Alexa users
 - 1 – 5 ratings, how much they would like to speak with this Socialbot again
 - Initial development period (April – June), Alexa Prize semifinals (July – Aug) , post semi-finals (Sept – Oct) and ongoing
- Amazon Analysis
 - Hand-labeling of conversations for various attributes (Venkatesh et al. NIPS 2017)
- Alexa Prize Finals
 - 1 – 5 ratings from Amazon-selected judges, with time as a tie-breaker

Evaluation

Alexa Prize Finals (Nov. 2017)

- Finalists: Czech Technical Univ., Univ. of Washington, Heriot-Watt Univ.
- 3 interactors, 3 judges
 - Interactors engage with the Socialbot in turn-by-turn conversation about popular topics
 - Judges listen to the interaction and give ratings (1 to 5)
 - Interactors end the conversation when 2 judges have requested to stop.
- Sounding Board results:
 - Average score: 3.17
 - Average duration: 10 minutes and 22 seconds
 - *“Two judges had clicked out just shy of the 20-minute mark.”* – James Vlahos, WIRED.

Agenda

- Amazon Alexa Prize
- Our Approach
- Evaluation
- Future Work

Future Work

Sounding Board faces and tries to address ...

Content Challenges

- Collect and understand online content
- Recommend content effectively and efficiently
- Present content through conversational interaction
- Exclude undesirable content

User Experience Challenges

- Understand user intent and reaction
- Keep the conversation grounded
- Deflect discussions about taboo topics

Future Work

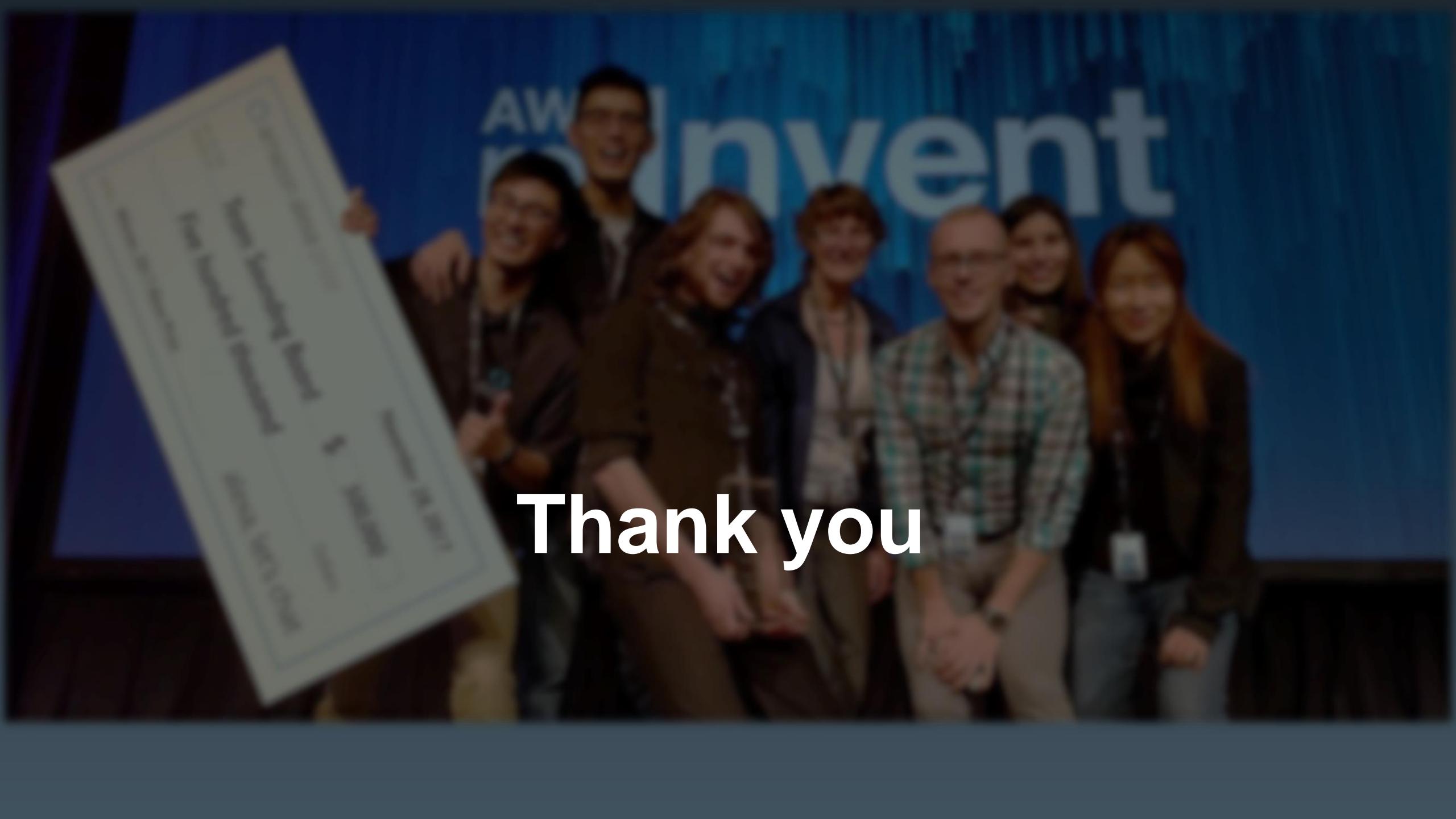
Research directions

- Personalized understanding and context carrying
 - Conversation history reveals user topic preference, like topic A -> like topic B
- Deeper discussion about news events and online content
 - Document-level knowledge graph and web-scale ontology
- Automatic evaluation and diagnosis of human-bot conversations
 - Not all good conversations are equally good
 - Human judgement has high-variance and is not informative enough
- The role of machine learning algorithms for system design
 - Human-in-the-loop learning paradigm and data annotation

Future Work

Beyond the conversational gateway to social media...

- Information-oriented (timely content important)
 - Conversational gateway to information for education
 - Voice-based help-desk information service
 - Debate/info-bot (e.g. for political campaigns)
- Chat-oriented (social interaction important)
 - Conversational “practice” for language learning
 - Companionship
 - Medical domains (diagnosis & therapeutics)
- Art exhibit

A group of seven people are standing behind a whiteboard in a presentation setting. The whiteboard has some faint, illegible text written on it. The background features a blue wall with the words "Award Invent" partially visible. A large, semi-transparent white rectangular box covers the bottom half of the image, containing the text "Thank you".

Thank you