

Graph Learning for AR/VR Intelligent Assistants

Xin Luna Dong, Meta Reality Lab

9/2022

This talk does not represent the company's point of view

Everyone Deserves An Assistant



What is A Virtual Intelligent Assistant?

Respond to commands



“Hey Siri, set a timer to 7pm”

“Ok, added to today’s reminders”

What is A Virtual Intelligent Assistant?

Control devices



“Hey Alexa, turn off bedroom lights”



What is A Virtual Intelligent Assistant?

Provide information



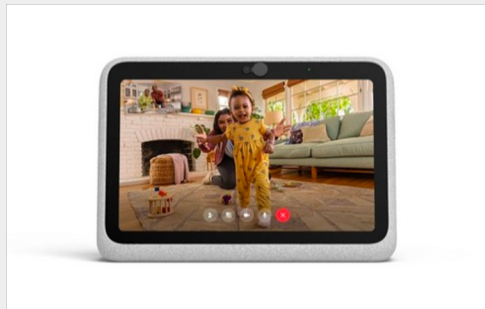
“Hey, Google, when does summer end?”

“Summer ends on Thursday, Sept 22”

Meta's Assistant

Empowering connection to people and experiences in your life

Facebook Portal



“Hey Portal”

- Hop on a call hands-free
- Get help with music, timers, alarms, weather, show photos from your Facebook profile, and more.

Meta Quest 2



“Hey Facebook” (double press the button on your controller)

“Who’s online?”--meet up with friends

“Open Beat Saber”--jump straight in the game, and more.

Ray-Ban Stories



“Hey Facebook, take a picture” -- capture moments hands-free

“Hey Facebook”--call friends on Messenger, manage device settings, and more.

What is An Ideal Virtual Intelligent Assistant?

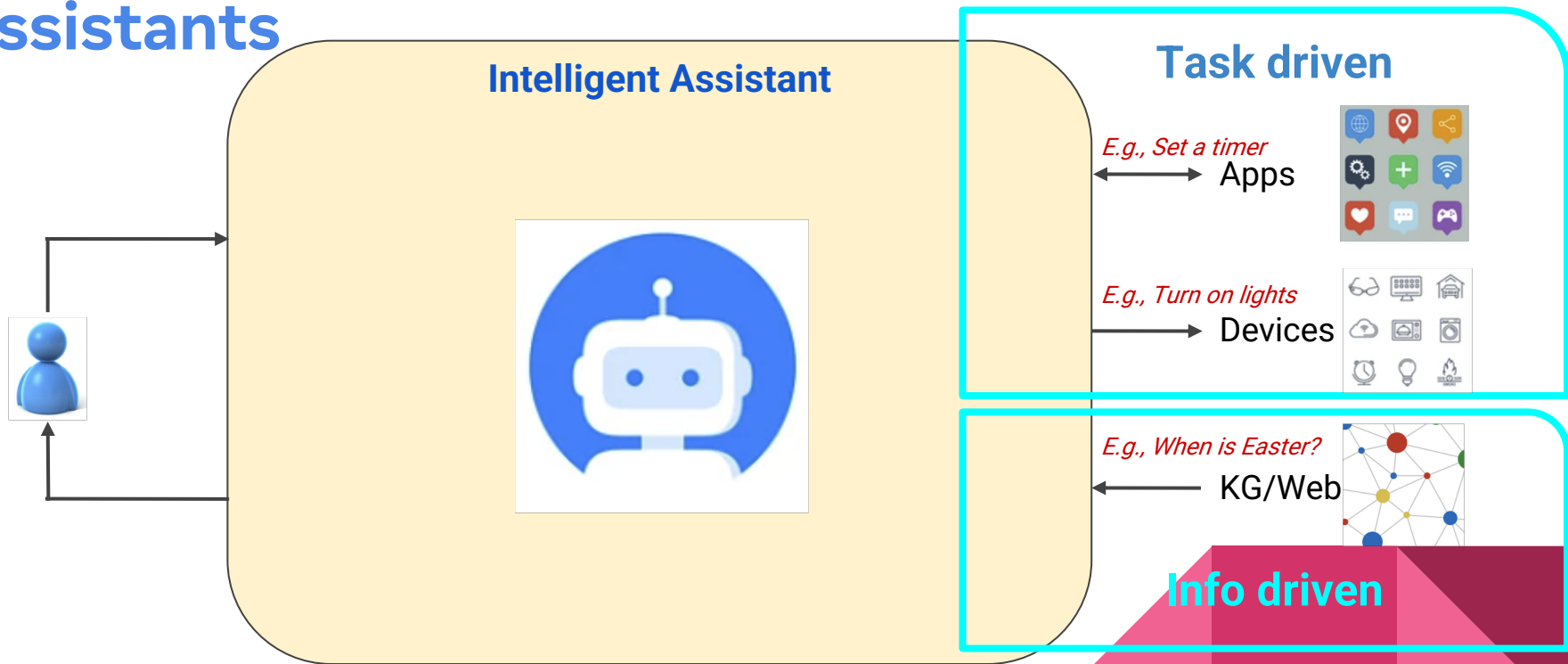
An *intelligent assistant* should be an agent that **knows you and the world**, can **receive your requests** or **predict your needs**, and provide you the **right services at the right time** with your permission.



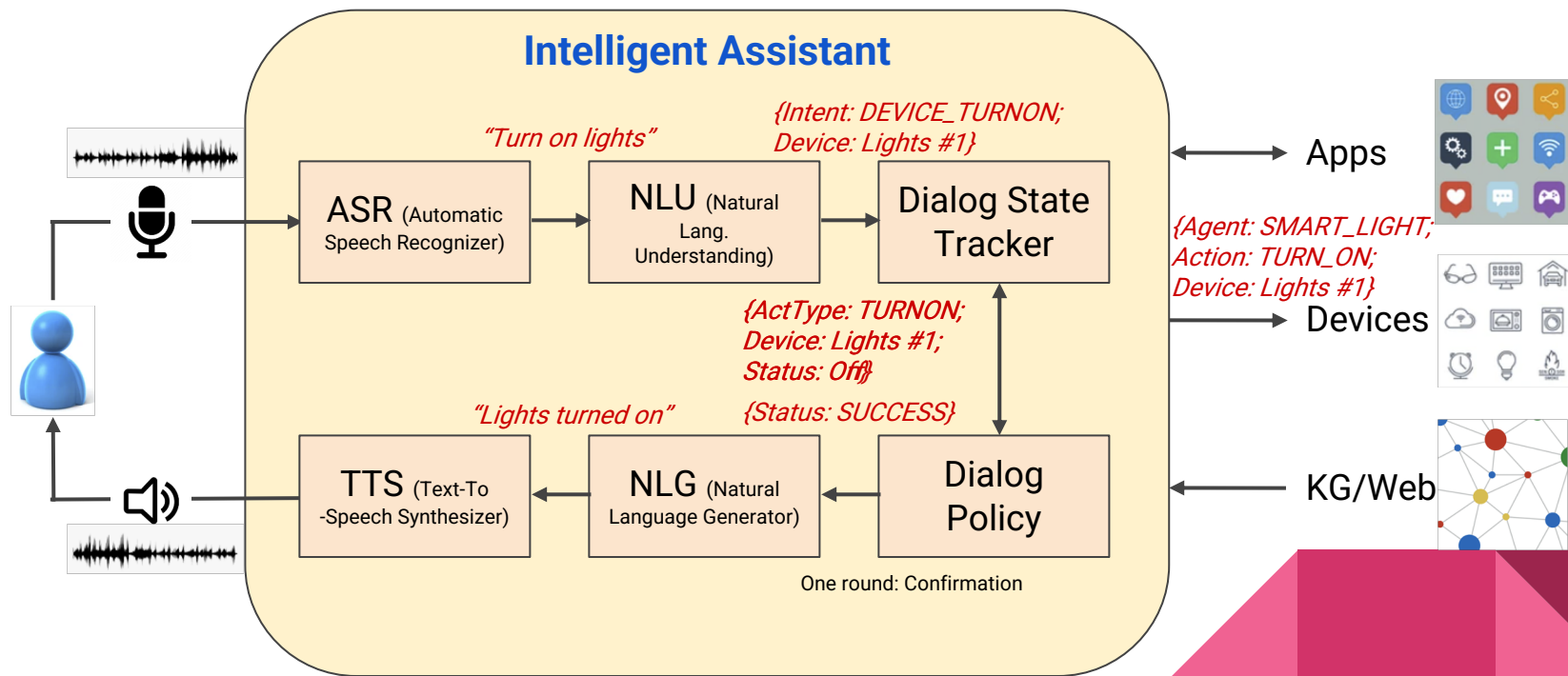


Technologies to Support Current Intelligent Assistants

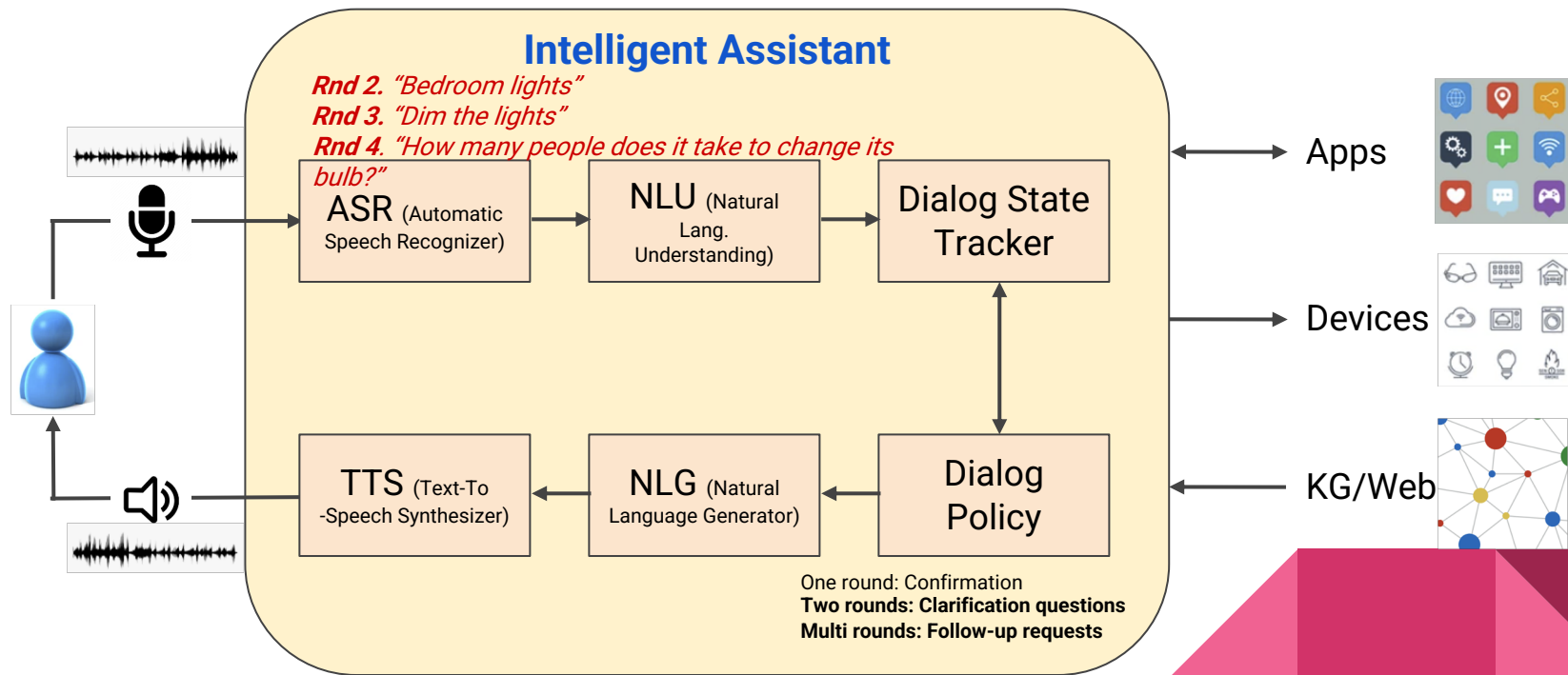
Two Types of Commands to Intelligent Assistants



Int. Asst. Is Essentially a Conversation System



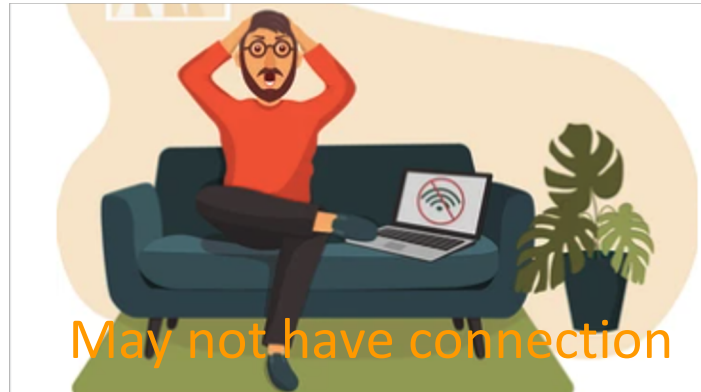
Int. Asst. Is Essentially a Conversation System





Challenges and Initial Solutions to AR/VR Assistants

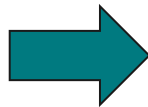
What Is Different for An AR/VR Assistant?



From Voice-Only to Multi-Modal



“How tall is Empire State Building?”

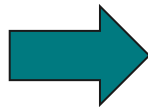


“What’s the name of this building and how tall is it?”

From Context-Agnostic to Context-Aware



"Show my shopping list"

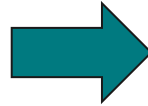


"Remember to buy apples and bananas at the grocery store around the corner"

From Reactive to Proactive

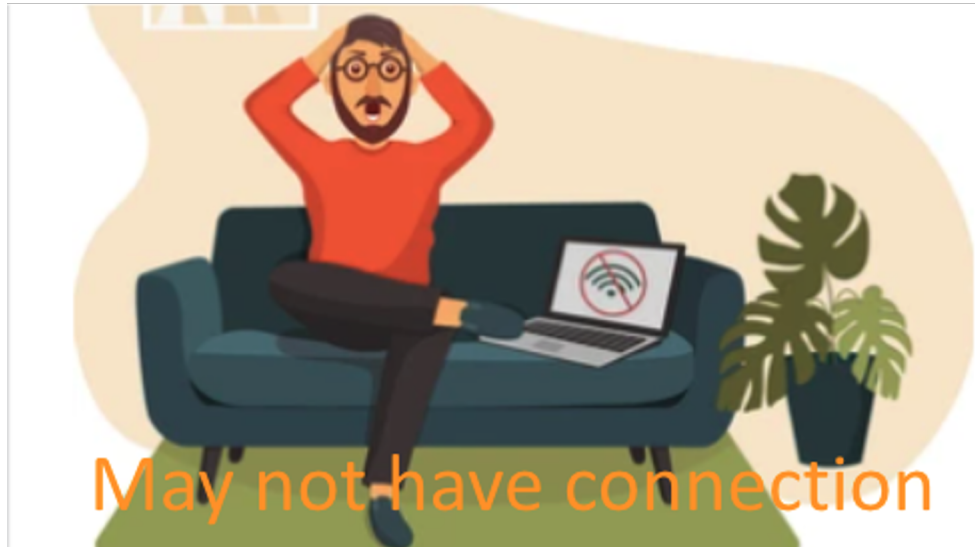


"What's the weather today?"



"Today is sunny, 70 degree. Would you like to play your favorite morning music?"

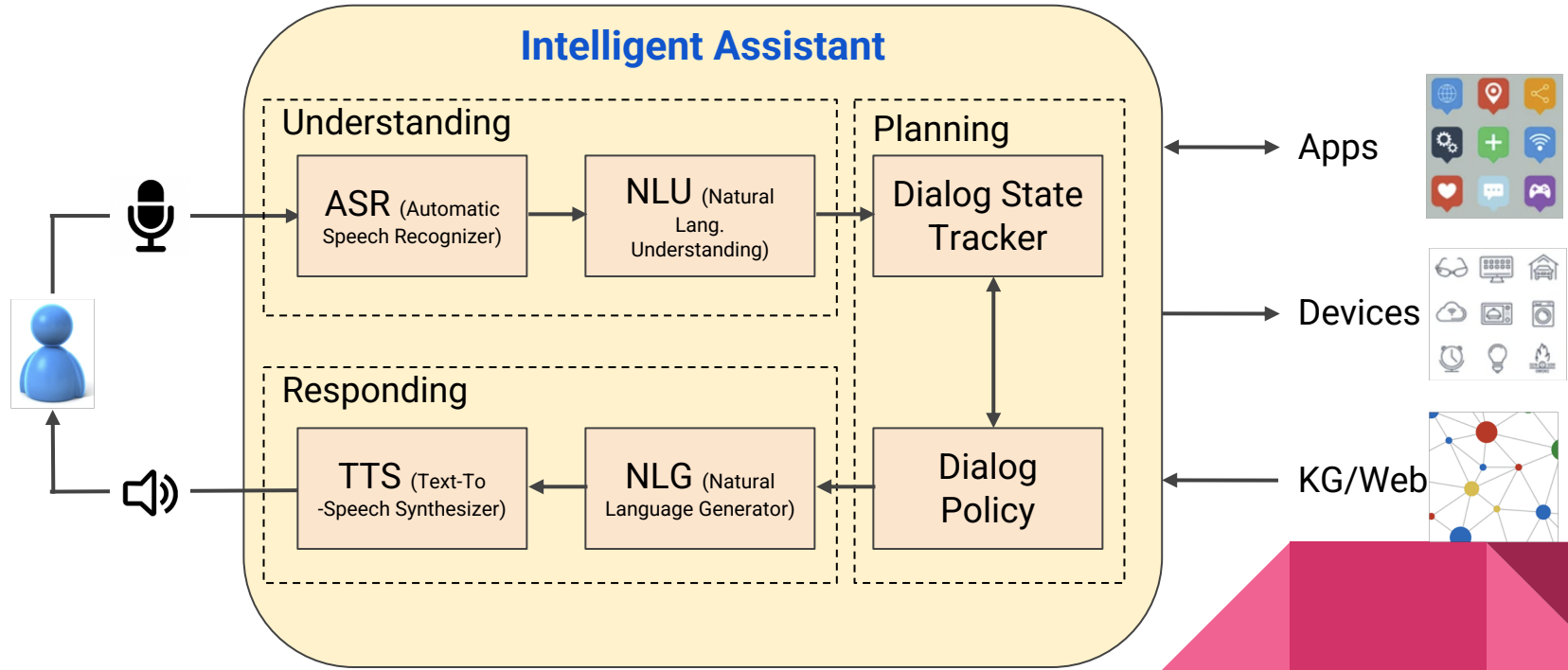
From Server-Side to On-Device



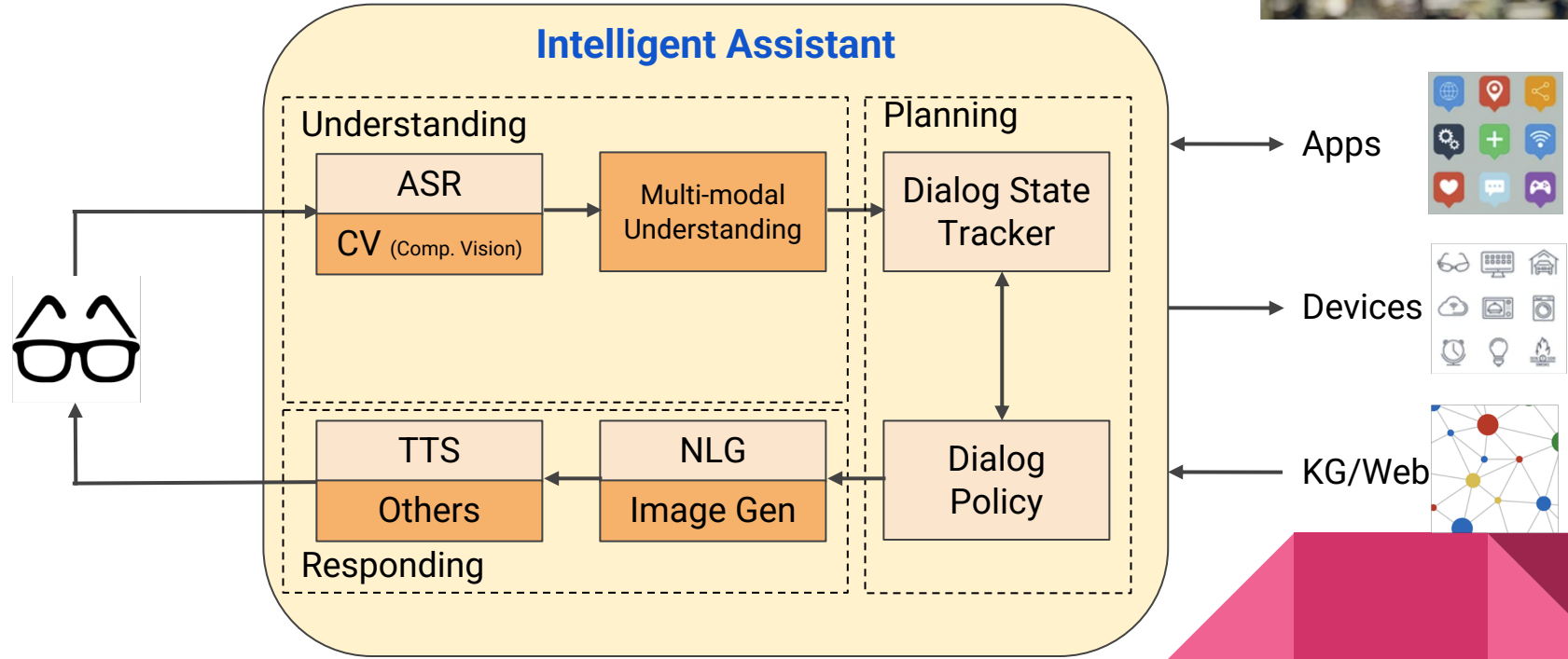
May not have connection

+ Privacy!!!

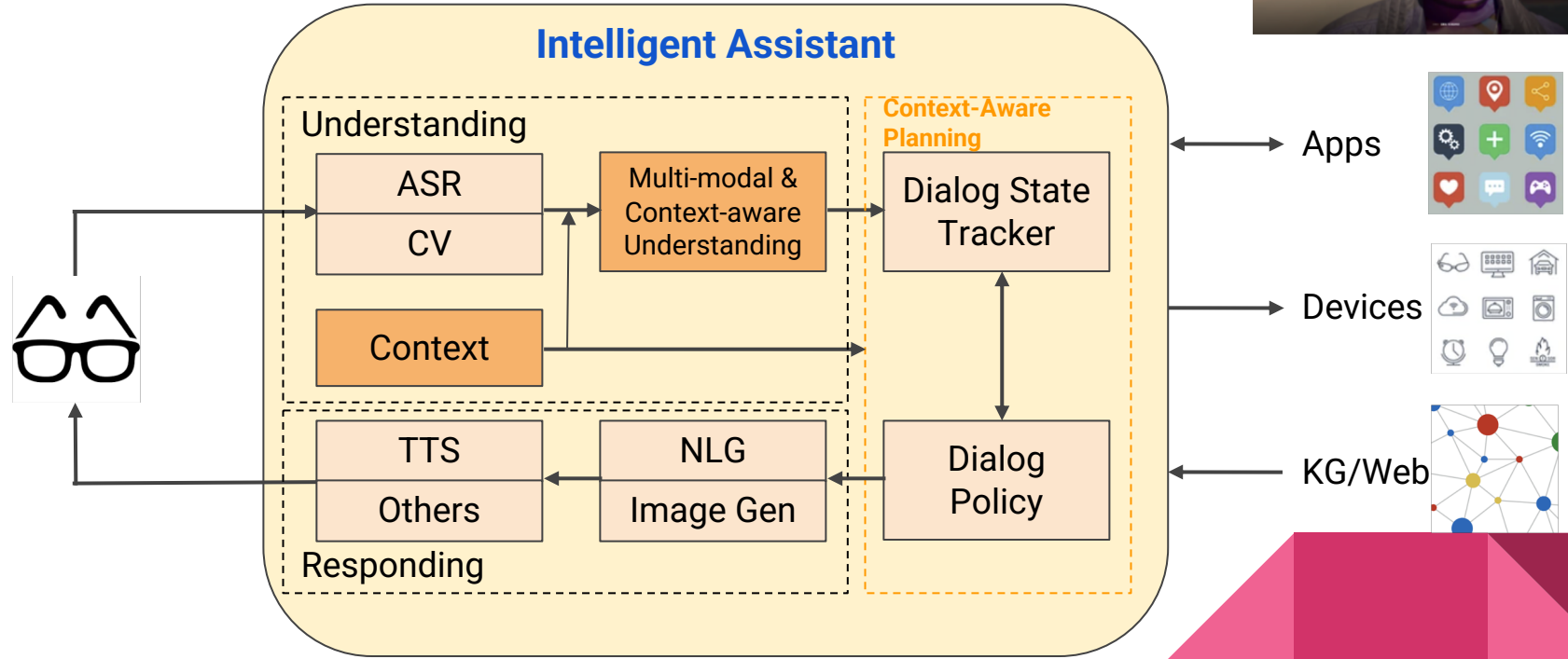
MultiModal, Context-Aware Assistant



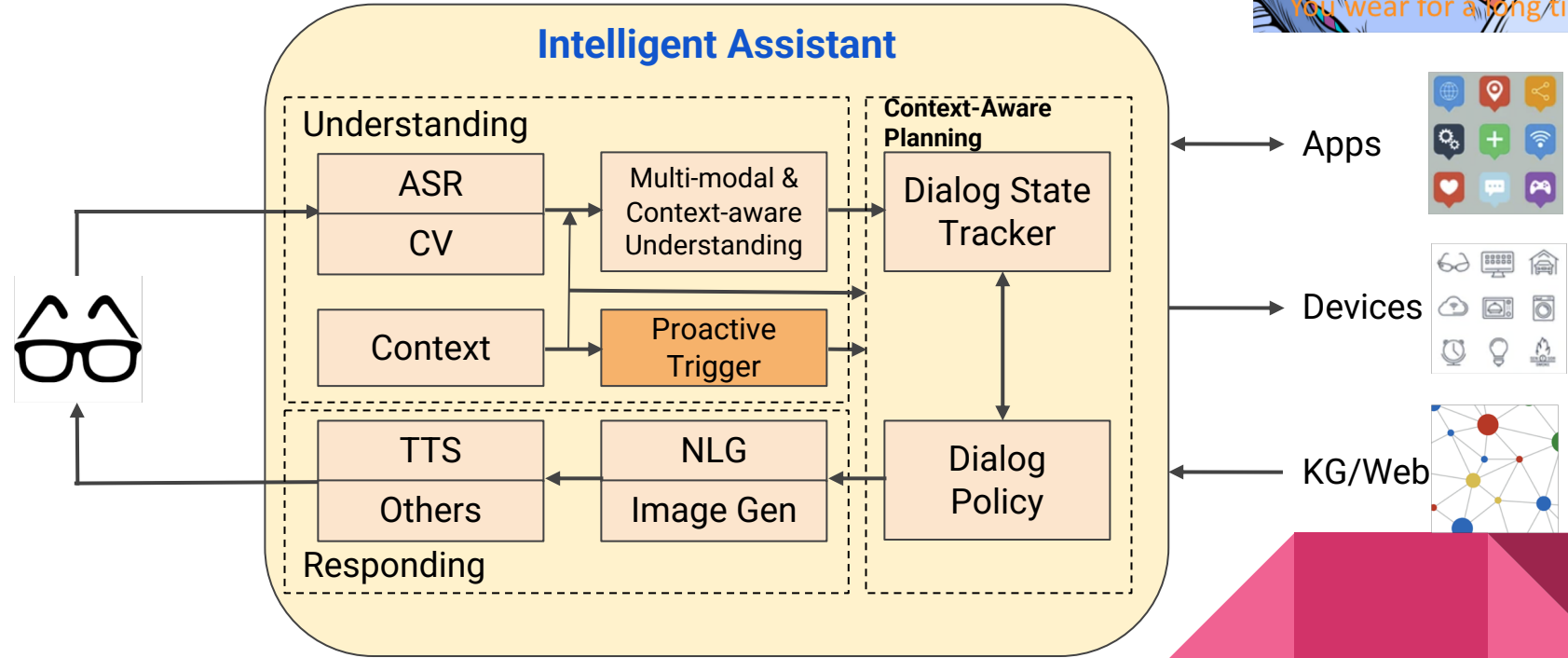
MultiModal, Context-Aware Assistant



MultiModal, Context-Aware Assistant



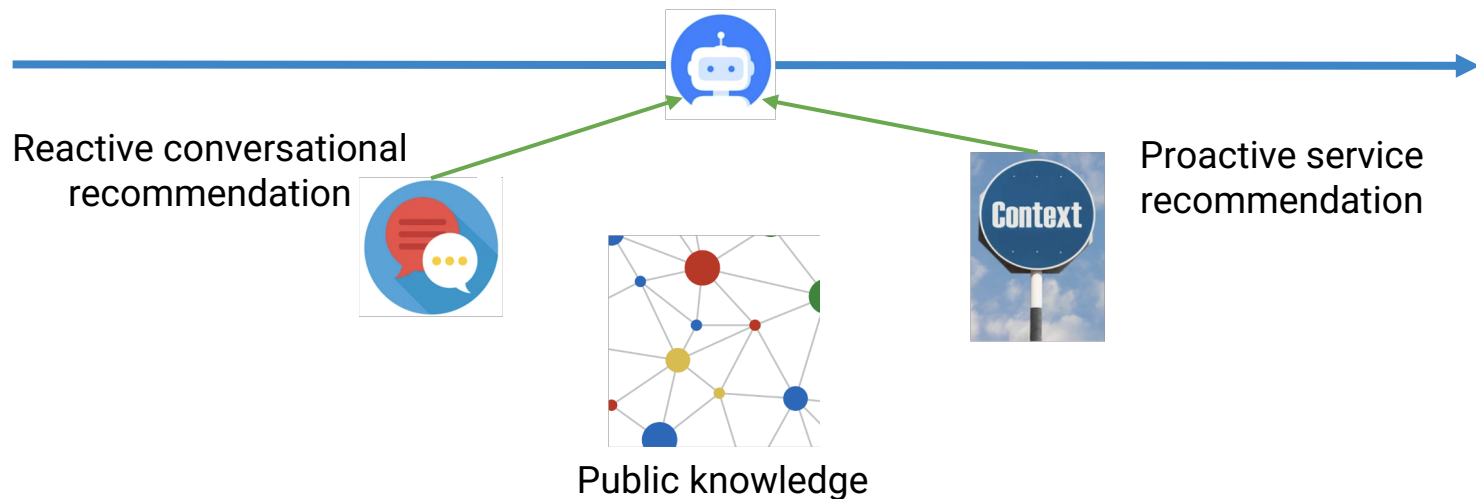
MultiModal, Context-Aware Assistant





Roles of Graph Learning in Enabling AR/VR Assistants

Assistant Recommendation



Challenges: Leveraging vast volume of public knowledge for reactive recommendation w/o sacrificing *latency*

4-1. Conv. Recom. w. Public KG

(a) Dialog

1 Can you recommend any classic books like *Catcher in the Rye*?

2 Do you prefer books by the same *author* or same *genre*?

3 I am interested in reading classic examples of *American literature*.

4 *Literary realism* is a common genre in classic *American literature*.

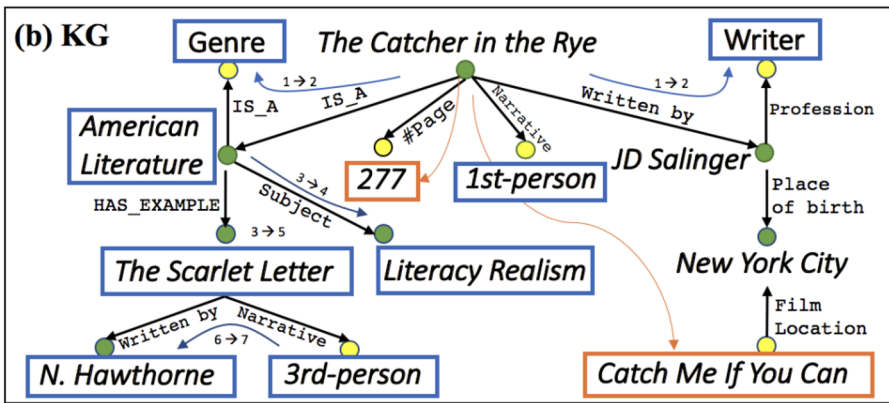
5 Do you prefer *First-person* or *Third-person narrative*?

6 I mostly prefer *third-person narrative*.

7 Consider reading the *Scarlet Letter*: a novel by *Nathaniel Hawthorne*.

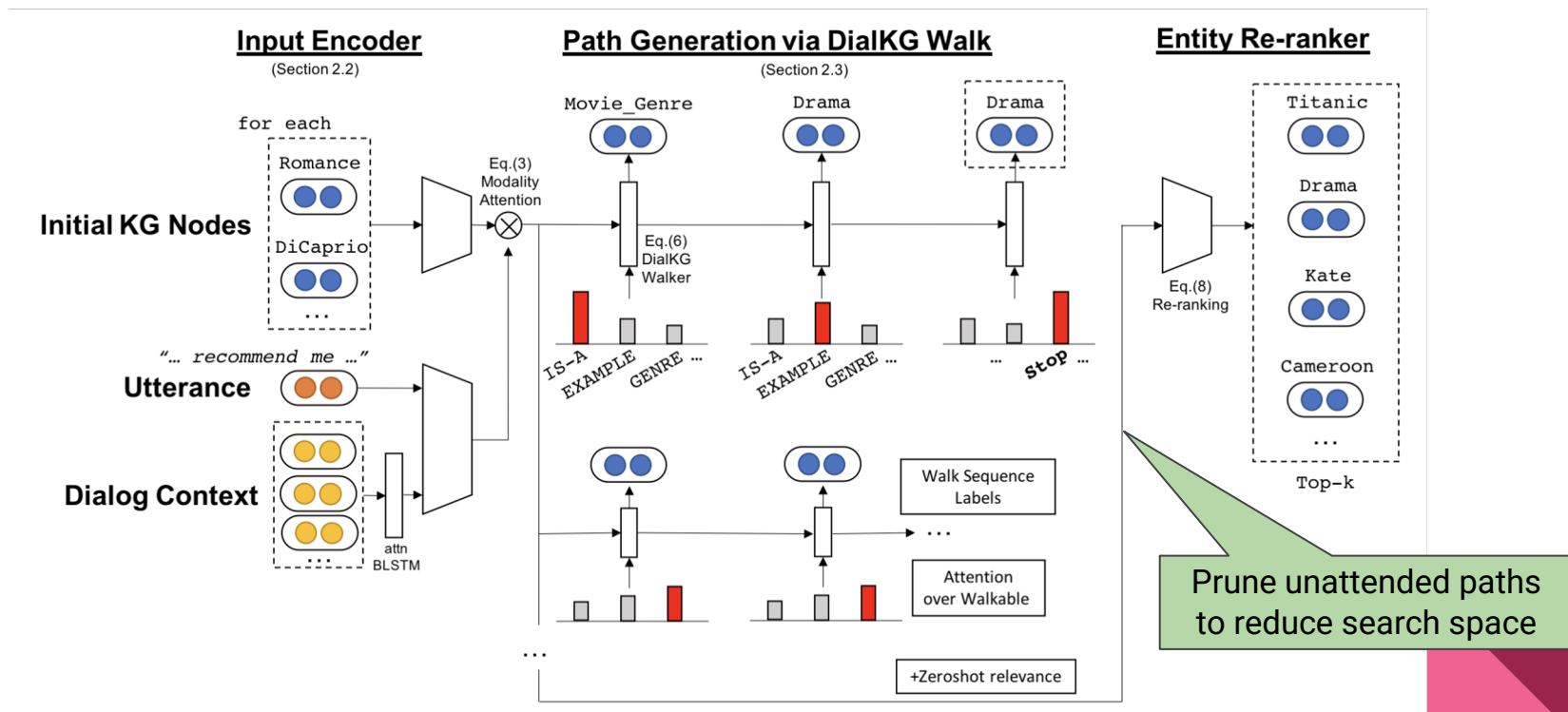
⋮

(b) KG

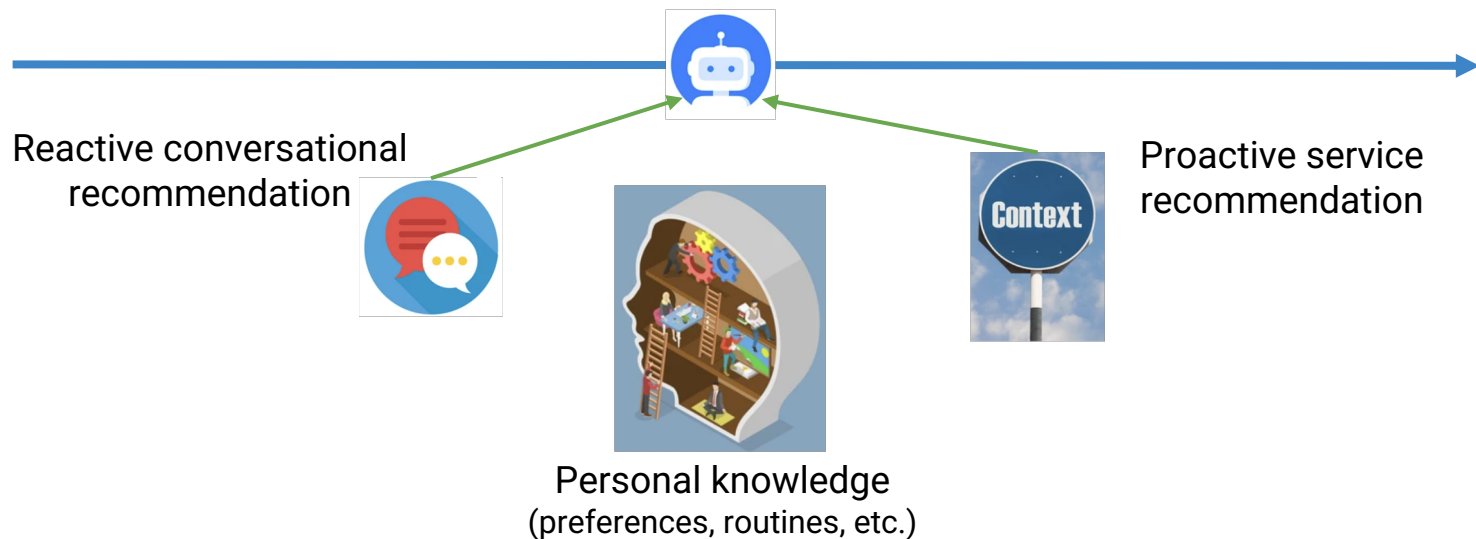


Find paths to follow when making suggestions or providing information

4-1. Conv. Recom. w. Public KG



Assistant Recommendation (Cont')



Challenges: Leveraging personal action log, preferences, routines, etc., to improve context-aware recommendation w/o sacrificing *privacy*.

4-2. Conv. Recom. w. Public+Personal KG

(a) Conversational Recommendations

1 Hello, I'm looking for a good place to eat.

2 Is Downtown Phoenix a good place to start your search as you've been there a few times?

3 Yes, I'd like something there please.

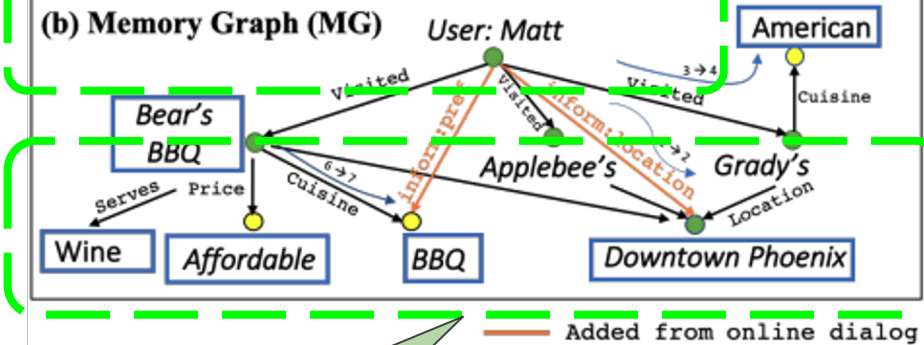
4 You ordered *American* food a few times, do you want something similar or feeling adventurous today?

6 No, I'm in the mood for a *BBQ* today.

7 *Bear's BBQ* got some *affordable* yet great *BBQ*. They also serve *wine*.

⋮

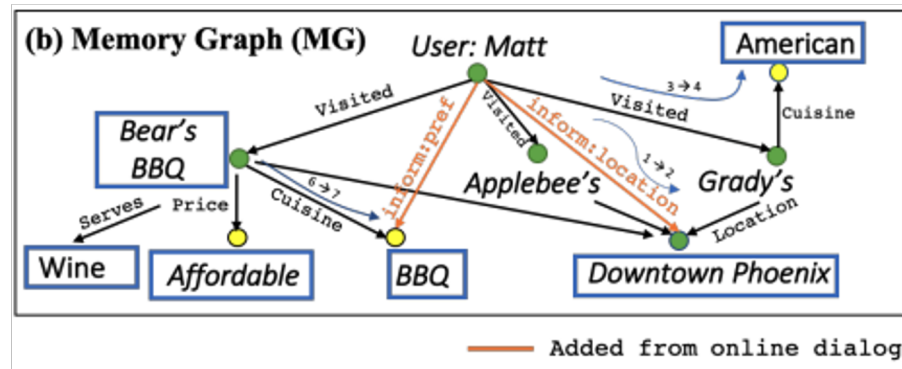
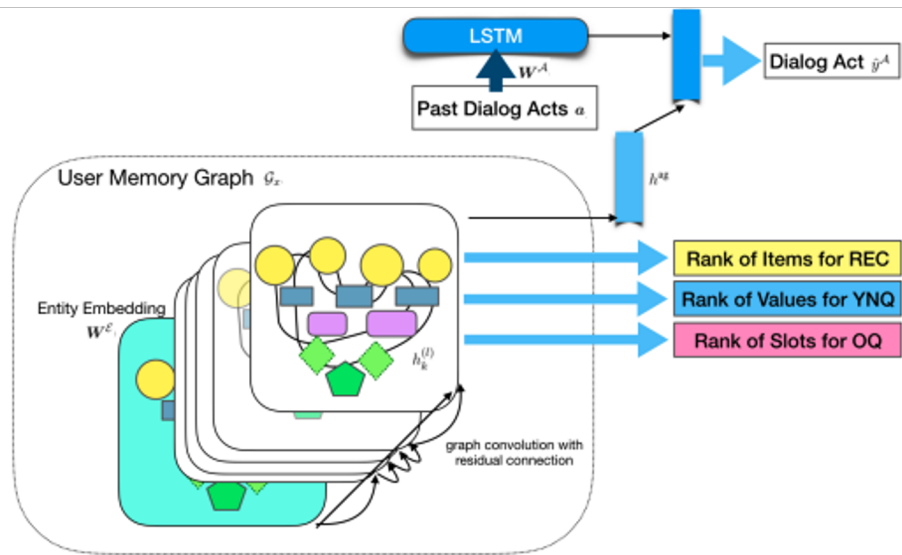
(b) Memory Graph (MG)



Personal KG

Public KG

4-2. Conv. Recom. w. Public+Personal KG



4-2. Conv. Recom. w. Public+Personal KG

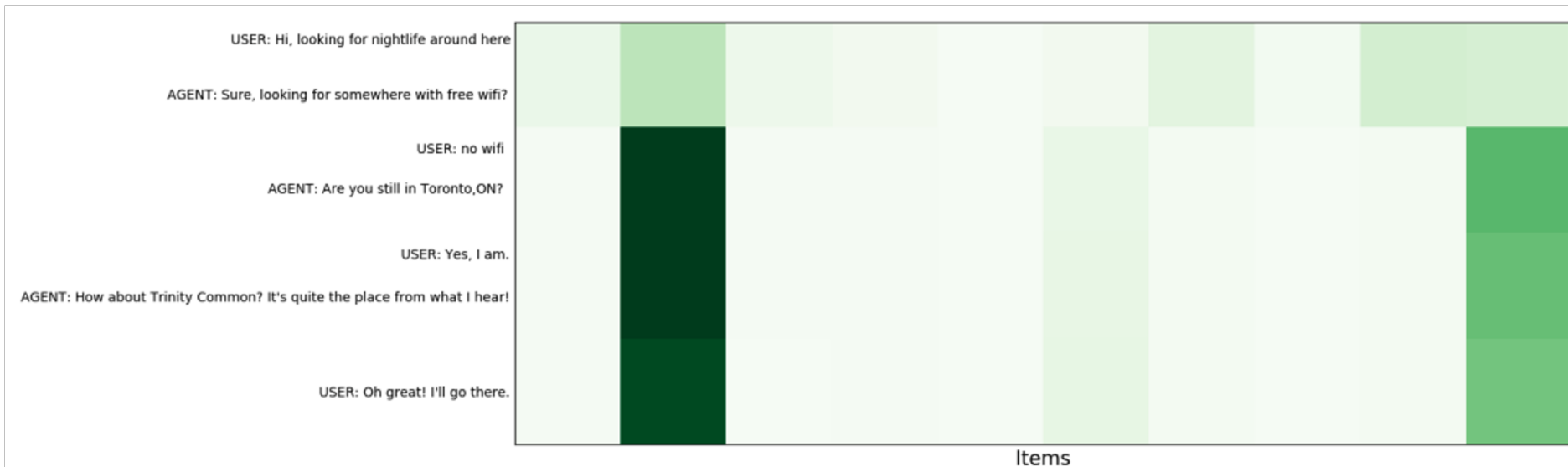


Figure 5: Visualization of item-level conversational reasoning, given an example dialog. Darker color indicates more salient items for recommendation at each given turn (row), predicted by our UMGR model.

4-3. Federated Learning

- ~~Push data to model~~ → Push models to data

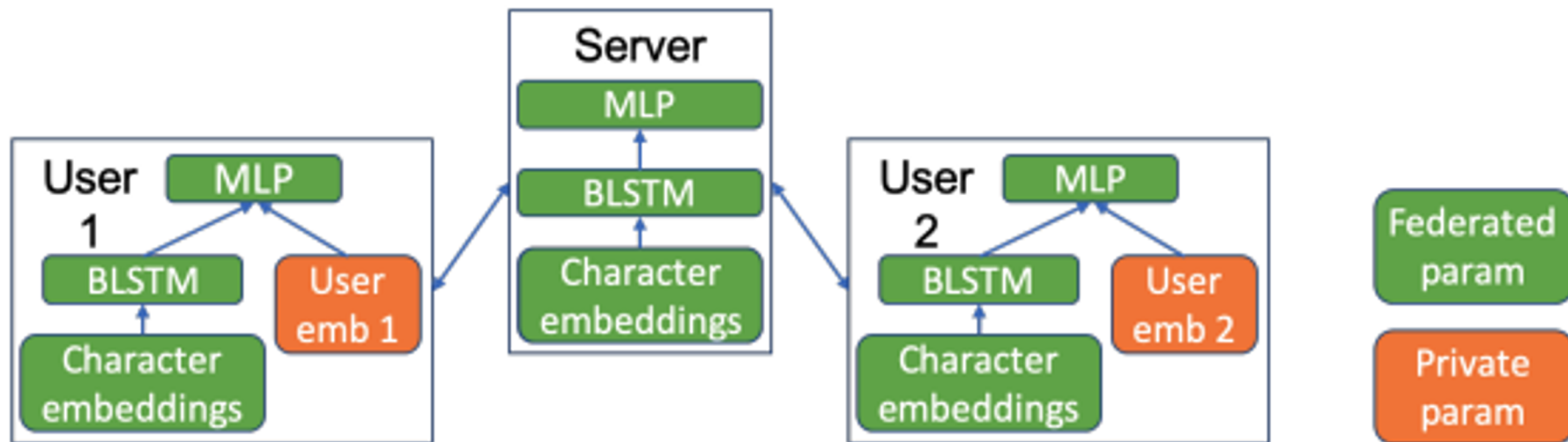
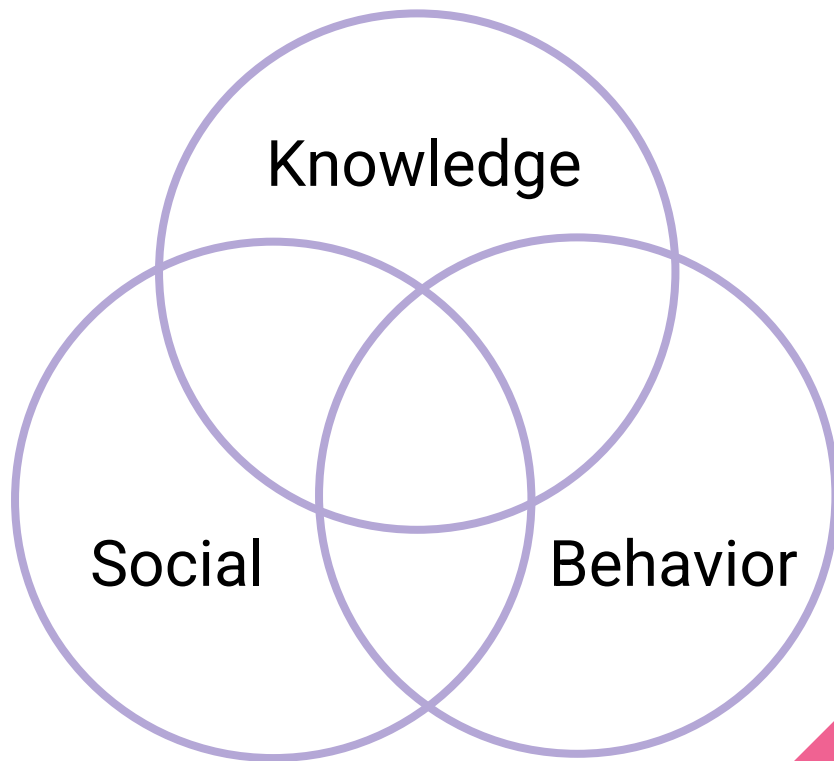


Figure 1: Personalized Document Model in FL.

Future of Graph Learning for AR/VR Assistant

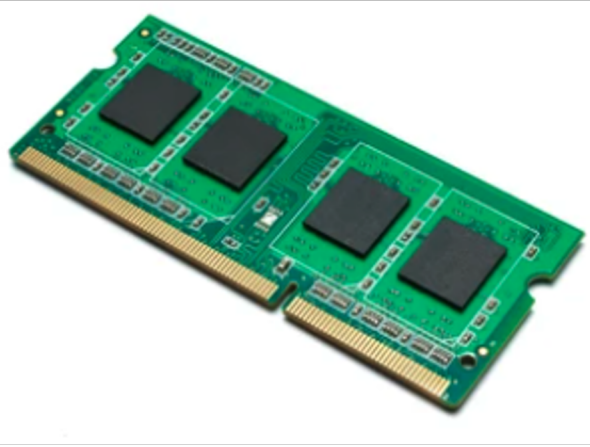
Multi-modal
Contextual AI
Proactive



Runtime
On-device
Privacy



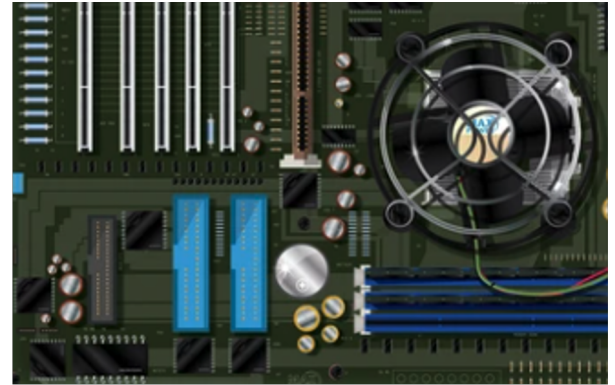
My Everyday Worries When Working on Devices



Memory




Battery



Thermal

Take-Aways

- An intelligent assistant should be an agent that *knows you and the world*, can *receive your requests* or *predict your needs*, and provide you *the right services at the right time* with your permission
 - An intelligent assistant is essentially a **conversation system**, *task-driven or information-driven*
 - Next-generation AR/VR assistants require new research on graph learning to support multi-modal & context-aware recommendations.
- 



Thank You

Q&A?